

L'EDI vu par Wikipedia.fr

Transfert et transformation

Table des matières

1	Introduction	1
1.1	Échange de données informatisé	1
1.1.1	Définition	1
1.1.2	Quelques organisations	1
1.1.3	Quelques normes EDI	2
1.1.4	Quelques protocoles de communication	2
1.1.5	Quelques messages courants	2
1.1.6	Articles connexes	4
2	Organismes de normalisation	5
2.1	Organization for the Advancement of Structured Information Standards	5
2.1.1	Membres les plus connus de l'OASIS	5
2.1.2	Standards les plus connus édictés par l'OASIS	6
2.1.3	Lien externe	7
2.2	GS1	7
2.2.1	Notes et références	7
2.2.2	Annexes	7
2.3	GS1 France	7
2.3.1	Lien externe	8
2.4	Comité français d'organisation et de normalisation bancaires	8
2.4.1	Liens externes	8
2.5	Groupement pour l'amélioration des liaisons dans l'industrie automobile	9
2.5.1	Présentation	9
2.5.2	Notes et références	9
2.5.3	Voir aussi	9
2.6	VDA	10
2.7	EDI-Optique	10
2.7.1	Voir aussi	10
3	Formats de données	11
3.1	Échange de données informatisées pour l'administration, le commerce et le transport	11
3.1.1	Organisations sectorielles	11
3.1.2	Voir aussi	12

3.2	electronic Business using XML	12
3.2.1	Introduction	12
3.2.2	Historique	13
3.2.3	Travail actuel	13
3.2.4	Vue d'ensemble conceptuelle d'architecture d'ebXML	15
3.2.5	Les secteurs d'activité concernés	16
3.2.6	Utilisations dans le cadre gouvernemental	18
3.2.7	Notes et références	18
3.2.8	Voir aussi	18
3.3	Electronic Data Interchange for the Electric Industry	19
3.3.1	Messages	20
3.3.2	Voir aussi	20
3.3.3	Liens externes	20
3.4	Échange télématique banque-clients	20
3.4.1	Notes et références	21
3.5	Electronic Banking Internet Communication Standard	21
3.5.1	Origines	21
3.5.2	Aspects techniques	21
3.5.3	Versions	22
3.5.4	Notes et références	22
3.5.5	Liens externes	22
3.6	Edifacas	22
3.6.1	Historique	22
3.6.2	EDI, technique de communication	23
3.6.3	Normes UN/CEFACT Edifact	23
3.6.4	Normes UN/CEFACT XML	23
3.6.5	Procédure d'attestation de conformité	23
3.6.6	Liens externes	24
4	Protocoles de communications	25
4.1	Protocole de communication	25
4.1.1	Concept	25
4.1.2	Exemple d'utilisation	25
4.1.3	Voir aussi	26
4.2	Réseau à valeur ajoutée	26
4.3	European Network Exchange (ENX) Association	26
4.3.1	Historique	26
4.3.2	Utilisation du réseau ENX	28
4.3.3	Fonctionnement du réseau ENX	28
4.3.4	Disponibilité mondiale	29
4.3.5	Liens	30
4.3.6	Références	31

4.4	X.25	32
4.4.1	Histoire	33
4.4.2	Couches	33
4.4.3	Notes et références	33
4.5	X.400	34
4.5.1	Liens externes	34
4.6	EDIINT	34
4.6.1	Certification Drummond	34
4.6.2	Autres sources d'information	34
4.7	Simple Mail Transfer Protocol	35
4.7.1	Principes d'envoi	35
4.7.2	Syntaxe type d'une session ESMTP	36
4.7.3	Les codes retour SMTP	36
4.7.4	Sécurité et problème du spam	36
4.7.5	Blocage du port 25 par les fournisseurs d'accès	37
4.7.6	Notes et références	37
4.7.7	Voir aussi	37
4.8	File Transfer Protocol	38
4.8.1	Histoire	39
4.8.2	Interopérabilité	39
4.8.3	Utilisation	39
4.8.4	Implémentations	39
4.8.5	Le protocole	40
4.8.6	Notes et références	43
4.8.7	Voir aussi	43
4.9	File Transfer Protocol Secure	44
4.9.1	Fonctionnement	44
4.9.2	Voir aussi	45
4.10	Secure File Transfer Protocol	45
4.10.1	Voir aussi	45
4.11	SSH File Transfer Protocol	45
4.11.1	Logiciels	46
4.11.2	Voir aussi	46
4.11.3	Liens externes	46
4.12	Applicability Statement 2	47
4.12.1	Avantages	47
4.12.2	Inconvénients	47
4.12.3	Introduction technique d'AS2	47
4.12.4	Voir aussi	47
4.12.5	Implémentations Open Source	48

5.1	Axway	49
5.1.1	Présentation	49
5.1.2	Évolution du chiffre d'affaires	49
5.1.3	Notes et références	50
5.1.4	Voir aussi	50
5.2	Generix	50
5.2.1	Historique	51
5.2.2	Logiciels	51
5.2.3	Notes et références	52
5.2.4	Lien externe	52
5.3	International Business Machines	52
5.3.1	Historique	52
5.3.2	Activités	54
5.3.3	Culture d'entreprise	61
5.3.4	Présidence de la société	62
5.3.5	Logo	64
5.3.6	Principaux actionnaires	65
5.3.7	Résultats financiers	65
5.3.8	Critiques	66
5.3.9	Notes et références	66
5.3.10	Voir aussi	69
5.4	Seeburger	69
5.4.1	Historique	70
5.4.2	Produits	71
5.4.3	Liens externes	71
5.5	Tibco Software	71
5.5.1	Principaux produits	71
5.5.2	Principales acquisitions	72
5.5.3	Notes et références	72
5.5.4	Liens externes	73
6	Compléments	74
6.1	Théorie de l'information	74
6.1.1	L'information selon Shannon, un concept de la physique mathématique	74
6.1.2	Le statut physique de la théorie de l'information	76
6.1.3	Développement de la théorie mathématique de l'information	76
6.1.4	Exemples d'information	77
6.1.5	Contenu d'information et contexte	78
6.1.6	Mesure de la quantité d'information	78
6.1.7	Limites de cette théorie	80
6.1.8	Notes et références	80
6.1.9	Voir aussi	81

6.2	Liste de normes ISO par domaines	81
6.2.1	Normes fondamentales	82
6.2.2	Normes de spécifications	83
6.2.3	Normes d'analyse et d'essais	83
6.2.4	Qualité	83
6.2.5	Banques et services financiers	83
6.2.6	Technologies de l'information et de la communication	84
6.2.7	Recherche d'information	87
6.2.8	Responsabilité sociétale des entreprises	88
6.2.9	Sécurité	88
6.2.10	Notes et références	89
6.2.11	Voir aussi	89
6.3	ISO 8601	89
6.3.1	Historique	90
6.3.2	Principe	90
6.3.3	Avantages	96
6.3.4	Notes et références	96
6.3.5	Liens externes	96
6.4	Code lieu-fonction	96
6.5	Codage des caractères	97
6.5.1	Principe	97
6.5.2	Historique	98
6.5.3	Jeux de caractères codés populaires, par pays	104
6.5.4	Autres codages de caractères populaires	106
6.5.5	Notes et références	107
6.5.6	Voir aussi	107
6.6	ISO 8859	107
6.6.1	Introduction	107
6.6.2	Caractères	108
6.6.3	Les parties du standard ISO 8859	108
6.6.4	Relations avec Unicode et UCS	110
6.6.5	État de développement	111
6.6.6	Voir aussi	111
6.7	American Standard Code for Information Interchange	112
6.7.1	Histoire	113
6.7.2	Principes	113
6.7.3	Internationalisation	114
6.7.4	Standardisation	114
6.7.5	Approximation, variantes et extensions	115
6.7.6	Influence	116
6.7.7	Description	117

6.7.8	Voir aussi	118
6.8	Extended Binary Coded Decimal Interchange Code	119
6.8.1	Historique	119
6.8.2	Exemple de codage	120
6.8.3	Transcodage de l'ISO 8859-1 vers l'EBCDIC	120
6.8.4	Voir aussi	121
6.9	UTF	121
6.10	UTF-8	122
6.10.1	Liens avec la norme internationale ISO/CEI 10646 et les standards Unicode et de l'internet	122
6.10.2	Description technique	122
6.10.3	Description	122
6.10.4	Avantages	126
6.10.5	Inconvénients	127
6.10.6	Histoire	128
6.10.7	Notes et références	131
6.10.8	Voir aussi	132
6.11	Adresse IP	133
6.11.1	Utilisation des adresses IP	133
6.11.2	Adresse IP et nom de domaine	134
6.11.3	Classe d'adresse IP	134
6.11.4	Sous-réseau	134
6.11.5	Agrégation des adresses	135
6.11.6	Base de données des adresses IP	135
6.11.7	Identification par adresse IP	137
6.11.8	Notes et références	138
6.11.9	Annexe	138
6.12	Port (logiciel)	139
6.12.1	Origine du mot	139
6.12.2	Explication métaphorique	139
6.12.3	Utilité	139
6.12.4	Attribution des ports	140
6.12.5	Notes et références	141
6.12.6	Voir aussi	141
7	Sources, contributeurs et licences du texte et de l'image	142
7.1	Texte	142
7.2	Images	146
7.3	Licence du contenu	149

Chapitre 1

Introduction

1.1 Échange de données informatisé

L'**Échange de Données Informatisées (EDI)**, ou en version originale *Electronic Data Interchange*, est le terme générique définissant un échange d'informations automatique entre deux entités à l'aide de messages standardisés, de machine à machine. L'EDI a été conçu à l'origine dans l'optique du « zéro papier » et afin d'automatiser le traitement de l'information : disposer rapidement d'une information exhaustive et fiable. Dans la pratique, l'EDI permet de réduire notablement les interventions humaines dans le traitement de l'information, et donc de le rendre effectivement plus rapide et plus fiable. La rapidité et la meilleure fiabilité des échanges par EDI permettent de fluidifier les informations échangées, et de réduire considérablement les coûts de traitement tout en améliorant la sécurisation des transactions. Ceci est maintenant indispensable selon la politique de gestion en **juste à temps** qui augmente le nombre de transactions à traiter dans un temps de plus en plus court.

Par exemple, une entreprise peut émettre ses commandes, ses avis d'expédition, ses factures par EDI, c'est-à-dire qu'elle va émettre des messages électroniques normalisés qui seront télétransmis vers les ordinateurs de ses partenaires commerciaux (**fournisseurs, clients, banques**) qui seront capables d'interpréter et d'intégrer automatiquement les données correspondantes dans leurs systèmes d'information respectifs sans papier ni intervention humaine. Ce processus permet la suppression des documents papier acheminés par la poste ou par fax et des saisies manuelles par des opérateurs.

1.1.1 Définition

Les messages normalisés au niveau de la syntaxe et du vocabulaire (sémantique) véhiculent des données essentiellement codifiées, par exemple, dans la Grande Distribution : les **codes lieu-fonction** d'une entreprise sont des identifiants attribués par EAN France **GS1 France** (anciennement **GENCOD France**). Il est à la forme numérique d'un **EAN 13**. Son premier chiffre est le code pays (3 pour la France) suivi de 01 (fournisseur) ou 02 (distributeur), puis le CNUF de l'entreprise, suivi d'un code interne choisi par l'entreprise pour différencier ses différents lieux-fonctions. Le 13^e et dernier chiffre est, comme pour tout EAN13, la **clé de contrôle**. Il en est de même pour les codes articles qui apparaissent sur les codes barres des produits de la Grande Distribution (EAN13).

Les standards EDI ont été définis vers le milieu des années 80.

1.1.2 Quelques organisations

- GED DE FRANCE, Répertoire des professionnels de l'EDI en France
- GS1 France, anciennement GENCOD EAN FRANCE, organisme de standardisation pour les biens de consommation
- ODETTE, organisme de standardisation pour l'industrie automobile en Europe
- VDA et GALIA, associations membres de ODETTE International, qui traitent des échanges dans l'industrie automobile respectivement en Allemagne et en France.

- UNECE, branche de l'ONU pour l'Union Européenne, et auteur entre autres des standards EDIFACT (Electronic Data Interchange for Administration, Commerce and Transport).
- Dans l'industrie du semi conducteur en Europe, Edifice, The European B2B forum for the Electronics Industry
- EDI / SANDRE, instance de normalisation des échanges de données sur l'eau
- AGRO EDI EUROPE, association de promotion et de normalisation des Échanges Électroniques Professionnels Agricoles

1.1.3 Quelques normes EDI

- Syntaxes XML : EbXML, RosettaNet (électronique), CAP (pièces détachées automobiles), CIDX (chimie), PIDX (pétrole), etc.
- EDIFACT de UNECE/ISO, normes et syntaxe, utilisés dans le monde entier et dans des secteurs d'activité très variés tels que le transport, l'industrie automobile ou pharmaceutique et la grande distribution. Beaucoup de normes sectorielles sont basées sur les messages EDIFACT. C'est le cas pour les normes GALIA dans l'automobile ou INOVERT dans le transport.
- EDIEL Utilisé dans le cadre des marchés libéralisés de l'énergie (Pays nordiques et Belgique)
- EAN, aujourd'hui GS1 (biens de consommation)
- ODETTE (Industrie automobile)
- VDA (industrie automobile, zone d'influence allemande)
- INOVERT (Transport, France)
- ANSI X12 (États-Unis)
- Tradacom (Grande-Bretagne)
- ADatP-3 (norme OTAN)
- EBICS (France)

1.1.4 Quelques protocoles de communication

- EDIINT AS1 (smtp), AS2 (http), AS3 (ftp)
- OFTP (protocole d'échange établi par Odette : Odette File Transfer Protocol)
- SMTP (protocole utilisé pour les courriels entre autres)
- X.400 (utilisé par exemple par les réseaux Atlas400 et Allegro en France)

1.1.5 Quelques messages courants

- Commandes ou ordres de livraison

Exemple de messages : DELJIT, DELFOR, DELINS, VDA4905, VDA4915

- Avis d'expédition

Exemple de messages : AVIEXP, DESADV

- Ordres de transport
- Factures

Exemple de messages : INVOIC

- États et mouvements de stock

Exemple de messages : MOUSTO, BALSTO, INVRPT

- Clôture du stock

Commandes : message ORDERS

APERAK : Accusé applicatif

BANSTA : (Banking Status)

CNTCND : Conditions contractuelles

COACSU : (Commercial Account Summary)

COMDIS : (Commercial Dispute)

CONDRA : (Drawing Administration)

CONTRL : Contrôle

CREMUL : (Multiple Credit Advice)

DEBMUL : (Multiple Debit Advice)

DELFOR : (Delivery Schedule)

DESADV : Avis d'expédition

DIRDEB : (Direct Debit)

DISPOR : Ordre de livraison

FINCAN : (Financial Cancellation)

FINSTA : (Financial Statement)

GENRAL : Message général

HANMOV : Ordre de livraison

IFCSUM : Ordre de transport

IFTMAN : (Arrival Notice)

IFTMBC : (Booking Confirmation)

IFTMBF : (Firm Booking)

IFTMIN : Instructions de transport

IFTSTA : Etat du transport

INSDDES : Instructions d'expédition

INVOIC : Facture / Avoir

INVRPT : Journal des mouvements

MSCONS : (Metered Services Consumption Report)

ORDCHG : Modification de commande

ORDERS : Commande

ORDRSP : Réponse de commande

OSTENQ : (Order Status Enquiry)

OSTRPT : (Order Status Report)

PARTIN : Fiche société

PAYMUL : (Multiple Payment Order)

PRICAT : Prix catalogue
PRODAT : Fiche produit
PROINQ : (Product Inquiry)
QUALITY : (Quality Test Report)
QUOTES : Devis
RECADV : Avis de réception de marchandises
REMADV : Avis de paiement
REPORT : Confirmation de livraison
REQOTE : Demande de devis
RETANN : Avis de retour de marchandises
RETINS : Instructions de retour de marchandises
SLSFCT : (Sales Forecast Report)
SLSRPT : Etat des ventes
TAXCON : (Tax Control)

1.1.6 Articles connexes

- Traçabilité
- Cycle de vie (commerce)
- Analyse du cycle de vie
- Série des normes ISO 14000, ISO 14048
- Code lieu-fonction
-  Portail de l'informatique

Chapitre 2

Organismes de normalisation

2.1 Organization for the Advancement of Structured Information Standards

 Pour les articles homonymes, voir Oasis (homonymie).

L'**OASIS** (acronyme de *Organization for the Advancement of Structured Information Standards*) est un consortium mondial qui travaille pour la standardisation de formats de fichiers ouverts basés notamment sur **XML**.

L'OASIS a été créé en 1993 et compte 3 500 membres faisant partie de 600 organisations dans 100 pays.

L'adhésion à **OASIS** est payante et les membres sont surtout de grosses entreprises mais une organisation moins dotée peut y adhérer si les membres le décident, comme a pu le faire **Debian**.

OASIS est structuré en plusieurs groupes de travail nommés les *Technical Committees*.

Les standards produits sont librement distribuables sur le réseau et disponibles en **PDF** et en général dans un format **XML**.

Contrairement au **W3C**, la plupart des standards OASIS utilisent le langage de schéma **Relax NG** et non pas **XML Schema** du **W3C**.

OASIS est l'une des quelques organisations autorisées par l'ISO à proposer ses standards suivant la procédure « PAS » (*Publicly Available Specification*, ou spécification publiquement disponible), qui évite qu'un comité technique de l'ISO ait à dupliquer celui de l'OASIS.

2.1.1 Membres les plus connus de l'OASIS

- Adobe
- AMD
- Airbus
- AOL
- BEA
- Boeing
- Citrix
- Computer Associates
- Dell
- Département de la Défense américain

- EDS
- Fujitsu
- General Motors
- Hewlett-Packard
- Hitachi
- IBM
- Intel
- ISIS Papyrus
- Mandriva
- Microsoft
- NEC
- NIST
- Nokia
- Novell
- Oracle
- PeopleSoft
- Reuters
- Rogue Wave Software
- Ricoh
- RSA Security
- SAP
- Sun Microsystems
- Thomson
- Verisign
- Vignette
- Visa
- WebMethods

2.1.2 Standards les plus connus édictés par l'OASIS

- DocBook
- OpenDocument
- ebXML
- SAML
- UDDI
- LegalXML
- PKI

- Relax NG
- XDI
- XRI
- WS-BPEL

2.1.3 Lien externe

- (en) Site officiel
-  Portail de l'informatique

2.2 GS1

GS1 (pour Global Standards^[1]) est un organisme mondial actif dans le domaine de la normalisation des méthodes de codage utilisées dans la chaîne logistique.

L'objectif de GS1 est d'établir des standards et des solutions permettant l'échange de biens de consommation sur la planète, visant ainsi toute la chaîne de distribution du producteur au consommateur.

GS1 a son siège à Bruxelles (Belgique).

2.2.1 Notes et références

[1] <http://www.publications.gs1.fr/Publications/Global-Standards-and-Solutions-in-Waste-Management>

2.2.2 Annexes

Article connexe

- GS1 France
- GS1 Canada

Lien externe

- Site officiel
-  Portail du commerce
-  Portail des technologies

2.3 GS1 France

GS1 France, anciennement **Gencod EAN France**, est un organisme de normalisation français, émanation locale de l'organisme international GS1. C'est une SA à Directoire et Conseil de Surveillance à but non lucratif.

Gencod est presque devenu, pour certains professionnels, un nom commun : les logisticiens ou les chefs de rayon emploient parfois le terme *Gencod* pour désigner les codes à barres. Gencod signifie originellement « Groupement d'études de normalisation et de codification ».

GS1 France a vocation à développer des standards permettant aux entreprises de communiquer entre elles. Par exemple, un industriel et un distributeur peuvent s'échanger des informations logistiques et commerciales grâce aux

standards diffusés par GS1 France. Les principes de codification qui sous-tendent les codes à barres constituent un standard parmi d'autres. GS1 France travaille également sur des langages fonctionnels comme EANCOM et GS1 XML]. On parle ainsi de « système » GS1 pour désigner l'ensemble des standards que maintient GS1 France : des codes, à leur symbolisation ou à leur inscription dans des puces électroniques RFID, jusqu'aux langages.

Historiquement associée au monde de la grande distribution, GS1 France couvre aujourd'hui de nombreux secteurs d'activités (construction, santé, défense, etc.). GS1 France est membre d'une organisation internationale, GS1 (anciennement EAN.UCC) et garantit la compatibilité de la majorité des solutions proposées avec celles maintenues par ses homologues. GS1 est présent dans cent trois pays et offre ses services à environ un million d'entreprises.

L'organisme a changé de nom le 10 février 2005 : Gencod EAN France s'appelle désormais GS1 France (suite au changement de nom de l'organisation internationale EAN.UCC, devenue GS1).

2.3.1 Lien externe

- Site officiel



- Portail de l'informatique



- Portail des télécommunications

2.4 Comité français d'organisation et de normalisation bancaires



Pour les articles homonymes, voir [Comité d'organisation](#).

Le **Comité français d'organisation et de normalisation bancaires** ou **CFONB** est un organisme professionnel qui a pour mission d'étudier et de résoudre, aux plans de l'organisation et des normes, les problèmes de caractère technique liés à l'activité bancaire. Ses travaux portent essentiellement sur les moyens et systèmes de paiement, mais concernent également le domaine des valeurs mobilières.

C'est par exemple le CFONB qui édicte les règles et les normes à appliquer relatives aux mentions et aux libellés devant figurer sur les extraits ou relevés de compte (intitulés des opérations) ou qui avait défini les protocoles ETEBAC (désormais obsolètes).

Aujourd'hui, le CFONB est impliqué dans la normalisation du protocole EBICS (une des alternatives, avec SWIFT, pour remplacer le protocole ETEBAC). Le CFONB est aussi impliqué dans la migration vers les virements et prélèvements européens SEPA, qui ont vocation à remplacer à partir du 1er février 2014 les virements et prélèvements français, autrefois normalisés par le CFONB.

2.4.1 Liens externes

- cfonb.org, le site web du CFONB
- Présentation sur cfonb.org



- Portail des associations



- Portail de l'économie

2.5 Groupement pour l'amélioration des liaisons dans l'industrie automobile

Le **Groupement pour l'amélioration des liaisons dans l'industrie automobile** (acronyme **GALIA**^[1]) est un organisme associatif français, régi par la loi du 1^{er} juillet 1901, de standardisation des moyens d'échange de produits et d'informations, créé en 1984 par et pour l'industrie automobile française.

2.5.1 Présentation

GALIA compte à ce jour environ 300 membres :

- sociétés industrielles (constructeurs automobiles, équipementiers et fournisseurs) ;
- sociétés de services (informatiques, prestataires logistiques, transporteurs) ;
- organisations.

Le groupement est par ailleurs membre de l'organisation européenne ODETTE, dont la mission est identique à la sienne pour l'Europe, et participe à des projets globaux en partenariat avec ses homologues américaine (AIAG), et japonaises (JAMA/JAPIA).

GALIA est connu auprès des opérateurs logistique du secteur automobile pour proposer un format d'étiquettes de traçabilité normalisé et utilisé, entre autres, par PSA et RSA ainsi qu'une série de plans d'emballage pour la maintenance de petits éléments. Galia a défini ainsi une gamme de bacs plastiques de différentes dimensions avec des critères dimensionnels et de fonctionnalités précis afin d'optimiser les manutentions et les volumes transportés et leurs équivalents cartons. Les fabricants de bacs et cartons doivent répondre alors aux cahiers des charges émis par Galia pour pouvoir apposer le nom de cet organisme sur leurs produits.

La norme Galia dans les emballages est principalement utilisé chez les constructeurs français. Les constructeurs germaniques utilisent pour la plupart les normes KLT et Europe. Parmi les fabricants de bacs plastique répondant aux critères Galia on peut Linpac, Utz et Gamma Wopla.

2.5.2 Notes et références

- [1] Michel Roux, Gilles Fleury, « Pilotez votre plateforme logistique : Les logiciels de gestion d'entrepôts ... », sur *books.google.fr* (consulté le 17 octobre 2014).

2.5.3 Voir aussi

Liens externes

- GALIA
- ODETTE
- AIAG
- JAMA
- JAPIA
-  Portail de l'automobile
-  Portail de la France

2.6 VDA

 Cette page d'homonymie répertorie les différents sujets et articles partageant un même nom.

VDA est une abréviation, qui signifie :

- *Verband der Automobilindustrie* (Union de l'industrie automobile) est, en Allemagne, l'organisme définissant les standards à appliquer à l'industrie automobile (en ce qui a trait à la qualité, la logistique, l'échange de données, etc.) ;
- *Volksbund für das Deutschtum im Ausland*, (en allemand : *l'association pour le germanisme à l'étranger*), est une organisation pangermaniste ;
- *Vallée d'Aoste*, région autonome et bilingue de l'Italie du nord ;
- *VDA - Verre Design Architecture*, intitulé d'une formation de master universitaire dont le thème de recherche est axé sur le matériau verre.
- Ville-d'Avray

VDA est un code, qui signifie :

- Volga-Dnepr Airlines (Russie), selon le code OACI des compagnies aériennes ;
selon le code AITA des compagnies aériennes, Volga-Dnepr Airlines a pour code VE.

2.7 EDI-Optique

 Pour les articles homonymes, voir [EDI](#).

EDI-Optique est une association française à but non lucratif créée en 1998, dont l'objectif est la standardisation des échanges de données informatisés (EDI) dans le secteur de l'optique ophtalmique. EDI-Optique compte plus de 60 membres verriers, lunetiers, laboratoires de contactologie, éditeurs de logiciels de gestion de point de vente et toutes centrales et enseignes françaises.

Les standards OPTO (OPTO v10, OPTO v11, OPTO v33, OPTO v34, OPTO v35) sont largement utilisés par la communauté des opticiens et par leurs fournisseurs. En 2006, EDI-Optique a collaboré avec l'UNOCAM pour définir le standard OPTOAMC d'échange d'informations entre les opticiens et les assurances maladie complémentaires.

En 2011, EDI-Optique a publié un nouveau standard de catalogue s'appuyant sur les normes ebXML. L'optique est ainsi l'une des premières filières de la distribution à adopter les technologies du future pour ses catalogues électroniques.

Les standards OPTO sont utilisés au-delà des frontières françaises.

2.7.1 Voir aussi

- Site officiel
-  Portail de l'informatique

Chapitre 3

Formats de données

3.1 Échange de données informatisées pour l'administration, le commerce et le transport

EDIFACT est l'acronyme pour **Échange de données informatisées pour l'administration, le commerce et le transport** (en anglais, *Electronic Data Interchange for Administration, Commerce and Transport*), qui est une norme des Nations unies décrivant des modalités techniques pour l'échange de données informatisé (EDI) dans différents secteurs industriels.

EDIFACT définit à la fois une *syntaxe* et un *contenu* pour les messages EDI.

La norme générale est adaptée par des organismes de normalisation nationaux et sectoriels, afin de mieux prendre en compte les besoins de chaque *branche d'activité*. En France on peut en particulier citer les normes GENCOD (grande distribution), INOVERT (transport), ODETTE (automobile) et ETEBAC (banque).

Au moins deux fois par an, la norme fait l'objet d'une mise à jour globale. L'ensemble de cette mise à jour fait l'objet de la création d'un nouveau répertoire des données et messages. Chaque répertoire est identifié par l'année (sur deux chiffres) ainsi qu'un numéro ou une lettre pour indiquer sa position dans l'année (1 ou 2 pour les anciens répertoires, A, B ou C pour les plus récents).

L'ISO a adopté EDIFACT par la norme ISO 9735.

3.1.1 Organisations sectorielles

Secteur bancaire

Dans le secteur bancaire français, EDIFACT est associé au protocole de transmission **ETEBAC 5** (www.etebac.com) pour *Échanges Télématiques Banque Client*. Cette norme d'échanges est édictée par le *Comité Français d'Organisation et de Normalisation Bancaires* (CFONB,) qui dépend de la Fédération Bancaire Française (FBF,)

Les principaux messages EDIFACT standardisés par le CFONB sont :

- PAYORD (répertoire 91.2), pour les virements domestiques et les VSOT.
- PAYEXT puis PAYMUL (répertoire 96 A), pour les virements commerciaux (VCOM).

Le CFONB préconise par ailleurs l'utilisation de fichiers de formats fixes qui sont bien plus simples à utiliser. Le format fixe de longueur 320 est désormais utilisable pour toute forme de virement (excepté les virements commerciaux qui sont associés au format fixe de longueur 400).

Dans le secteur bancaire mondial le message PAYMUL du répertoire 96 A est parfois proposé par les banques pour la formalisation des paiements domestiques et internationaux.

Le message FINSTA, destiné au relevés de comptes, est par contre rarement utilisé en raison de sa complexité.

Secteur agricole

La normalisation de messages dématérialisés s'organise via l'association Agro EDI Europe , fondée en France en 1992.

Secteur automobile

- Site officiel de ODETTE.

Secteur aérien

- Site officiel de IATA

3.1.2 Voir aussi

Articles connexes

- Échange de données informatisées

Liens externes

- Site officiel UN/EDIFACT de la Commission économique pour l'Europe.



- Portail du management



- Portail de l'informatique



- Portail des télécommunications

3.2 electronic Business using XML

ebXML (prononcer i-bi-ix-em-el, comme dans eBusiness, bien que la prononciation eu-bé-ix-em-el soit très répandue), abréviation de l'anglais *Electronic Business using eXtensible Markup Language* signifiant *commerce électronique en utilisant XML*, est une suite de spécifications basées sur le langage XML utilisable pour le commerce électronique.

Cette suite a été éditée par OASIS et UN/CEFACT sous la forme de la spécification technique ISO/TS 15000.

L'objectif est de fournir une infrastructure globale, ouverte, fondée sur XML, permettant d'assurer les échanges électroniques professionnels (EEP ou B2B en anglais) de manière interopérable.

ebXML est couplé avec le registre UDDI pour la recherche de services web.

3.2.1 Introduction

Complémentaire à EDIFACT (*Electronic Data Interchange for Administration, Commerce and Transport*) qui est plutôt réservé aux échanges importants entre partenaires stables, ebXML s'adresse aux entreprises de toutes tailles, en permettant de définir les échanges électroniques entre l'entreprise et ses divers partenaires, comme les catalogues électroniques...

La spécification prétend se situer dans le prolongement de la démarche déjà initiée depuis une trentaine d'années, appelée EDIFACT.

L'architecture d'ebXML est conçue pour être couplée avec celle du registre UDDI. Elle est basée sur l'utilisation commune de composants et de profils de partenaires commerciaux ("trading partner profiles"). Ces concepts sont directement mis en œuvre dans les spécifications d'ebXML.

ebXML est issu des technologies de l'ooEDI (échange de données informatisé orienté objet), des concepts méthodologiques d'UML/UMM, de la mouvance XML et du projet « future vision » du groupe ANSI X12 EDI.

Ces travaux sont déjà intégrés à la version actuelle d'ebXML. L'activité continue pour prendre en compte les avancées d'autres groupes tels que l'OMG et OASIS – notamment la spécification BCM d'OASIS (Business-Centric Methodology ou Méthodologie Orientée Métier).

3.2.2 Historique

En Europe, les gouvernements ont pris conscience des enjeux du commerce électronique vers 1997. En France, Francis Lorentz a commencé en octobre 1997 une mission sur le commerce électronique, qu'il a présentée en mai 1998.

ebXML a débuté en 1999 comme une initiative commune entre l'UN/CEFACT (United Nations Centre for Trade Facilitation and Electronic Business) et OASIS (Organization for the Advancement of Structured Information Standards). Un comité de coordination commun regroupe des représentants de chacun des deux organismes pour maintenir l'effort. Des réunions trimestrielles des groupes de travail ont été tenues entre novembre 1999 et mai 2001. À la session plénière finale, un protocole d'accord a été signé par les deux organismes, fractionnant la responsabilité des diverses caractéristiques mais continuant la supervision par un comité de coordination commun.

Le projet original a envisagé cinq couches de spécifications de données, y compris des normes XML pour :

- Processus métier (Business Process, BP),
- Accords de Protocole de Collaboration (CPA),
- Composants communs de données (Core data components, CDC, ou Core components),
- Transmission de messages,
- Enregistrements et dépôts.

À la fin des travaux des deux organismes, le livrable a été soumis au comité technique 154 de l'ISO pour approbation. L'Organisation internationale de normalisation (ISO) a approuvé les cinq caractéristiques suivantes d'ebXML en tant que la norme d'ISO 15000, sous le titre général de “Commerce électronique en langage de balisage extensible (ebXML)” (en anglais : “Electronic business eXtensible Markup Language”) :

- ISO 15000-1 : accord de collaboration de profil d'associé ebXML,
- ISO 15000-2 : spécifications de service “messages” d'ebXML,
- ISO 15000-3 : modèle d'enregistrement d'information ebXML,
- ISO 15000-4 : spécification d'enregistrement des services ebXML,
- ISO 15000-5 : Spécification Technique des Composants Communs ebXML (CCTS) 2.2.

Toutes ces normes sont disponibles uniquement en anglais.

Les comités techniques OASIS et UN/CEFACT assurent la responsabilité de maintenir et de faire progresser les caractéristiques ci-dessus.

L'initiative freebXML.org a été établie pour favoriser le développement et l'adoption du logiciel de gestion de la base ebXML en open source.

ebXML a engendré d'autres travaux sur l'échange de données et de métadonnées, avec le SDMX (Statistical data and metadata exchange).

3.2.3 Travail actuel

Les travaux continuent (2006-2007) sur le renforcement et la maturation d'ebXML, en particulier dans les secteurs des services web et des architectures orientées services (SOA, Service Oriented Architectures).

Transmission de messages (ebMS)

Les dernières spécifications de transmission de messages d'ebXML (v3.0) peuvent être regardées comme une spécialisation des services Web pour des échanges business-business (B2B)^[1]. La conformité aux normes de services web n'enlève pas la raison d'être des logiciels personnalisés du web prévus pour la transmission des messages. Des modes variés de traitement de messages doivent être soutenus au-delà de la notion de service : un message peut être aligné et utilisé par un processus d'affaires, être mis en attente pour un traitement par lot, être expédié à une application, être routé par un bus SOA, etc. v3.0 réutilise les caractéristiques des services Web qui fournissent des fonctions au niveau du protocole (sécurité, fiabilité), et construit les dispositifs qui satisferont les besoins d'affaires : en-têtes standard d'affaires, modèle d'échange pull (non simplement push ou question/réponse), soutien des transactions d'affaires (comme dans UMM), autorisation de message au niveau applicatif.

Processus d'affaires et collaboration (ebBP)

L'ebBP (suite de BPSS) vise la surveillance des processus de collaboration d'affaires entre partenaires^[2]. En décembre 2006, l'ebXML BPSS (Business Process Schema Specifications ou spécifications de schéma de processus d'affaires) ou l'ebBP v2.0.4 a été approuvé comme un standard OASIS. La soumission à l'ISO est prévue.

L'ebBP v2 améliore l'utilisation d'ebXML, des services web et des possibilités reliées par SOA telles que contexte et l'appui de collaboration multiple. L'ebBP assiste les communautés pour l'échange de fichiers de base, de transmission de messages, et/ou les échanges basés sur SOAP. Il spécifie les activités économiques - simples ou complexes, de niche ou composées - et la qualité des contrats de services tels que la fiabilité, la sécurité, et l'émission d'avis d'exception ou d'avis de réception prévus. Les modèles de transaction, les expressions conditionnelles, les variables sémantiques d'affaires et d'autres fonctions soutiennent les activités variées dans des collaborations binaires ou multi-parties constituant l'environnement de commerce électronique.

Le perfectionnement et l'adoption de l'ebBP ont été conduits principalement par des communautés d'utilisateurs dans des domaines tels que l'eGouvernement, la santé, les services financiers, la comptabilité et autres. Un éditeur open source, freebxmlbp, est disponible sur SourceForge.net.

Profil et accord de protocole de collaboration (CPPA)

Le comité technique CPPA (Collaboration Protocol Profile and Agreement) ebXML a actuellement deux initiatives principales. La première concerne l'accomplissement d'une mise à jour de la version 2 de l'ebXML CPPA, et la seconde est une note sur la mise en place de services pour l'échange et la conclusion de CPAs (accords de protocoles de collaboration)^[3].

La mise à jour de CPPA inclut plusieurs clarifications éditoriales et modifications techniques de schéma qui ont été approuvées depuis la publication de la version 2 (ISO 15000-1 de l'ebXML CPPA : 2004). Une annexe normative définit comment des codes d'ISO 6523 sont mis en correspondance des URIs qui identifient des « types de parties ». Les changements principaux de cette mise à jour des spécifications de CPPA se situent dans un prolongement qui se sert de groupes de substitution pour l'extension des CPPs et CPAs pour rencontrer d'autres approches de processus d'affaires et/ou du réseau de transport. L'application de cette extensibilité à d'autres modèles de protocoles de transmission de messages (EDIINT, WS-* et ebMS 3.0) est fournie.

Des applications d'incorporation de WSDL et WS-Policy sont également en projet. Les spécifications de processus d'affaires d'ebBP d'OASIS sont adaptées en conséquence dans la nouvelle approche d'extensibilité. Des exemples de la version 2 de CPP ou CPA seront validés suivant le nouveau schéma après le changement de l'espace de noms (namespace) et quelques modifications éditoriales en conformité avec les préfixes d'attribut et les choix par défaut du nouveau schéma. En fin de compte, une structure alternative pour CollaborationRole permet un contenu de modèle beaucoup "plus plat". Les projets de négociation ont été activement mis à jour pour se servir du document descripteur de négociation dans un processus considérablement simplifié d'accord de collaboration. Le document en est aux étapes initiales du processus rédactionnel.

Enregistrement (registre) et dépôt

Les spécifications d'enregistrement d'ebXML OASIS ont été développées en vue de réaliser des enregistrements et des dépôts interopérables, avec une interface qui permette tant la soumission, que l'interrogation et la récupération du

contenu de l'enregistrement et du dépôt^[4]. Un enregistrement/dépôt est un composant important d'une infrastructure de SOA.

Les spécifications d'enregistrement d'ebXML ont évolué jusqu'à fournir un ensemble très riche de possibilités pour satisfaire aux exigences des SOA en matière de gouvernance et de gestion fédérée d'information agrégée. Les dispositifs incluent la gestion des contrôles d'accès et d'identité, les recherches prédéfinies et les recherches spécifiques, la version, les taxonomies définies par l'utilisateur, types de relation étendus aussi bien que les types de relations définis pour l'utilisateur, le support des engagements dont HTTP (REST) et SOAP API, l'abonnement et l'avis d'événement, les questions, la validation de WSDL, et les prolongements standards nécessaires à certains cas d'utilisation spécifiques. La version 3.0 des normes d'enregistrement d'ebXML (ISO 15000-3 et -4) est disponible en téléchargement^[5].

Composants de noyau (CCTS)

Article détaillé : [Core Component Technical Specification](#).

Au sein de l'UN/CEFACT TMG, le travail continue sur des composants de noyau (en anglais *core component*), appelés quelquefois composants communs.

Ces travaux s'intéressent aussi à la technologie relative telle que l'assemblage de contexte et de contenu.

Le composant commun est un concept retenu dans le cadre commun d'interopérabilité européen, et en France dans le référentiel général d'interopérabilité RGI dans sa partie Le volet sémantique.

3.2.4 Vue d'ensemble conceptuelle d'architecture d'ebXML

Structure en trois couches

Tandis que les spécifications d'ebXML adoptées par l'ISO et OASIS cherchent à fournir les procédures formelles XML qui peuvent être mises en application directement - l'architecture d'ebXML est concentrée sur les concepts et les méthodologies qui peuvent plus largement être appliqués pour permettre à des praticiens d'améliorer des solutions d'échanges électroniques professionnels.

La spécification ebXML comporte trois couches :

- la couche business information,
- la couche business process,
- la couche partner discovery.

Cette architecture globale, ouverte et interopérable permet des utilisations par les moteurs de recherche sémantiques.

Source : ebXML technical architecture, Vancouver, 2000

Couche “business information”, contexte

Article détaillé : [Composant commun](#).

Lors de la modélisation des processus, les messages sont définis en utilisant des composants communs (*core components*), qui appartiennent également au registre d'annuaire UDDI. Ces composants sont réutilisables comme en EDIFACT. Ainsi, des éléments communs à plusieurs messages, tels que l'adresse, les moyens de contacts, les coordonnées bancaires, etc. ne sont définis qu'une seule fois. Ils figurent dans le type de donnée businessEntity d'UDDI

Des composants de base peuvent être élémentaires (code postal, nom, montant, etc.) ou agrégés (adresse, personne physique, etc.). Les composants de base et les principes de construction de blocs réutilisables sont décrits dans le document CCTS.

Les composants de base peuvent varier en fonction de différents éléments tels que le pays, le secteur, le message, etc. Les informations associées à une personne physique, dans un cadre médical sont assez différentes de celles utilisées dans un cadre professionnel (votre auditrice ne voudra peut-être pas vous donner sa date de naissance).

EbXML permet d'adapter les composants de base en fonction de leur utilisation à travers la notion de contexte. Il définit huit dimensions contextuelles : le processus, la localisation géopolitique (pays, région, etc.), le rôle du partenaire, le rôle du tiers, le secteur, le produit ou le service, la législation / réglementation, les possibilités des systèmes.

Outre les composants communs élémentaires et agrégés, qui portent une sémantique métier, il existe une troisième catégorie, le composant commun type, qui ne porte pas de sémantique.

La spécification technique de composants (CCTS, Core Component Technical Specification) continue dans UN/CEFACT. Elle fait l'objet de la spécification ISO/TS 15000-5.

Son cousin UBL (Universal Business Language), est une spécification développée par un groupe OASIS pour mettre en application des transactions spécifiques de XML basées sur l'application des principes de CCTS aux transactions typiques de la chaîne logistique, telles que la facture, la commande, la demande de transport, etc. UBL est déjà utilisé et mis en œuvre par des administrations d'Europe du nord et d'Espagne.

Couche “processus d'affaires”

Article détaillé : Processus d'affaires.

ebXML intègre une méthodologie et une infrastructure complète de support des échanges électroniques. Il impose une modélisation des processus d'affaires et des messages permettant la réutilisation des parties déjà définies (processus ou messages).

ebXML utilise une méthodologie créée par l'UN/CEFACT, UMM (UN/CEFACT Modelling Methodology) qui définit les processus et les messages en utilisant le standard de notation de modélisation UML (Unified Modeling Language) de l'OMG (Object Management Group).

Les processus d'affaires sont décrits à travers un langage BPSS (Business Process Schema Specification) qui est exprimé en XML [BPSS].

Couche “partenaires”

Le principe d'ebXML est de faciliter l'interopérabilité entre les entreprises, avec des architectures agiles.

La description XML BPSS montre l'enchaînement de tous les échanges possibles, avec leurs règles métier, dans un scénario pouvant être multipartite (notion de **chorégraphie des Services Web WS-*** – choreography), mais ne traite pas des processus internes aux entreprises (notion d'**orchestration**). Ainsi, un processus complet d'achat international modélisera les échanges de messages entre les différents partenaires (acheteur, vendeur, transporteurs, douanes, assurances, banques, etc.) vus de manière globale mais excluant les adaptations internes aux partenaires (traitement de la commande vers les stocks, la fabrication, intégration dans les PGI, etc.).

Génération de messages

Lors de la conception du message, un diagramme de classes UML est défini pour le message. Ce diagramme est alors utilisé pour générer un schéma XML [XML Schemas] du message, à travers un mapping UML Core Components défini dans le document [UML-CC].

La description du message utilise les composants de base, adaptés au contexte du message. Ce message est indépendant de la syntaxe utilisée pour l'échange : EDIFACT ou ebXML. Le schéma XML du message doit respecter les spécifications XML Design and Naming Rules [XML-NDR].

3.2.5 Les secteurs d'activité concernés

La spécification générale est adaptée par des organismes de normalisation nationaux et sectoriels, afin de mieux prendre en compte les besoins de chaque branche d'activité.

Secteur de la comptabilité, de la fiscalité et du reporting

L'Association Edificas s'est intéressée très tôt au traitement de la **comptabilité** et à la communication des informations financières afin d'emmagasiner des gains de productivité et de faciliter la fluidité des échanges sur le sol national français. En fait, depuis 1987. Cela s'est traduit très rapidement par l'adoption de la norme **EDIFACT**, rejoignant ainsi le monde de l'industrie, du commerce et de la finance auquel la **comptabilité** est naturellement associée, mais également le monde de la normalisation car il était souhaitable d'asseoir les projets sur des concepts solides en dehors de toute notion de **propriété intellectuelle**.

L'initiative ebXML de la **comptabilité** dépend d'un certain nombre de modèles (voir tableau 1) dont les principaux concernent le **journal**, le **plan comptable**, le **grand livre**, la **balance comptable**, le **bilan** et le **reporting** (comptable, financier et fiscal). D'autres sont venus se rajouter et plus particulièrement les plans de regroupement permettant d'obtenir à partir de la **comptabilité** les agrégats informationnels destinés à alimenter le **reporting**.

Depuis 1998, avec l'émergence de **XML**, la démarche s'effectue parallèlement dans les deux mondes : celui d'**EDIFACT** et celui d'**ebXML**. Dans ce dernier domaine, deux projets sont menés actuellement et qui concernent notamment l'**écriture comptable** : le projet **ENTREC** dans le monde **EDIFACT**, et le projet *Accounting entry* dans le monde **ebXML**.

En complément de cette initiative, il faut relever la présence du "jeton comptable" électronique, émanant de l'**écriture comptable ebXML**, et qui est destiné à être accroché à tout message **ebXML** devant faire l'objet d'une écriture en **comptabilité**, ce qui est le cas actuellement de la chaîne commerciale et des **marchés publics**.

Tous ces travaux sont réalisés au sein du **TBG12 Accounting & Audit** de l'**UN/CEFACT**. la **comptabilité**, la **fiscalité** et le **reporting** (Edificas).

Secteur du commerce

En France, le commerce (**GENCOD**) est un secteur très actif.

Secteur de la banque et de la finance

Le secteur de la **banque** et la **finance** est pris en charge par **CFONB**, et **SWIFT**. Le secteur de l'assurance est pris en charge par **CFNA**, et **GEMA** pour les mutuelles d'assurance

Secteur de l'agriculture

Le secteur de l'**agriculture** est pris en charge au niveau européen par **Agro EDI Europe** .

Secteur de la construction

Le secteur de la construction est pris en charge par **EDIBUILT**. Les équipements techniques du bâtiment sont traités par **EDIBATEC**.

Secteur de la pharmacie

Le secteur de la **pharmacie** est pris en charge par **EDIPHARM**.

Textile habillement

Le secteur de la chaussure est pris en charge par **EDI Chaussure**, la filière **Textile/Habillement/Distribution** par **EDI-TEX**.

Autres secteurs d'activités

Les huissiers (**CNHJ**), l'édition (**DILICOM**), la gestion de l'eau (**OICE**), etc.

3.2.6 Utilisations dans le cadre gouvernemental

Asie

La Chine a adopté ebXML comme standard national pour le commerce électronique.

Voir : Catalog of Chinese standards

États-Unis

Les États-Unis n'emploient pas ebXML pour l'administration fédérale ni pour celle des États. Le gouvernement fédéral dispose pour les grandes fonctions régaliennes de registres de métadonnées qui s'appuient sur la norme ISO 11179 (définition des concepts et organisation).

Union européenne

Bien que la vocation initiale des spécifications de l'UN/CEFACT soit une application au commerce électronique, les concepts d'ebXML sont proposés dans l'Union européenne pour les systèmes d'e-gouvernement en tant que cadre commun d'interopérabilité. En France, le Référentiel Général d'Interopérabilité RGI fait référence à une méthodologie inspirée des concepts d'ebXML : la méthode UML/XML.

3.2.7 Notes et références

- [1] http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ebxml-msg
- [2] http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ebxml-bp
- [3] http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=ebxml-cppa
- [4] http://www.oasis-open.org/committees/tc_home.php?wg_abbrev=regrep
- [5] <http://www.docs.oasis-open.org/regrep/v3.0/regrep-3.0-os.zip>

3.2.8 Voir aussi

Bibliographie

- UN/Cefact's Modelling Methodology (UMM) : UMM Meta Model – Foundation Module Version 1.0, Technical Specification 2006-10-06 ;
- UN/Cefact – ebXML Core Components Technical Specifications version 2.01 – ISO 15000-5 ;
- UN/Cefact Business Requirements Specification version 1.5 (Cefact/ICG/005) ;
- Unified Modeling Language (UML version 1.4) ;
- [BPSS] UN/CEFACT Business Process Schema Specification – Version 1.10 – 2003-10-18 , pour d'autres version, : Menu "Download", puis sélectionner "Business Process Working Group, puis "BPSS".
- [CCTS] – ebXML Core Components Technical Specifications Version 2.01 – ISO 15000-5 :2005 – , pour d'autres versions, <http://www.untmg.org> : Menu "Download", puis sélectionner "Core Components Working Group, puis " Core Components Technical Specifications ".
- [UML-CC] UML Profile for Core Components – Version candidate 1.0 (fondée sur CCTS version 2.1) – 2006-10-31 - , pour d'autres versions, : Menu "Download", puis sélectionner "Business Process Working Group, puis "UML profile for Core Components".
- [XML-NDR] XML Naming and Design Rules – Version 2.0 – 2006-02-17 - , pour d'autres versions :

Articles connexes

- Commerce électronique
- Composant commun
- CCTS
- ISO/TS 15000
- Référentiel Général d'Interopérabilité
- Processus d'élaboration d'une norme ISO
- Métadonnée
- Registre de métadonnées
- ISO/CEI 11179
- Catégorie:Métadonnées

Liens externes

- (en) ebXML.org Site officiel
- Site sur l'ebXml
- Centre des nations unies pour la facilitation des procédures commerciales et le commerce électronique
- EdiFrance
- (en) List of Web service markup languages
- (en) Simple introduction to ebXML
- (en) RosettaNet
- (en) freebXML.org
- (en) freebXML Registry open source project
- (en) OASIS homepage
- (en) OASIS News : ISO Approves ebXML OASIS Standards
- (en) UN/CEFACT homepage
- (en) ISO/TS 15000
-  Portail de l'informatique

3.3 Electronic Data Interchange for the Electric Industry

EDIEL est une norme décrivant des modalités techniques pour l'échange de données informatisé (EDI) dans le secteur de la distribution de l'énergie. Elle a été créée par l'organisation Nordel qui regroupe les transporteurs d'électricité des pays nordiques.

Elle est implémentée dans le cadre des marchés libéralisés du gaz et de l'électricité en Belgique, dans les pays nordiques, en Allemagne et aux Pays-Bas.

La norme permet aux différents acteurs du marché (fournisseurs, gestionnaires de réseaux de distribution, responsables d'équilibre et gestionnaires de réseaux de transport) d'échanger des données.

3.3.1 Messages

La norme utilise les messages suivants :

- UTILTS utilisé pour transmettre des consommations mesurées, estimées ou agrégées
- UTILMD contient la description de l'infrastructure (*Master data* en anglais, ou données de référence) ; point d'accès au réseau électrique, compteur, caractéristiques utilisateur...
- APERAK Message de confirmation
- INVOIC utilisé pour la facturation de l'utilisation du réseau (*Grid fee* en anglais)

3.3.2 Voir aussi

- Marché de l'électricité

3.3.3 Liens externes

- www.nordel.org Site de Nordel
- www.ediel.org Site du forum nordique EDIEL.

-  Portail de l'informatique

-  Portail de l'énergie

3.4 Échange télématique banque-clients

Échange télématique banque-clients (ETEBAC) est un protocole de télétransmission bancaire.

Ce protocole permet l'échange de fichiers entre les banques et leurs clients. Ce protocole existe depuis 1991^[1]. La version 5 a été créée en 1995^[2].

Les fichiers échangés sont regroupés en deux catégories :

- sens client ⇒ banque (ou sens aller) : il s'agit des fichiers de remises (virements, prélèvements, LCR (lettres de change relevées), etc.)
- sens banque ⇒ client (ou sens retour) : il s'agit des fichiers de relevés (extraits de comptes journaliers, extraits de comptes mensuels, relevés de prélèvements à payer, relevés de prélèvements impayés etc.)

La communication téléphonique (RTC ou RNIS) est toujours à l'initiative du client. Les fichiers sont ainsi, soit déposés sur le serveur par le client (sens aller) soit récupérés sur le serveur par le client (sens retour).

À l'heure actuelle, seules deux versions de ce protocole sont encore utilisées : ETEBAC3 et ETEBAC5.

- ETEBAC3 n'est pas sécurisé (la communication n'est pas chiffrée et l'utilisateur n'est pas formellement authentifié) et oblige le client à confirmer ses transferts (sens aller) de fichiers par fax à la banque.
- ETEBAC5 est sécurisé par l'utilisation de cartes à puce utilisant l'algorithme RSA, les fichiers sont chiffrés et l'émetteur authentifié. La banque vérifie la signature et émet un accusé de réception signé lui aussi.

Les standards ETEBAC et les formats de fichiers échangés (sauf formats propriétaires banques) ont été définis par le CFONB.

Il est à noter que France Télécom a prévu d'arrêter la maintenance du réseau X.25 sur lequel transitent les informations au format ETEBAC. L'arrêt du réseau X.25 était dernièrement prévu pour fin 2011, il est une nouvelle fois repoussé, pour le 30 juin 2012.

En remplacement de ce protocole, les solutions retenues par le CFONB sont le web-banking, le protocole EBICS ou le protocole SWIFTNet par l'intermédiaire du réseau SWIFT^[3].

3.4.1 Notes et références

[1] « www.gsit.fr »

[2] « archives GSIT »

[3] « Migration ETEBAC vers EBICS et SWIFTNET », sur www.cfonb.org

-  Portail de la finance
-  Portail de l'informatique
-  Portail des télécommunications

3.5 Electronic Banking Internet Communication Standard

EBICS (ou **Electronic Banking Internet Communication Standard**) est un protocole de communication sécurisé permettant l'échange de fichiers entre des clients et des établissements bancaires.

3.5.1 Origines

EBICS a été conçu par le **ZKA (en)** (*Zentrale Kreditausschuss*) et a été ensuite choisi par le Comité français d'organisation et de normalisation bancaires (CFONB), dans le cadre du remplacement d'ETEBAC. Ses modalités de fonctionnement et de distribution sont analogues à ces anciens protocoles. Les modalités d'utilisation pour la communauté française ont été précisées dans un « Guide de mise en œuvre »^[1].

3.5.2 Aspects techniques

EBICS est un protocole sur IP, utilisant des messages XML véhiculés par le standard HTTPS. EBICS utilise les standards RSA, AES et X.509 pour ses aspects cryptographiques.

EBICS utilise trois paires de clés RSA :

- La clé de signature pour créer la signature électronique du fichier d'ordre.
- La clé de chiffrement, pour transmettre la clé AES ayant servi pour chiffrer le fichier d'ordre (on a donc un chiffrement hybride).
- La clé d'authentification, pour signer le message XML EBICS.

L'implémentation française utilise des certificats X.509 pour la gestion de ces trois paires de clés.

EBICS permet aussi bien de transmettre des fichiers XML comme ceux utilisés par SEPA que les fichiers des anciens formats fixes. Avant d'être transmis, un fichier d'ordre est zippé, chiffré, puis encodé en Base64. Le résultat est ensuite découpé en segments de 1 Mo maximum, et chacun de ces segments est inséré dans un message XML.

La transaction EBICS se fait toujours à l'initiative du client, que ce soit dans le sens du client vers la banque (virements, prélèvements...) ou dans le sens de la banque vers le client (relevés de compte...).

3.5.3 Versions

Deux versions du protocole sont prévues en France :

- EBICS T, destiné à remplacer ETEBAC3
- EBICS TS, avec signature, destiné à remplacer ETEBAC5

3.5.4 Notes et références

[1] (fr) « Migration ETEBAC vers EBICS et SWIFTNET », sur www.cfonb.org (consulté le 28 avril 2011)

3.5.5 Liens externes

- Site officiel du protocole EBICS
- Site de qualification gratuit, par l'éditeur Elcimai Financial Software, du protocole EBICS
- Compendium EBICS
- Le PPI EBICS blog

-  Portail de l'informatique

-  Portail des télécommunications

-  Portail de la finance

3.6 Edificas

Edificas (Échange de données informatisé fiscales, informationnelles, comptables, analytiques et d'audit, et sociales) est une association française regroupant des professionnels de la comptabilité et de l'informatique dont l'objectif est de réaliser des travaux techniques de normalisation pour le compte de ses membres.

Accréditée comme communauté sectorielle comptable par l'organisme de coordination EDIFRANCE, EDIFICAS bénéficie d'une reconnaissance par la Commission européenne et d'un statut européen délivré par l'e-BES et international par l'ONU (CEFACT).

Ses travaux l'ont conduit dans les années 90 à normaliser l'écriture comptable (message ENTREC), le plan comptable (message CHACCO), le grand livre (message LEDGER), la balance comptable (message BALANC) et le reporting (message INFENT) dans le monde EDIFACT. Son objectif à présent est de réaliser la même normalisation dans le monde ebXML (*accounting entry* en cours de test).

3.6.1 Historique

Créée en 1992, l'association EDIFICAS, pôle national de compétence EDI en matière de comptabilité, d'audit et de social, regroupe des partenaires économiques français, privés comme publics, intéressés ou impliqués dans la mise en place de la stratégie EDI, au plan national et européen comme international.

Cette association, constituée à l'initiative de l'Ordre des experts-comptables, a pour objet de promouvoir l'EDI en matière financière, informationnelle, comptable et d'audit, analytique et sociale. À ce titre, EDIFICAS s'aide de normes internationales appelées EDIFACT tout en préparant déjà sa migration vers les standards XML.

À ce jour, EDIFICAS a déjà délivré de nombreuses attestations de conformité concernant EDI-TDFC et EDI-TVA (téléprocédures des déclarations fiscales annuelles et de TVA), étudie une procédure EDI concernant l'Impôt sur le revenu et s'appête à mettre en œuvre la téléprocédure EDI-COMPTA pour faciliter les échanges de messages comptables.

3.6.2 EDI, technique de communication

L'EDI offre une technique de communication qui se propage actuellement dans tous les secteurs de l'économie (agriculture, assurance, automobile, chimie, commerce, comptabilité, construction, finance, informatique, juridique, transport, secteur public, etc.) et également dans toutes les fonctions de l'entreprise (achats, ventes, banques, production, transport, reporting, consolidation, etc.).

Il s'appuie sur une norme internationale qui, par son langage appelé **EDIFACT**, vise à résoudre les problèmes tenant à l'hétérogénéité du parc informatique (**matériel** et **logiciel**) et à l'intégration des nouvelles **technologies de l'information et de la communication**. D'ores et déjà, des ponts sont lancés sur les recommandations **XML** pour assurer la pérennité des services.

La pratique de l'EDI permet de remplacer les systèmes actuels de communication sur support papier ou magnétique, à la fois coûteux et peu efficaces, par des échanges de **messages structurés** entre les différentes **applications informatiques**. Elle favorise l'intégration de la comptabilité existant déjà dans les différentes chaînes fonctionnelles de l'entreprise.

Il ne faut pas confondre EDI, technique de communication, et **EDIFACT**, un métalangage. Il faut par contre opposer EDI et **EFI**. En EDI, on peut trouver des métalangages comme **EDIFACT**, **ebXML** et **XBRL**.

3.6.3 Normes UN/CEFACT Edifact

EDIFACT : Electronic Data Interchange For Administration Commerce and Transport. Cette norme internationale, utilisée pour l'Echange de données informatisé (**EDI**), est la norme **UN/CEFACT** maintenue et coordonnée par le Centre pour la facilitation des procédures et pratiques dans l'administration, le commerce et les transports (**CEFACT**) en collaboration avec **ISO/TC 154-UN/CEFACT**.

La norme **EDIFACT** est utilisée de façon prédominante dans le monde entier. Aux États-Unis, elle est désormais préconisée pour les échanges avec l'administration.

En France, cette norme a été adoptée par la DGI dès 1997 suivant en cela la lettre circulaire du Premier ministre.

Les normes **BALANC** (*balance*), **CHACCO** (*plan de comptes*), **ENTREC** (*écriture comptable*), **INFENT** (*comptes annuels* ou *liasse fiscale* ou *dossier de gestion*) et **LEDGER** (*grand livre*) ont été mises au point dans cet esprit. Elles permettent d'importer (ou d'exporter) des fichiers comptables d'un système informatique à un autre : reprise d'une balance ou en cas de changement de système d'exploitation ou de progiciel de comptabilité.

3.6.4 Normes UN/CEFACT XML

Les normes **UN/CEFACT XML** viennent d'être approuvées pour les messages suivants : **ENTRY** pour l'écriture comptable, **LEDGER** pour le grand livre, **CHART OF ACCOUNTS** pour le plan de comptes, **FINANCIAL REPORTING** pour les comptes annuels et les états financiers et **ACCOUNTING MESSAGE** pour identifier les intervenants dans un interchange, le dossier traité et les paramètres du dossier traité.

En préparation, il reste à normaliser la pièce comptable (**BUNDLE COLLECTION**), message qui servira également pour l'archivage électronique, la liste des journaux utilisés par une entreprise (**ACCOUNTING JOURNAL**) (Prévision 2011).

Ainsi, l'ensemble de la comptabilité et de l'organisation comptable d'une entreprise pourra être immatérialisé en **XML**.

3.6.5 Procédure d'attestation de conformité

Afin de faciliter la phase d'exploitation de la procédure **EDI-TDFC**, notamment en période fiscale, et d'éviter des anomalies de syntaxes **EDIFACT** dans les messages, une procédure obligatoire d'attestation de conformité de la structure des fichiers est mise en place.

Cette procédure donne lieu à la délivrance d'une attestation de conformité émise par **EDIFICAS**. Le répertoire des attestations de conformité délivrées est détenu par **EDIFICAS**. Il est consultable sur son site Web.

3.6.6 Liens externes

- EDIFICAS France
- EDIFICAS Europe
- CLEEP - Comité de Liaison pour les Echanges Electroniques Professionnels
- UN/CEFACT et UN/EDIFACT
- Database EDIFACT Syntax Version 4 Release 1 et Documents EDIFACT Syntax Version 4 Release 1

-  Portail des entreprises
-  Portail de l'économie
-  Portail de la France

Chapitre 4

Protocoles de communications

4.1 Protocole de communication

Dans les réseaux informatiques et les télécommunications, un **protocole de communication** est une spécification de plusieurs règles pour un type de communication particulier.

Initialement, on nommait **protocole** ce qui est utilisé pour communiquer sur une même couche d'abstraction entre deux machines différentes. Par extension de langage, on utilise parfois ce mot aussi aujourd'hui pour désigner les règles de communication entre deux couches sur une même machine.

Les protocoles de communication les plus utilisés sont les **protocoles réseau**.

4.1.1 Concept

Communiquer consiste à transmettre des informations, mais tant que les interlocuteurs ne lui ont pas attribué un sens, il ne s'agit que de données et pas d'information. Les interlocuteurs doivent donc non seulement parler un langage commun mais aussi maîtriser des règles minimales d'émission et de réception des données. C'est le rôle d'un protocole de s'assurer de tout cela. Par exemple, dans le cas d'un appel téléphonique :

1. l'interlocuteur apprend que vous avez quelque chose à transmettre (vous composez son numéro pour faire sonner son combiné) ;
2. il indique qu'il est prêt à recevoir (vous attendez qu'il décroche et dise « Allo ») ;
3. il situe votre communication dans son contexte (« Je suis Marc. Je t'appelle pour la raison suivante... ») ;
4. un éventuel *destinataire final* peut y être identifié (« Peux-tu prévenir Michel que... ») ;
5. le correspondant s'assure d'avoir bien compris le message (« Peux-tu me répéter le nom ? ») ;
6. les procédures d'anomalies sont mises en place (« Je te rappelle si je n'arrive pas à le joindre ») ;
7. les interlocuteurs se mettent d'accord sur la fin de la communication (« Merci de m'avoir prévenu »).

Cette *métacommunication* n'est autre que la mise en œuvre de protocoles.

Mais vous avez déjà implicitement observé un *autre* protocole, avec une autre couche de communication, en attendant d'avoir la tonalité pour composer le numéro de votre correspondant. Et les standards téléphoniques de départ et d'arrivée, pour leur part, se sont coordonnés entre eux aussi : autant de protocoles superposés et qui s'ignorent sans souci entre eux, puisque chacun concerne une couche d'abstraction et son **homologue** à distance.

4.1.2 Exemple d'utilisation

Lorsqu'on est en session X Window avec un ordinateur distant sur une ligne RNIS, la communication est payée au temps. On peut dans la plupart des cas déconnecter la session de bas niveau RNIS au bout de quelques secondes

d'inactivité tout en maintenant la connexion de niveau supérieur **TCP/IP**. Ainsi lorsqu'un message **TCP/IP** est envoyé, le pilote RNIS réétablit la communication en moins de deux secondes, donnant l'illusion d'une continuité de liaison avec des prix bas. Pour **TCP/IP**, la liaison n'aura jamais semblé être coupée.

4.1.3 Voir aussi

Article connexe

- Interopérabilité
-  Portail de l'électricité et de l'électronique
-  Portail de l'informatique
-  Portail des télécommunications

4.2 Réseau à valeur ajoutée

Un **réseau à valeur ajoutée** (en anglais *Value Added Network* ou **VAN** ) est un réseau de télécommunication qui se distingue par le fait qu'il assure certaines fonctions de traitement de l'information en plus de son transport.

Exemples de fonctions de traitement de l'information : messageries vocales, consultations de banques de données, liaisons interbancaires.

-  Portail des télécommunications

4.3 European Network Exchange (ENX) Association

L'Association **European Network Exchange Association** (en abrégé : *Association ENX*) est un regroupement d'associations, de constructeurs et de sous-traitants automobiles européens

4.3.1 Historique

L'association

L'Association ENX, fondée en 2000 ^[1], est une association de droit français selon la loi de 1901. Ses sièges sont situés à Boulogne-Billancourt (France) et Francfort-sur-le-Main (Allemagne).

Les 15 membres de l'association, également représentés au conseil d'administration ENX, sont **AUDI**, **BMW**, **BOSCH**, **CONTINENTAL**, **DAIMLER**, **FORD**, **MAGNA**, **PSA PEUGEOT CITROEN**, **RENAULT**, **VOLKSWAGEN** ainsi que les associations automobiles **ANFAC** (Espagne), **GALIA** (France), **SMMT** (Royaume-Uni) et **VDA** (Allemagne). L'association peut décider d'accueillir de nouveaux membres sur simple demande, leur nombre total étant toutefois limité conformément aux statuts de l'association.

Domaines d'activité

ENX est une association à but non lucratif qui régit le réseau de communication ENX standard au niveau juridique et d'organisation. Ce réseau offre aux entreprises participantes une plate-forme d'échange d'informations et d'initiation de projets préconcurrentiels de coopération dans le domaine des technologies de l'information. La motivation principale des industriels français et allemands lors de la création du standard, était de protéger la propriété intellectuelle tout en réduisant les coûts et en simplifiant les échanges de données dans l'industrie automobile.

Un des bénéfices avancé en faveur de la création d'une « Communauté de Confiance » pour les industriels est que, même si les entreprises protègent leurs propres infrastructures, des problèmes peuvent survenir dans les cas où des solutions de chiffrement ou d'authentification doivent être mises en œuvre entre différentes sociétés et doivent être reconnues réciproquement comme étant dignes de confiance. Les deux parties se retrouvent très souvent dans une impasse, lorsque chacune d'entre elles cherche à mettre en œuvre ses propres mécanismes, si ce n'est plus tôt. Cette situation est bien illustrée par l'exemple du chiffrement des courriels, en cas de conflit de règles de sécurité relatives à l'utilisation d'applications partagées et des milliers d'échanges de données non chiffrées.

Le chiffrement des courriels en est un très bon exemple, avec le conflit de règles de sécurité relatives à l'utilisation conjointe de l'application avec des milliers de connexions non chiffrées. La solution : une infrastructure commune digne de confiance^[2]. FORD cite ainsi l'utilisation du réseau ENX pour communiquer avec les fournisseurs, comme un exemple sur la manière de réaliser des économies considérables par la consolidation et la standardisation^[3].

La mise en œuvre d'exigences industrielles pour la sécurité informatique entre les entreprises est un autre champ d'action. Les sujets suivants en font partie :

- Messagerie sécurisée (entre les entreprises),
- Cloud computing sécurisé (entre les entreprises),
- Protection de la propriété intellectuelle lors de développements collaboratifs (par exemple en utilisant l'ERM (Enterprise Rights Management)).

L'Association ENX est membre du projet ERM.open de l'Association ProSTEP iViP et a déjà pris une part active au projet SP2 précédent avec Adobe, BMW, FH Augsburg, Continental, Daimler Fraunhofer IGD, Microsoft, PROSTEP, Siemens PLM, TU Darmstadt, TAC, Volkswagen et ZF Friedrichshafen^[4].

Le projet SkIdentity, auquel participe l'Association ENX, a été désigné le 1^{er} mars 2011 lors du salon IT CeBIT à Hanovre par le Ministère de l'Économie et de la Technologie comme étant l'un des 12 gagnants du concours technologique organisé par le Ministère, « un cloud computing sûr pour les moyennes entreprises et le secteur public - Trusted Cloud ». Avec la mise en place du programme Trusted Cloud, le Ministère veut encourager « le développement et l'expérimentation de solutions cloud innovantes, sécurisées et juridiquement conformes »^[5].

Présidents de l'Association ENX

Les présidents de l'Association ENX sont

- Clive Johnson, Ford-Werke GmbH (depuis juin 2013)
- Prof. Dr. Armin VORNBERGER (octobre 2005 - mai 2013)
- Hans-Joachim HEISTER, Ford-Werke GmbH (juillet 2001 - octobre 2005)
- Prof. Dr. Gunter ZIMMERMEYER, Verband der Automobilindustrie e.V. (juillet 2000 - juillet 2001)

Les adhésions de l'Association ENX

L'Association ENX est membre des associations et organisations suivantes

- Automotive Industry Action Group (AIAG), Southfield, Michigan
- Bundesverband Informationswirtschaft, Telekommunikation und neue Medien e.V. (BITKOM)
- ProSTEP iViP e.V.

Des adhésions réciproques existent avec ANFAC, GALIA et SMMT

4.3.2 Utilisation du réseau ENX

Scénario d'utilisation

Le réseau de communication (du même nom) de l'industrie automobile européenne se base sur les normes établies par l'Association ENX concernant la sécurité, la disponibilité et l'interopérabilité. Ce réseau industriel garantit un échange sécurisé des données de développement, de contrôle de production et logistiques dans l'industrie automobile européenne.

L'industrie automobile est marquée par une forte coopération internationale et la nécessité pour les entreprises de coordonner des processus étroitement liés qui exigent une coordination précise et des échanges de données continus entre les partenaires. Cela rend nécessaire les « concepts de réseau mondial intégré »^[6]. ENX est décrit comme une plate-forme constituant la base préalable pour de tels modèles de production coopérative^[7].

Une nouvelle orientation a été adoptée fin 2002. L'objectif était d'orienter le développement technique en fonction des exigences des utilisateurs sur une base cohérente, notamment pour les petites et moyennes entreprises^[8]. La mise en œuvre a duré plusieurs années. En juin 2004, les utilisateurs français déploraient l'absence de solutions d'entrée de gamme rentables dans le portfolio de France Télécom^[9].

En mars 2011, plus de 1500 entreprises de l'industrie automobile et d'autres secteurs dans plus de 30 pays utilisent le réseau ENX, disponible dans le monde entier.

Tous les protocoles et applications IP peuvent être utilisés sur le réseau. Le panel va de l'échange de données EDI classique jusqu'à la réalisation de vidéoconférences en passant par l'accès à des bases de données et aux échanges de courriels sécurisés. L'utilisation de protocoles de transfert de données EDI tels qu'OFTP (Odette File Transfer Protocol), OFTP2 et AS2 est répandue sur le réseau ENX. L'OFTP2, développé à partir de 2004, peut également être utilisé sur le réseau Internet public.

Selon la presse spécialisée, certains constructeurs automobiles exigent depuis 2010 l'utilisation d'OFTP2 par Internet, « plusieurs dizaines de milliers de fournisseurs » seraient concernés. Dans ce média accessible à chacun, le transfert de données sensibles devrait entraîner un investissement en sécurité beaucoup plus important, les coûts de changement et d'implémentation sont difficiles à estimer^[10].

Utilisation par les constructeurs automobiles

Liste des constructeurs automobiles qui ont une connexion ENX :

Enregistrement : un pré-requis avant utilisation

L'enregistrement auprès de l'Association ENX est la condition préalable à l'utilisation du réseau ENX^[11]. L'enregistrement peut être demandé directement auprès de l'Association ENX ou de l'un de ses représentants.

Représentativité de l'Association ENX

L'Association ENX est représentée, dans certains pays et dans certains secteurs, par des associations et des organisations industrielles (appelées ENX Business Center). Ces organisations servent entre autres d'interlocuteurs pour les utilisateurs dans la langue du pays concerné, traitent les demandes d'enregistrement et sont responsables des autorisations initiales de nouveaux utilisateurs dans leur secteur de représentation respectif.

Avec ce mode de représentation, l'Association ENX donne la possibilité aux associations industrielles et aux organisations similaires de gérer des groupes d'utilisateurs de manière autonome et indépendante.

-

4.3.3 Fonctionnement du réseau ENX

Fonctionnement du réseau et des connexions

Fonctionnement au travers d'opérateurs de télécommunication certifiés Le réseau ENX satisfait aux exigences de qualité et de sécurité des réseaux d'entreprises, tout en étant également aussi ouvert et flexible aux constructeurs automobiles, aux fournisseurs et à leurs partenaires de développement que le réseau Internet public^[12]. Les échanges de données entre les utilisateurs ENX s'effectuent au moyen d'un VPN (Virtual Private Network) chiffré à travers les réseaux des prestataires de communication certifiés dans ce but par l'Association ENX^[13].

Le premier opérateur certifié était la filiale de Deutsche Telekom, DeTeSystem^[14]. Orange Business Services, Telefonica, Infonet ont suivi, ainsi que Verizon Business en 2007. En 2010, avec ANXeBusiness, BCC et Türk Telekom, ce sont trois autres entreprises qui ont se sont soumises avec succès à la certification ENX. Selon les renseignements de l'Association ENX, Open Systems AG est un opérateur supplémentaire, actuellement en cours de certification.

Les offres des opérateurs certifiés sont interopérables et fournis dans un environnement concurrentiel.

Vue d'ensemble des opérateurs certifiés conformément au standard ENX

-

Processus de certification Selon l'Association ENX, la certification s'effectue en deux étapes^[15]. Au cours de la première étape appelée phase de conception, l'Association ENX vérifie que le modèle d'exploitation ENX entreprise par l'opérateur répond aux spécifications techniques ENX. Lors de la seconde étape, l'opérateur implémente son modèle d'exploitation. En plus de l'organisation interne, la vérification porte également sur le contrôle de l'interopérabilité IPsec. Outre la vérification de l'organisation interne, l'interopérabilité IPsec est également testée dans le laboratoire d'interopérabilité IPsec appelé « ENX IPsec Lab »^[16]. De plus, le chiffrement ENX est mis en œuvre ainsi que la connexion aux opérateurs déjà certifiés, par l'intermédiaire de points de peering privés, appelés "Points d'Interconnexion ENX". Pour finir, l'application et le respect des spécifications ENX sont examinés lors d'un test pilote. Une préparation adéquate par l'opérateur doit permettre de procéder à la certification payante dans un délai de 3 à 4 mois environ.

Fonctionnement des éléments centraux en arrière-plan

À la demande et sous le contrôle de l'Association ENX sont réalisés des services centraux qui servent à simplifier les connexions (« Interconnectivité ») entre les différents opérateurs certifiés et l'interopérabilité des matériels de chiffrement utilisés. Ils comprennent, les points appelés "Points d'interconnexion" ("ENX POI"), le laboratoire d'interopérabilité IPsec ("ENX IPsec Lab") et l'infrastructure à clés publiques ("ENX PKI") au sein du Centre de confiance ENX (ENX Trust Center).

Les "Points d'Interconnexion" sont établis avec une redondance géographique, reliés entre eux et exploités dans les centres informatiques des régions suivantes :

- Région Rhin-Main, Allemagne
- Île-de-France, France
- Côte Est des États-Unis

Ces éléments de fonctionnement centraux ne sont pas visibles pour les utilisateurs. Les clients reçoivent leur propre connexion comprenant un routeur IP, du matériel de chiffrement, du matériel de codage, un chiffrement de bout en bout et ininterrompu de chaque communication et des contrats de service individuels directement auprès de l'opérateur de télécommunication certifié choisi.

4.3.4 Disponibilité mondiale

Réseau JNX et ANXeBusiness en Amérique du Nord

JNX (Japanese Network Exchange) est un réseau de l'industrie automobile japonaise, similaire à l'ENX sur le plan technique et d'organisation. Son contrôle est exercé par le Centre JNX relié aux associations automobiles japonaises JAMA et JAPIA. JNX et ENX ne sont pas liés.

En revanche, entre le standard ENX et le réseau américain ANX développé dès les années 90, il existe des différences techniques, d'organisation et commerciales considérables.

Connexion entre l'Europe et l'Amérique du Nord

ENX en tant que standard commun depuis 2010 Le 26 avril 2010, l'Association ENX et le ANX eBusiness ont annoncé la connexion de leurs réseaux afin de créer un standard mondial dans l'industrie automobile^{[17],[18]}. De cette alliance est né un réseau industriel transatlantique regroupant plus de 1500 entreprises et son fonctionnement effectif a démarré avec la fin de la phase pilote, le 26 mai 2010.

D'après les déclarations concordantes de l'Association ENX et ANX eBusiness Corp., seul le standard ENX est utilisé pour les connexions transatlantiques tant en Europe qu'en Amérique du Nord. Dans leurs publications, ANX et ENX qualifient l'interconnexion comme étant gratuite pour les utilisateurs^[19].

Différences entre ENX et ANX Le réseau Nord Américain, appelé Automotive Network Exchange (ANX), est exploité par ANXeBusiness Corp. Également initié à l'origine par l'industrie automobile et exploité par un consortium, il a, à l'inverse d'ENX, été vendu et exploité par la suite en tant qu'entreprise de services classique à but lucratif^[20].

ANX est un réseau physique dans lequel la disponibilité occupe le premier plan. En général, ANX est basé tout d'abord sur l'exploitation de connexions de lignes fixes avec des garanties de disponibilité élevée. ANX offre, avec le produit supplémentaire « TunnelZ »^[21], une gestion optionnelle du tunnel VPN qui n'est pas utilisée par tous les constructeurs et fournisseurs connectés au réseau. Dans le réseau ANX classique, la gestion des clés de codage s'effectue au moyen de Pre-Shared Key (PSK), tandis que le niveau de chiffrement est limité au DES.

ENX s'établit en tant que Service de Sécurité Managé qui intègre en permanence une gestion de tunnel standardisée, une infrastructure à clés publiques (PKI) basée sur un Trust Center et des mécanismes d'authentification et de chiffrement basés sur réseaux différents (de privé à public).

En fait, il n'y a dans le réseau ANX qu'un seul opérateur face au client, l'entreprise ANXeBusiness elle-même, alors que les services ENX sont proposés par différentes entreprises en concurrence.

Afin de relier malgré tous les réseaux, ANXeBusiness continue d'exploiter son propre réseau séparément et indépendamment d'ENX, mais il fournit à chaque utilisateur ANX qui le souhaite une connexion native ENX, incluant tous les critères de sécurité et de service exigés, via son propre réseau physique. ANX s'est soumis, dans ce but, à la certification et la surveillance par l'Association ENX et fait office d'opérateur certifié ENX^[22].

Synthèse Avec la certification de ANXeBusiness comme opérateur ENX, ENX et ANX utilisent les différences d'organisation précitées entre le consortium industriel à but non lucratif (ENX) d'une part et un opérateur (ANX) d'autre part, afin de connecter les deux réseaux. On ne peut pas pour autant parler d'interopérabilité réciproque étant donné qu'ANX prend en charge le standard ENX. Pour ANX, il est probable que de nouvelles perspectives de marché découlent de l'accès potentiel à tous les utilisateurs ENX. Dans le même temps, on peut supposer que le rapprochement avec ANX facilitera l'entrée sur le marché d'opérateurs ENX supplémentaires aux États-Unis et créera de la concurrence.

4.3.5 Liens

Réseaux de communication de l'industrie automobile (organisations à but non lucratif)

- Association ENX : présente dans le monde entier
- JNX Center : Japon

Informations des opérateurs de télécommunication certifiés conformément au standard ENX (solutions commerciales)

- ANXeBusiness Corp.
- BCC : ENX Connect
- KPN : ENX – Automotive Industry Services
- OPEN SYSTEMS : ENX Global Connect
- NUMLOG : Partenaire expert d'ORANGE BUSINESS SERVICES pour les connexions ENX

- ICDCS : Partenaire expert d'ORANGE BUSINESS SERVICES pour les connexions ENX
- TURK TELEKOM : TT ENX
- T-SYSTEMS : Extranet Solution : Intégrer en toute sécurité les partenaires et fournisseurs
- VERIZON : Certification ENX

Associations membres d'ENX

- ANFAC
- GALIA
- SMMT
- OSD
- VDA

Autres organisations informatiques dans l'industrie automobile

- ODETTE INTERNATIONAL LTD. – Organisation à but non lucratif développant des standards et des solutions pour la logistique et l'EDI
- Consortium SASIG - Consortium à but non lucratif de l'industrie automobile dans le domaine de l'échange de données de développement

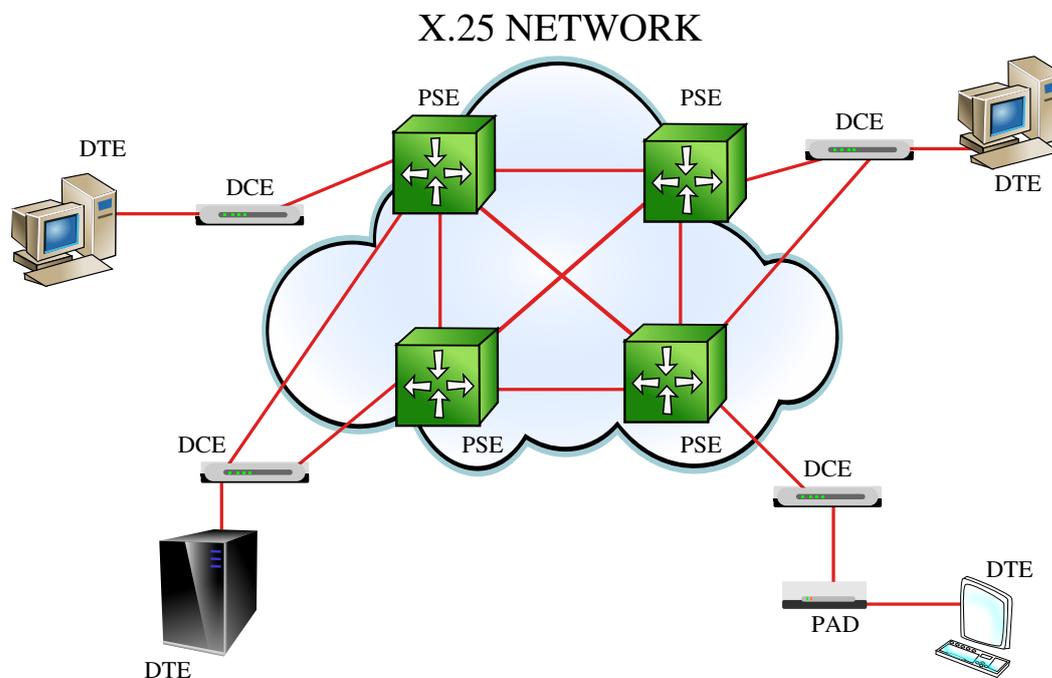
4.3.6 Références

- [1] Europäische ENX-Gesellschaft unter Federführung des VDA gegründet In : Automobilindustrie, 19. Juli 2000
- [2] *Kampf der Beliebigkeit - Combat du hasard.* (PDF) dans : *Automobilindustrie*, septembre 2009 (Interview mit Lennart Oly - interview avec Lennart Oly)
- [3] *Ich erwarte, dass wir in der IT "more-for-less" liefern werden - J'espère que nous livrerons « toujours plus pour moins cher » dans l'informatique.* dans : *Automotive IT*, 10. decembre 2009 (Interview mit Hans-Joachim Heister - interview avec Hans-Joachim Heister)
- [4] ProSTEP iViP Projekt Secure Product Creation Processes - Projet ProSTEP iViP Sécuriser les processus de création de produit (SP²)
- [5] BMWi : Sichere Internet-Dienste - Sicheres Cloud Computing für Mittelstand und öffentlichen Sektor (Trusted Cloud) - BMWi : Services internet sécurisés - Cloud computing sûr pour les moyennes entreprises et le secteur public (Trusted Cloud)
- [6] Jürgen Appel : Kommunikation über Firmengrenzen hinweg - Communiquer au-delà des frontières de l'entreprise dans : *Computerwoche*, 17. septembre 2004
- [7] Helga Biesel : IT im Automobilbau : Spannende Entwicklungen - L'informatique dans l'industrie automobile : des développements passionnants dans : *Computerwoche*, 17. septembre 2004
- [8] ENX wechselt auf die Überholspur - ENX passe à la vitesse supérieure *Computerwoche*, 13. septembre 2002
- [9] Aurélie Barbeaux : L'extranet européen de l'automobile déçoit dans : *Usine Nouvelle*, 10. juin 2004
- [10] An OFTP2 kommt kein Zulieferer vorbei - Aucun fournisseur n'échappe à OFTP2 dans : *Automobilwoche* 5. août 2010
- [11] ENX FAQ Chapitre 5 – Enregistrement
- [12] Zugang für Unbefugte verboten : Porsche setzt auf Branchennetzwerk ENX - Pas d'accès non autorisé : Porsche mise sur le réseau ENX de l'industrie dans : *Computerwoche*, 3. août 2001
- [13] Alexandra Haack dans : Sichere Tür zu Tür Kommunikation - Une communication sécurisée de bout en bout dans : *Automobilwoche*, 9. septembre 2006

- [14] VDA : Weg frei zum European Network Exchange - DeTeSystem und Automobilindustrie unterzeichnen Rahmenvertrag - La voie est libre pour le Réseau d'Echange Européen – DeTeSystem et l'industrie automobile signent un accord cadre
- [15] Comment devenir un « opérateur de communication certifié » ?
- [16] Netzsicherheit der europäischen Automobilindustrie - La sécurité du réseau dans l'industrie automobile européenne dans : Automobilindustrie, 7. fevrier 2003
- [17] Klaus-Dieter Flörecke : *Globaler Standard im Datenaustausch - Standard mondial dans l'échange de données.* dans : *Automobilwoche*, 27. avril 2010
- [18] *ENX und ANX - ENX et ANX.* dans : *Automobilindustrie*, 28. avril 2010
- [19] David Barkholz : *Les réseaux transatlantiques s'allient pour faciliter le partage des données automobiles* dans : *Automotive News*, 26. juillet 2010
- [20] AIAG : *AIAG vend ses actifs et opérations ANX alors que le réseau réalise son projet de devenir une offre commerciale reconnue dans le monde entier.* (PDF)
- [21] Produktbeschreibung ANX TunnelZ - Description du produit ANX TunnelZ
- [22] ANX Certified - ANX certifié

-  Portail des associations
-  Portail de l'automobile
-  Portail du génie mécanique

4.4 X.25



Exemple de réseau X25

X.25 est un protocole de communication normalisé par commutation de paquets en mode point à point offrant de nombreux services^[1]. Après avoir été exploité, en France, par la société Transpac filiale de France Télécom qui en détenait le monopole, c'est sous sa nouvelle dénomination **Orange Business Services** que la commercialisation et la maintenance en a été assuré jusqu'en juin 2012, date de fin d'exploitation technique et commerciale. Cette fermeture a entraîné l'arrêt des services minitel qui s'appuyaient sur ce réseau X25.

4.4.1 Histoire

Fin 1971, face aux progrès du datagramme dans les esprits, Alain Profit, responsable du projet Hermes et chef du groupement ITD (Informatique et Transmission de Données) au CNET, lance l'étude et la réalisation d'un réseau expérimental à commutation par paquets dont il confie la responsabilité à Rémi Després^[2]. Ce projet deviendra six ans plus tard la norme X.25. Entre-temps, il est freiné en 1974 par le choix de Bell Canada d'opter pour le concept de datagramme pour son projet de réseau public Datapac^[3].

4.4.2 Couches

X.25 intègre les trois couches basses du modèle OSI (*Open Systems Interconnection*) :

- niveau 1 : couche physique
- niveau 2 : couche liaison
- niveau 3 : couche réseau
- niveau 3½ : couche de réseau imperméable, dans la version LAPB—2 de H. Fernandez et R.FAURE^[réf. souhaitée]

Cette norme a été établie en 1976 par le CCITT (puis reprise par l'UIT-T) pour les réseaux à commutation de paquets, à la demande de 5 pays qui l'utilisaient pour leurs réseaux publics de communication : Transpac pour la France, EPSS pour la Grande-Bretagne, Datapac pour le Canada, DCS pour la Belgique, et Telenet pour les États-Unis.

X.25 définissait l'interface entre un ETTD (Équipement terminal de traitement de données) et un ETCD (Équipement terminal de circuit de données) pour la transmission de paquets. Il fixait donc les règles de fonctionnement entre un usager du réseau et le réseau lui-même.

X.25 a été supplanté par la croissance des réseaux IP et d'Internet. Cette norme était réputée fiable et sécurisée (il n'y avait aucune perte de donnée grâce aux nombreux contrôles et aux éventuelles retransmissions d'éléments perdus, et les paquets arrivaient dans l'ordre tels qu'ils avaient été émis). Cependant, sa complexité en faisait un protocole consommateur en ressources, mal adapté à la croissance des échanges de données.

Le protocole X.25 était aussi utilisé dans des réseaux tels que le réseau allemand de transmission de données « Datex-P », le réseau de la navigation aérienne français, dans le protocole radio AX.25 (utilisé par les radioamateurs, et notamment pour l'APRS), ainsi que dans beaucoup d'établissements bancaires (protocole ETEBAC) pour gérer les guichets automatiques bancaires. En 2013, l'utilisation de ce protocole est devenu anecdotique (encapsulation de type X.25 over TCP/IP).

4.4.3 Notes et références

[1] (en) « X.25 Virtual Circuits - Transpac in France - Pre-Internet Data Networking »

[2] *Les débuts du réseau public français de commutation de données par paquets : TRANSPAC*, par Guy Pichon, ancien directeur de Transpac, page 46

[3] “Les débuts du réseau public français de commutation de données par paquets : TRANSPAC”, par Guy Pichon, ancien directeur de Transpac, page 49



Portail de l'informatique



Portail des télécommunications

4.5 X.400

X.400 est un protocole de courrier électronique normalisé par l'UIT. Il n'a jamais connu de déploiement réellement significatif. Toutefois il est toujours utilisé aujourd'hui en tant que réseau à valeur ajoutée dans les échanges EDI. Il a fait l'objet de nombreuses polémiques, à propos de sa comparaison avec SMTP, ce dernier l'ayant nettement emporté.

Ses adresses sont de la forme /C=FR/O=Bull/CN=Support.

Le principal héritage de **X.400** est le vocabulaire, MTA (*Mail Transport Agent*) ou MUA (*Mail User Agent*), toujours utilisé aujourd'hui.

En France, Orange Business Services commercialise son réseau X.400 sous le nom d'Atlas 400 ou plus simplement Atlas.

4.5.1 Liens externes

- Page de l'UIT sur X.400

-  Portail des télécommunications

4.6 EDIINT

Le protocole **EDIINT** (EDI via INternet) est un protocole d'échange de données informatisées disponible en deux versions, l'une en communication courriel (AS1, sur SMTP), l'autre en communication web (AS2, sur HTTP). Une troisième version est en préparation basée sur le protocole FTP (AS3, sur FTP).

EDIINT a été conçu par des utilisateurs EDI et pour le transport des données via Internet.

EDIINT est conçu sur la base d'un protocole général (MIME), sécurisé (chiffrement et signature électronique) et répondant aux nécessités des échanges électroniques de documents (accusé de réception / non-réception, reçu de non-répudiation). EDIINT permet de détecter tout changement apporté à un document électronique pendant son transport grâce à sa technologie de chiffrement. La signature électronique, elle aussi chiffrée, valide que l'émetteur est bien celui qu'il dit être. Enfin, l'émetteur reçoit un accusé de réception et de bon déchiffrement, appelé *Message Disposition Notification* (MDN). Cet accusé garantit le bon déroulement de la transaction de bout en bout et est lui aussi signé et chiffré.

4.6.1 Certification Drummond

Le Drummond Group est une entreprise privée américaine qui réalise des certifications pour son propre compte (label "Drummond Certified™") et pour le compte d'associations sectorielles.

- Liste des solutions certifiées Drummond

4.6.2 Autres sources d'information

- Groupe de travail GS1 France

- Projet open source EDIINT AS2

-  Portail des télécommunications

-  Portail de l'informatique

4.7 Simple Mail Transfer Protocol

Simple Mail Transfer Protocol (SMTP), littéralement « protocole simple de transfert de courrier ») est un protocole de communication utilisé pour transférer le courrier électronique (courriel) vers les serveurs de messagerie électronique.

SMTP est un protocole assez simple (comme son nom l'indique). On commence par spécifier l'expéditeur du message, puis le ou les destinataires d'un message, puis, en général après avoir vérifié leur existence, le corps du message est transféré. Il est possible de tester un serveur SMTP en utilisant la commande telnet sur le port 25 d'un serveur distant.

Le SMTP commence à être largement utilisé au début des années 1980. Il est alors un complément à l'**UUCP**, celui-ci étant plus adapté pour le transfert de courriers électroniques entre des machines dont l'interconnexion est intermittente. Le SMTP, de son côté, fonctionne mieux lorsque les machines qui envoient et reçoivent les messages sont interconnectées en permanence.

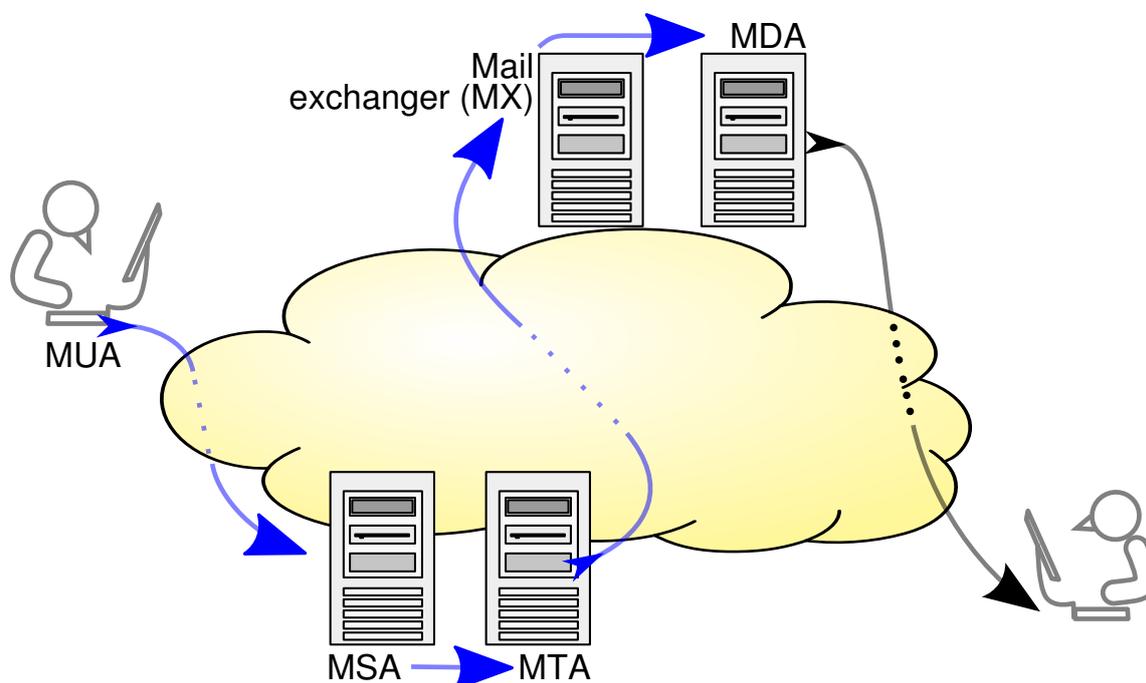
Le logiciel **Sendmail** est l'un des premiers, sinon le premier serveur de messagerie électronique à utiliser SMTP. Depuis, la plupart des clients de messagerie peuvent utiliser SMTP pour envoyer les messages. Certains nouveaux serveurs sont apparus, comme **Postfix**, **Qmail** de Daniel J. Bernstein, **Exim** et **Exchange** de Microsoft.

Comme le protocole utilisait du texte en **ASCII** (7 bits), il ne fonctionnait pas pour l'envoi de n'importe quels octets dans des fichiers binaires. Pour pallier ce problème, des standards comme **MIME** ont été développés pour permettre le codage des fichiers binaires au travers de SMTP. Aujourd'hui, la plupart des serveurs SMTP acceptent le MIME sur 8 bits, ce qui permet de transférer des fichiers binaires presque aussi facilement que du texte simple.

SMTP utilise **TCP** pour le transfert des données.

SMTP ne permet pas de récupérer à distance des courriels arrivés dans une boîte aux lettres sur un serveur. Les standards **Post Office Protocol (POP)** et **IMAP** ont été créés dans ce but.

4.7.1 Principes d'envoi



Principe d'envoi via SMTP

Le transfert de messages entre serveurs de messagerie électronique se fait généralement sur le port 25 qui est le port standard enregistré auprès de l'IANA. Les serveurs utilisent les enregistrements **MX** des serveurs DNS pour acheminer le courrier.

Les clients de messagerie utilisaient aussi le port 25 (smtp) pour soumettre des messages en utilisant le protocole SMTP. Mais la nécessité de mieux contrôler les envois des clients, en particulier par l'authentification, a conduit à

l'attribution du port 587 (submission)^[3].

Les administrateurs de serveur peuvent choisir si les clients utilisent le port TCP 25 (SMTP) ou le port 587 (soumission), tel que formalisé dans la RFC 6409^[4] (RFC 2476 précédemment), pour relayer le courrier sortant vers un serveur de messagerie. Les spécifications et de nombreux serveurs supportent les deux. Bien que certains serveurs prennent en charge le port 465 (historique) pour le SMTP sécurisé en violation des spécifications, il est préférable d'utiliser les ports standard et les commandes ESMTP (Extended SMTP) standard selon la RFC 3207^{[5],[6]}, si une session sécurisée doit être utilisée entre le client et le serveur.

4.7.2 Syntaxe type d'une session ESMTP

Le test par `telnet` mentionné ci-dessus donnerait un *dialogue* du genre (les messages du serveur sont en rouge) :

```
telnet smtp.xxxx.xxxx 25 Connected to smtp.xxxx.xxxx. 220 smtp.xxxx.xxxx SMTP Ready HELO client 250-smtp.xxxx.xxxx
250-PIPELINING 250 8BITMIME MAIL FROM : <auteur@yyyy.yyyy> 250 Sender ok RCPT TO : <destinaire@xxxx.xxxx> 250 Recipient ok. DATA 354 Enter mail, end with "." on a line by itself Subject : Test Corps du
texte . 250 Ok QUIT 221 Closing connection Connection closed by foreign host.
```

Notons que la fin du texte est repérée par un point seul sur sa ligne. Lorsque le texte doit contenir un point seul sur sa ligne, il est donc nécessaire de le doubler (`<CR><LF>..<CR><LF>`).

4.7.3 Les codes retour SMTP

Comme on le constate dans l'exemple ci-dessus, il existe une syntaxe précise pour envoyer les messages et une série de codes retour sur 3 chiffres pour indiquer le statut de la demande. Le premier chiffre du code retour indique le statut global de la demande, les 2 autres chiffres donnent le détail du statut :

- code 2 : la demande a été exécutée sans erreur ;
- code 3 : la demande est en cours d'exécution ;
- code 4 : indique une erreur temporaire ;
- code 5 : la demande n'est pas valide et n'a pas pu être traitée.

Messages les plus courants :

4.7.4 Sécurité et problème du spam

Article détaillé : [Lutte anti-spam](#).

Une des limitations de SMTP vient de l'impossibilité d'authentifier l'expéditeur. Pour ceci, l'extension **SMTP-AUTH** a été définie. Malheureusement, l'impossibilité d'imposer largement SMTP-AUTH a rendu ce protocole impuissant face au phénomène du **spam**.

Le spam est dû à un certain nombre de facteurs dont : l'implémentation de logiciels **Mail Transfer Agent** (MTA) ne respectant pas les standards, les failles de sécurité dans les systèmes d'exploitation autorisant les spammeurs à contrôler à distance des PC utilisateurs pour leur faire envoyer du spam et enfin un manque d'intelligence de certains MTA.

Afin de lutter efficacement contre ce phénomène, il existe deux approches : modifier profondément SMTP (voire le remplacer) ou bien lui adjoindre d'autres protocoles pour combler ses lacunes. Modifier SMTP de manière importante, ou le remplacer complètement, ne paraît pas faisable, à cause de l'importance du réseau de serveurs déjà installé. Malgré tout, des solutions alternatives ont été développées comme **Internet Mail 2000** (en) ou **ePost**.

Une autre approche consiste à créer des systèmes visant à assister les opérations du protocole SMTP. Le groupe de recherche anti-spam (ASRG) de l'IRTF (en), travaille actuellement sur l'authentification des courriers électroniques dans le but de fournir un système flexible, léger, extensible, et évolutif. L'ensemble de ces recherches ont abouti au protocole **MARID** (en) en 2004 ainsi qu'au protocole **DomainKeys Identified Mail** en 2006.

4.7.5 Blocage du port 25 par les fournisseurs d'accès

En 2006, l'AFA recommande aux fournisseurs d'accès internet (FAI) de bloquer les paquets TCP/IP sortant à destination du port 25^[8]. L'idée développée est qu'« un utilisateur résidentiel ne devrait pouvoir émettre ses messages électroniques que via le serveur de son fournisseur de messagerie électronique. »

À l'époque entre 50% et 80% du spam était généré par des ordinateurs infectés^[9].

En France, les principaux FAI ont suivi cette recommandation : Orange bloque le port 25 depuis juin 2007^[10], Free depuis décembre 2006 (c'est une option, le blocage peut être désactivé)^[11], AOL depuis 2003.

La pratique aujourd'hui est la soumission du message par l'utilisateur au serveur de messagerie en utilisant du SMTP authentifié (port 587). Le port 25 sert uniquement aux serveurs SMTP entre eux.

4.7.6 Notes et références

[1] (en) Request for comments n° 821.

[2] (en) Request for comments n° 5321.

[3] (en) « Message Submission for Mail », Request for comments n° 4409, 2006.

[4] (en) Request for comments n° 6409.

[5] (en) Request for comments n° 3207.

[6] La RFC 3207 spécifie uniquement le port 25 bien connu et le port de soumission, qui est le port 587, pour la commande STARTTLS, le précurseur d'une session SMTP chiffrée en utilisant TLS. Il n'est fait aucune mention du port non officiel 465.

[7] (en) Chris Porter, Email Security with Cisco IronPort, Indianapolis, Cisco Press, 2012, 539 p. (ISBN 978-1-58714-292-5), p. 14

[8] Recommandation de l'AFA pour lutter contre le spam

[9] L'AFA publie ses recommandations contre le spam, PC INpact (05/05/2006)

[10] Problème de mails avec Orange ? La cause au port 25 bloqué, PC INpact (28/06/2007)

[11] Free bloque dorénavant le port 25

4.7.7 Voir aussi

Articles connexes

- courrier électronique
- exim
- IMAP
- POP
- LMTP
- Postfix
- qmail
- sendmail

Liens externes

- (en) [RFC 1870](#) – SMTP Service Extension for Message Size Declaration
- (en) [RFC 2505](#) – Anti-Spam Recommendations for SMTP MTAs
- (en) [RFC 2920](#) – SMTP Service Extension for Command Pipelining
- (en) [RFC 3030](#) – SMTP Service Extensions for Transmission of Large and Binary MIME Messages
- (en) [RFC 3207](#) – SMTP Service Extension for Secure SMTP over Transport Layer Security
- (en) [RFC 3461](#) – Simple Mail Transfer Protocol (SMTP) Service Extension for Delivery Status Notifications (DSNs)
- (en) [RFC 3463](#) – Enhanced Mail System Status Codes
- (en) [RFC 3464](#) – An Extensible Message Format for Delivery Status Notifications
- (en) [RFC 3834](#) – Recommendations for Automatic Responses to Electronic Mail
- (en) [RFC 4954](#) – SMTP Service Extension for Authentication
- (en) [RFC 5068](#) – Email Submission Operations : Access and Accountability Requirements
- (en) [RFC 5321](#) – Simple Mail Transfer Protocol
- (en) [RFC 5322](#) – Internet Message Format
- (en) [RFC 6409](#) – Message Submission for Mail
- (en) [RFC 6522](#) – The Multipart/Report Media Type for the Reporting of Mail System Administrative Messages
- (en) [RFC 6530](#) – Overview and Framework for Internationalized Email
- [Le protocole SMTP](#)
- [Tester un serveur SMTP](#)
-  [Portail de l'informatique](#)

4.8 File Transfer Protocol

 Pour les articles homonymes, voir [FTP](#).

File Transfer Protocol (protocole de transfert de fichiers), ou **FTP**, est un protocole de communication destiné à l'échange informatique de fichiers sur un réseau TCP/IP. Il permet, depuis un ordinateur, de copier des fichiers vers un autre ordinateur du réseau, ou encore de supprimer ou de modifier des fichiers sur cet ordinateur. Ce mécanisme de copie est souvent utilisé pour alimenter un site web hébergé chez un tiers.

La variante de FTP protégée par les protocoles SSL ou TLS (SSL étant le prédécesseur de TLS) s'appelle **FTPS**.

FTP obéit à un modèle *client-serveur*, c'est-à-dire qu'une des deux parties, le *client*, envoie des requêtes auxquelles réagit l'autre, appelé *serveur*. En pratique, le serveur est un ordinateur sur lequel fonctionne un logiciel lui-même appelé *serveur FTP*, qui rend publique une arborescence de fichiers similaire à un système de fichiers UNIX. Pour accéder à un serveur FTP, on utilise un logiciel client FTP (possédant une interface graphique ou en ligne de commande).

FTP, qui appartient à la couche application du modèle OSI et du modèle ARPA, utilise une connexion TCP.

Par convention, deux ports sont attribués (*well known ports*) pour les connexions FTP : le port 21 pour les commandes et le port 20 pour les données. Pour le FTPS dit *implicite*, le port conventionnel est le 990.

Ce protocole peut fonctionner avec IPv4 et IPv6.

4.8.1 Histoire

FTP est issu de la RFC 114^[2] créée le 16 avril 1971. Cette spécification fut remplacée par la RFC 765^[3] en juin 1980. Elle fut elle-même rendue obsolète par la RFC 959^[4] en octobre 1985, version finale de la spécification.

Plusieurs RFC viennent compléter cette spécification, comme la RFC 2228^[5] de juin 1997 pour l'ajout d'extensions de sécurité ou la RFC 2428^[6] de septembre 1998 qui ajoute la prise en charge du protocole IPv6 et définit un nouveau type de mode passif.

4.8.2 Interopérabilité

Le protocole FTP ne permet pas toujours d'assurer l'interopérabilité entre plates-formes différentes et régions différentes par une gestion adéquate de l'encodage des noms de fichiers. Seuls les logiciels serveur et client respectant le standard RFC 2640 en donnent la garantie^[7], grâce à l'utilisation de l'encodage UTF-8 et accessoirement d'une nouvelle commande LANG permettant de choisir la langue des messages retournés par le serveur lors de la session FTP. L'encodage UTF-8 permet d'encoder les noms des fichiers provenant de n'importe quel pays, bien qu'un encodage plus spécifique puisse toujours être utilisé localement par le serveur, la conversion vers l'UTF-8 restant à sa discrétion.

4.8.3 Utilisation

Pour accéder à un serveur FTP, on utilise un logiciel (client FTP). Ces logiciels existent avec ligne de commande ou avec une interface graphique. Le standard FTP est si répandu que ces logiciels sont à présent inclus avec les dernières distributions Windows & Linux.

L'utilisation en ligne de commande, sous Windows comme sous Linux, se fait généralement au moyen de la commande *ftp adresse_du_serveur* saisie dans une console. (ici, ftp est le nom du logiciel, et le paramètre le nom du serveur)

Dans les interfaces graphique, comme les navigateurs Web, la forme usuelle est utilisée, à savoir : *ftp://adresse_du_serveur* . Ici, ftp est le nom du protocole, suivi du nom du serveur.

Sous Windows (7, vista), un logiciel client FTP est installé sur la machine, car inclus dans Windows. On peut y accéder à travers les éléments *poste de travail* puis *Ajouter un emplacement réseau*.

4.8.4 Implémentations

Logiciels clients de FTP

Logiciels libres Article détaillé : Logiciel libre.

- GNU inetutils : paquet logiciel GNU contenant un client FTP en ligne de commande
- ftp (en ligne de commande sous Unix/Linux/Windows)
- cURL (en ligne de commande sous Linux/OS X/Windows)
- Cyberduck (pour Mac OS X et Windows)
- FileZilla (pour Linux, Mac OS X et Windows)
- FireFTP (extension pour Firefox)
- gFTP (pour GNOME)
- NcFTP (Windows et systèmes de type UNIX)
- Yafc

Logiciels propriétaires Article détaillé : [Logiciel propriétaire](#).

- FTP Rush (Windows 2000, Windows XP (32/64 bits), Windows 2003 (32/64 bits), Windows Vista (32/64 bits), Windows 7(32/64 bits)

FTP Rush prend en charge les protocoles FTP, FXP, SFTP et TFTP. C'est un partagiciel.

- [WebDrive](#) (Windows 2000, Windows XP (32/64 bits), Windows 2003 (32/64 bits), Windows Vista (32/64 bits), Windows 7(32/64 bits)
- CuteFTP (pour Mac OS X et Windows)
- Fetch (pour Mac OS X)
- [Secure FTP Client \(en\)](#) (pour Java SE)
- Steed (pour Windows)
- [Transmit \(en\)](#) (pour Mac OS X)
- [Yummy FTP \(en\)](#) (pour Mac OS X)

Logiciels serveurs de FTP

Logiciels libres Article détaillé : [Logiciel libre](#).

- FileZilla Server
- vsftpd
- proftpd
- pure-ftpd

Logiciels propriétaires Article détaillé : [Logiciel propriétaire](#).

- Titan FTP Server

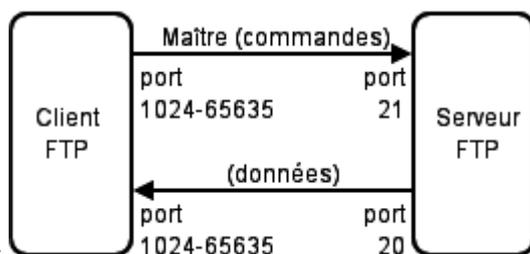
4.8.5 Le protocole

Le protocole utilise deux types de connexions TCP :

- Une connexion de *contrôle* initialisée par le client, vers le serveur (port 21 en général), pour transmettre les commandes de fichiers (transfert, suppression de fichiers, renommage, liste des fichiers...).
- Une connexion de *données* initialisée par le client ou le serveur pour transférer les données requises (contenu des fichiers, liste de fichiers).

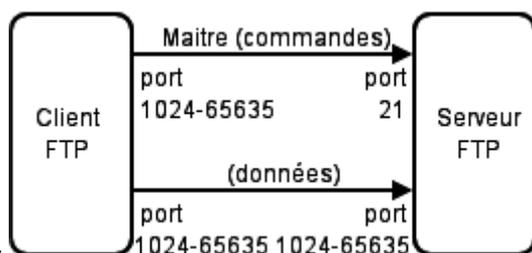
Établissement des connexions

FTP peut s'utiliser de deux façons différentes.



Mode actif :

En **mode actif**, c'est le client FTP qui détermine le port de connexion à utiliser pour permettre le transfert des données. Ainsi, pour que l'échange des données puisse se faire, le serveur FTP initialisera la connexion de son port de données (port 20) vers le port spécifié par le client. Le client devra alors configurer son pare-feu pour autoriser les nouvelles connexions entrantes afin que l'échange des données se fasse. De plus, il peut s'avérer problématique pour les utilisateurs essayant d'accéder à des serveurs FTP lorsque ces utilisateurs sont derrière une passerelle NAT. Étant donnée la façon dont fonctionne le NAT, le serveur FTP lance la connexion de données en se connectant à l'adresse externe de la passerelle NAT sur le port choisi. Certaines passerelles NAT n'ayant pas de correspondance pour le paquet reçu dans la table d'état, le paquet sera ignoré et ne sera pas délivré au client.



Mode passif :

En **mode passif**, le serveur FTP détermine lui-même le port de connexion à utiliser pour permettre le transfert des données (data connexion) et le communique au client. En cas de présence d'un pare-feu devant le serveur, celui-ci devra être configuré pour autoriser la connexion de données. L'avantage de ce mode est que le serveur FTP n'initialise aucune connexion. Ce mode fonctionne sans problème avec des clients derrière une passerelle NAT. Dans les nouvelles implémentations, le client initialise et communique directement par le port 21 du serveur ; cela permet de simplifier les configurations des pare-feu serveur.

La connexion de contrôle

Cette connexion fonctionne en mode texte et est donc aisée à simuler avec Telnet.

Le client envoie une commande sous la forme d'une ligne de texte terminée par un retour à la ligne (CR suivi de LF, soit `\r\n,0D0A` en hexadécimal, ou le caractère 13 et le caractère 10).

Par exemple, la commande suivante demande le téléchargement du fichier "fichier.txt" :

```
RETR fichier.txt
```

N.B. : Les commandes telles que `GET` ou `PUT` ne sont pas reconnues dans le protocole FTP, mais souvent utilisées par les logiciels de client FTP.

À la suite de l'envoi de la commande, le client reçoit une ou plusieurs réponses du serveur. Chaque réponse est précédée d'un code décimal permettant au client FTP de traiter la réponse qui peut comporter une ou plusieurs lignes de texte.

Pour l'exemple précédent, si le serveur trouve le fichier demandé, il envoie au client :

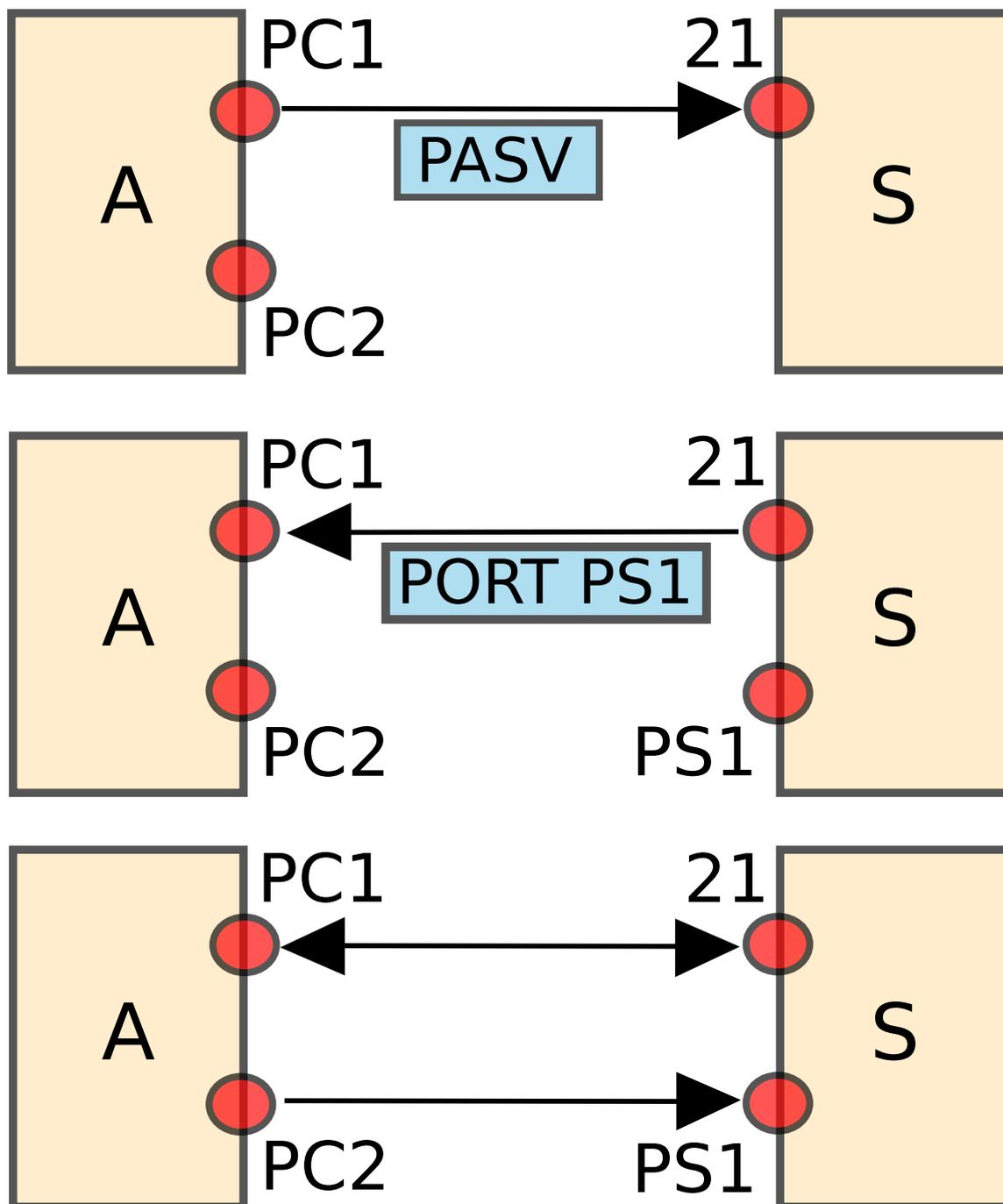
```
150 File status okay ; about to open data connection.
```

Selon ce que le client et le serveur sont convenus, l'un des deux écoute sur le port TCP convenu, et l'autre s'y connecte pour établir la connexion de données. Puis le serveur envoie au client le contenu du fichier demandé, ferme la connexion de données, et envoie la réponse suivante sur la connexion de contrôle :

```
226 Closing data connection.
```

La connexion de contrôle réutilise le protocole de contrôle du terminal NVT introduite au départ pour Telnet, sous une forme simplifiée. Cela signifie en particulier que le code de code décimal 255 sert à introduire des séquences de contrôle de l'affichage.

La connexion de données



Connexion de données

La connexion de données est établie pour la durée de transmission de données (contenu de fichiers, ou liste de fichiers). En général, elle est établie pour le transfert de données d'une seule commande, à moins qu'un autre mode de transmission soit sélectionné et supporté par le serveur.

La commande PASV indique au serveur qu'il doit attendre passivement la connexion en écoutant un port TCP. Le port écouté par le serveur est indiqué dans la réponse :

227 Entering Passive Mode ($h1,h2,h3,h4,p1,p2$).

Où $h1$ à $h4$ sont 4 nombres entiers entre 0 et 255 représentant l'adresse IP du serveur, et $p1$ et $p2$ représentent le port TCP où le serveur attend la connexion, sous la forme de deux entiers entre 0 et 255 ($\text{port_TCP} = p1 * 256 + p2$).

Dans le cas contraire où le client attend la connexion sur un port TCP, il indique sous la même forme le port écouté en envoyant la commande PORT :

PORT *h1,h2,h3,h4,p1,p2*

Si tout se passe bien, le serveur répond :

200 Command okay.

Mode de transfert

Lors du transfert de fichier sur la connexion de données, deux modes peuvent être utilisés :

- Le mode binaire : le fichier est transmis tel quel.
- Le mode ASCII : uniquement destiné aux fichiers texte. Le fichier est examiné et des transformations apportées pour conserver un format correct. Par exemple, la fin de ligne est représentée par le caractère <LF> sur un système UNIX, et par la paire <CR><LF> sous Windows. Une machine Windows recevant un fichier texte par FTP récupère donc finalement un fichier avec des <CR><LF> en mode ASCII et des <LF> en mode binaire. Ce mode a donc ses avantages, mais peut être source de corruption de fichiers (non texte) pendant le transfert si on utilise un client ancien / en ligne de commande, incapable de s'adapter au type de fichier. Il faut alors basculer en mode binaire (en utilisant généralement la commande BIN) avant le transfert, afin de le conserver intact.

Transfert entre deux serveurs

Article détaillé : File eXchange Protocol.

La spécification du protocole FTP^[4] précise qu'il est possible d'effectuer un transfert de fichiers directement entre deux serveurs FTP.

4.8.6 Notes et références

[1] RFC 3659

[2] RFC 114

[3] RFC 765

[4] RFC 959

[5] RFC 2228

[6] RFC 2428

[7] <http://www.linuxfromscratch.org/blfs/view/cvs/introduction/locale-issues.html#locale-wrong-filename-encoding> Beyond Linux® From Scratch - Version 2012-08-12 - Chapter 2. Important Information - Going Beyond BLFS - Locale Related Issues

4.8.7 Voir aussi

Articles connexes

- Protocole réseau passant difficilement les pare-feu
- Clients FTP
- Autres protocoles de transfert de fichiers :
 - Trivial File Transfer Protocol (TFTP)

- Secure File Transfer Protocol (SFTP)
 - File Transfer Protocol over SSL (FTPS)
 - File eXchange Protocol (FXP)
 - CFT (Cross File Transfer)
-
- Variable d'environnement pour la langue

Liens externes

- Le protocole FTP
- (en) RFC 1579 — Firewall-Friendly FTP
- (en)Exemple d'un serveur et d'un client FTP pour les étudiants
-  Portail de l'informatique
-  Portail des télécommunications
-  Portail des réseaux informatiques

4.9 File Transfer Protocol Secure

Le **File Transfer Protocol Secure**, abrégé **FTPS**, est un protocole de communication destiné à l'échange informatique de fichiers sur un réseau TCP/IP, variante du FTP *sécurisé* avec les protocoles SSL ou TLS. Il permet au visiteur de vérifier l'identité du serveur auquel il accède grâce à un *certificat d'authentification*. Il permet également de chiffrer la communication.

4.9.1 Fonctionnement

Il y a deux méthodes pour invoquer le chiffrement SSL/TLS avec FTP, soit de manière explicite, soit implicite.

FTP avec chiffrement explicite

La connexion s'effectue sur le port 21, le port de commande FTP standard, et puis soit :

- la commande "AUTH TLS" (anciennement "AUTH TLS-C") demande au serveur de chiffrer le transfert de commande en TLS, et le chiffrement du transfert de données se fait par la commande "PROT P".
- (non standard) la commande "AUTH SSL" (anciennement "AUTH TLS-P") demande au serveur de chiffrer le transfert de commande et de données en SSL.

Cette approche est compatible avec les serveurs ou clients FTP ne supportant pas le chiffrement SSL/TLS, auquel cas une connexion non chiffrée pourra être utilisée ou bien refusée.

Le schéma d'URI est ftpes :// ou simplement ftp ://

Les RFC 4217 et RFC 2228 décrivent ce mécanisme explicite avec chiffrement TLS.

FTP avec chiffrement implicite

La connexion au serveur se fait sur le port 990 qui est le port de commande et sur lequel la négociation SSL/TLS s'effectue. Le port de données est le 989 et est lui aussi chiffré.

Le schéma d'URI est ftps ://

Cette approche plus ancienne que la méthode explicite n'est pas soutenue par l'IETF.

Cette approche est semblable au fonctionnement du HTTPS décrit dans la RFC 2818 car la négociation SSL/TLS se fait lors de la connexion.

4.9.2 Voir aussi

Articles connexes

- (fr) Protocole réseau passant difficilement les pare-feu
- (en) Liste des serveurs FTP, FTPS (FTP sur SSL/TLS), et SFTP (FTP sur SSH)
- (en) Liste des clients et comparatifs

Lien externe

- (en) Liste des ports TCP et UDP sur le site IANA
- (en) Exemple de support des protocoles ftps :// et ftpes :// dans un client et serveur FTP (Filezilla)
-  Portail des télécommunications
-  Portail de la sécurité informatique

4.10 Secure File Transfer Protocol

Secure File Transfer Protocol (« protocole sécurisé de transfert de fichiers ») est une variante du protocole FTP qui « tunnelise » la session à travers une connexion *Secure Shell (SSH)* pour la sécuriser.

Comme FTP utilise des connexions TCP multiples (ce qui est une exception dans la liste des protocoles TCP/IP), il est particulièrement difficile de réaliser ce *tunneling*.

Le nom de *Secure FTP* ne doit pas être confondu avec d'autres méthodes de sécurisation de FTP, tel que FTPS (SSL) avec certificat d'authentification et chiffrement.

4.10.1 Voir aussi

- Sftp (Unix)
- Protocole réseau passant difficilement les pare-feu
-  Portail de la sécurité informatique

4.11 SSH File Transfer Protocol

 Pour d'autres protocoles de communication que celui-ci, voir SFTP.

Dans le contexte de Secure Shell (SSH), **SFTP** décrit ces deux choses-ci :

- un protocole de communication fonctionnant au-dessus de SSH pour transférer et gérer des fichiers à distance ;
- un programme en ligne de commande qui implémente la partie cliente de ce protocole de communication, comme celui fourni par OpenSSH.

Comparé au précédent protocole scp, le protocole SFTP supporte beaucoup plus d'opérations sur des fichiers à distance. Il se comporte plus comme un protocole de système de fichiers. Il est censé être plus indépendant de la plateforme d'utilisation ; par exemple, avec scp, l'extension des wildcards (*) spécifiés par le client sont à la charge du serveur, qui en fait ce qu'il veut, alors que l'architecture de SFTP évite ce genre de problèmes.

Le programme sftp apporte une interface similaire au programme ftp. Le protocole SFTP n'est pas FTP au-dessus de SSL (visitez File Transfer Protocol over SSL), c'est un nouveau protocole conçu intégralement par le groupe de travail IETF SECSSH. Il n'existe pas de RFC décrivant le protocole SFTP, mais seulement un brouillon (draft).

Certaines implémentations du programme scp utilisent en fait le protocole SFTP à la place du protocole scp.

sftp est souvent associé au protocole (et au programme) SSH-2, parce qu'ils ont été conçus en même temps par le même groupe. Cependant, il est possible de le faire fonctionner sur SSH-1, et certaines implémentations le font.

4.11.1 Logiciels

- OpenSSH - le serveur SSH libre
- Nautilus et Konqueror - les explorateurs de fichier des bureaux GNOME et KDE
- WinSCP - un client SCP/SFTP libre pour Windows
- FileZilla Client - un client FTP/SFTP/FTPS libre multiplate-forme (interface graphique)
- KFTPGrabber un client FTP/SFTP libre sous linux (interface graphique)
- Core-FTP Lite - un client FTP/SFTP
- Yafc - un client FTP/SFTP pour UNIX
- PuTTY - un client SSH libre
- Lftp - un client multiprotocole dont SFTP/SCP/FISH, libre
- Cyberduck - un client FTP/SFTP pour Mac OS X
- Mindterm - un client SSH/SFTP/SCP en Java, donc multiplate-forme
- MySecureShell - solution pour configurer un serveur SFTP
- Total Commander - avec plugin dédié - un client SFTP pour Windows
- Tectia - un client/serveur propriétaire multi-plateforme (interface graphique)

4.11.2 Voir aussi

4.11.3 Liens externes

- (en) Commons VFS du "Apache Jakarta project" - Bibliothèque Java - client SFTP
- (en) Draft IETF sur SFTP

-  Portail de l'informatique

-  Portail de la cryptologie

-  Portail de la sécurité informatique

4.12 Applicability Statement 2

AS2 (*Applicability Statement 2*) est une spécification décrivant une méthode de transport de données électroniques sécurisée et fiable au travers d'Internet, basée sur le protocole HTTP et le standard S/MIME (RFC 3852).

La RFC 4130 décrit cette spécification.

Les données peuvent être en lien avec de l'EDI (Échange de données informatisé) mais peuvent très bien être de tout autre type. AS2 spécifie le mode de connexion, de livraison, de validation et d'acquittement des données. Ce mode de communication enveloppe le message qui est envoyé ensuite par Internet. La sécurité des communications est assurée par des certificats numériques et du chiffrement.

L'implémentation d'AS2 nécessite deux machines, un client et un serveur, reliés tous deux à Internet.

Le client peut lui-même être un serveur pour recevoir des données. Le client envoie des données au serveur (*trading partner*), puis à réception, l'application envoie un acquittement (ou MDN - *Message Disposition Notification*) à l'émetteur.

4.12.1 Avantages

- Élimination ou réduction des coûts liés aux RVA (Réseau à valeur ajoutée).
- Développé pour pousser les données de façon sécurisée et fiable par Internet.
- Connectivité rapide et fiable.
- Le chiffrement assure que seuls l'émetteur et le destinataire peuvent interpréter les données.
- Les signatures numériques assurent l'authentification : seuls les messages des émetteurs autorisés sont acceptés.
- Un algorithme de hachage assure l'intégrité des données en détectant si le document a été altéré pendant le transport.

4.12.2 Inconvénients

- Il faut posséder une adresse publique, une connexion permanente à Internet et un pare-feu.
- Il est impossible d'aller chercher les données à l'extérieur.
- Le coût d'un logiciel permettant d'utiliser AS2 si on n'utilise pas un logiciel open source.
- Il faut gérer les certificats utilisés pour les connexions sécurisées.
- Fonctionne seulement sur des réseaux TCP/IP.

4.12.3 Introduction technique d'AS2

- AS2 crée une enveloppe pour les données, qui est ensuite envoyée par Internet en utilisant les protocoles standards.
- Les données sont transmises par le protocole HTTP, en requête POST, avec ou sans l'utilisation de SSL, à une adresse IP statique.
- Les données peuvent être transmises signées et chiffrées.
- Pas de répudiation des données (Non-répudiation).

4.12.4 Voir aussi

- Échange de données informatisées

4.12.5 Implémentations Open Source

- Hermes Message Service Handler
- m-e-c as2
- OpenAS2
- BabelAS2
- AS2Secure
-  Portail d'Internet

Chapitre 5

Quelques acteurs

5.1 Axway

Axway

Logo d'Axway

Axway est un éditeur de logiciel, qui fournit des solutions progicielles et des services dans les domaines de l'intégration d'applications d'entreprise (IAE), du Business to business (B2B), de la supervision des activités métier (BAM), du pilotage de procédure d'entreprise (BPM) et de l'architecture orientée services (SOA).

5.1.1 Présentation

Axway sert aujourd'hui^[Quand?] 11 000 clients dans 100 pays.

En 2012, il s'agit du sixième éditeur français^[3].

Cette société a été créée en 2001 comme une filiale de Sopra Group, regroupant une grande partie des activités de cette dernière dans le domaine du progiciel. Elle a racheté plusieurs entreprises dans les années qui ont suivi, et en particulier : la société américaine Cyclone Commerce en 2006^[4], un département allemand d'Atos-Origin d'environ 150 personnes en 2007^[5] et la société américaine Tumbleweed en 2008.

Ses deux sièges sont à Phoenix, aux États-Unis, et à Paris. Ses centres de développement sont situés en France, à Annecy et à Paris, en Roumanie à Bucarest, en Bulgarie à Sofia et aux États-Unis à Scottsdale. Elle dispose également de différents centres de support dont un à Noida en Inde.

Le 14 juin 2011, **Axway** se sépare de Sopra Group et est introduit en bourse à Paris^[6].

Le 19 juin 2014, Axway annonce la finalisation de l'acquisition de Systar^[7], une société d'édition de logiciel de supervision des activités métier (BAM). Les 150 employés de Systar, rejoignent alors les 1 750 employés de la société, portant l'effectif du groupe à 1 900 personnes.

5.1.2 Évolution du chiffre d'affaires

La pertinence de cette section est remise en cause, considérez son contenu avec précaution. En discuter ?

- 2001 : 45 M€
- 2002 : 50 M€
- 2002 : 70 M€
- 2003 : 74 M€
- 2004 : 74 M€

- 2005 : 85 M€
- 2006 : 117 M€
- 2007 : 145 M€
- 2008 : 171 M€
- 2009 : 182 M€
- 2010 : 208 M€
- 2011 : 217 M€
- 2012 : 224 M€
- 2013 : 237.5 M€

5.1.3 Notes et références

[1] Chiffres clés sur le site officiel

[2] À propos de Systar

[3] Palmarès des éditeurs de logiciels français 2012 — place actualisée suite à l'introduction en bourse

[4] Le Monde Informatique - Axway s'offre l'Amérique

[5] SICAVonline - Sopra : Axway va acquérir les activités "BtoB" d'Atos Origin en Allemagne

[6] Introduction en bourse d'Axway au NYSE Euronext Paris

[7] Axway : finalisation de l'acquisition de Systar. sur BFMTV.com

5.1.4 Voir aussi

Lien externe

- Site officiel
-  Portail des entreprises
-  Portail de l'informatique
-  Portail de la Savoie

5.2 Generix

Generix Group est un éditeur de logiciels de gestion français coté à la Bourse de Paris.



5.2.1 Historique

La société Generix est créée par Bernard Becquart en 1990 dans le département du Nord, à Villeneuve-d'Ascq. À son introduction en bourse au nouveau marché en 1998, la société et le progiciel de gestion intégré (ERP) portaient volontairement le même nom (Générix).

Elle développa dans un premier temps un module de gestion commerciale, pour l'administration des ventes, des achats et de gestion des stocks. Le logiciel s'est par la suite étoffé avec le support de la comptabilité et de la logistique.

En 2006, cette entreprise, cotée au nouveau marché, se retrouvera en 12^e position au classement^[1] des éditeurs français avec un CA de 46 millions d'euros.

Generix devient Generix Group en 2007, après le rachat d'Influe puis d'Infolog. Son progiciel de gestion intégré est alors rebaptisé *Generix Collaborative Enterprise*. Ce logiciel dont le nom est associé à l'histoire de l'entreprise est finalement cédé en 2014 à Aurea (Austin, Texas)^{[2],[3]}.

5.2.2 Logiciels

Influe

- Gestion partagée des approvisionnements (GPA)
- Échange de données informatisé[EDI]
- WebEDI
- Portail Web (clients, fournisseurs)
- Intégration
- Dématérialisation fiscale

Agil

- gestion des magasins

Infolog

- WMS Gestion d'entrepôt
- TMS Gestion de transport

- GTS Gestion de la traçabilité et suivi des retraits rappel
- GLS Gestion des flux logistiques
- EM Gestion des alertes, des évènements, des indicateurs
- YM Gestion de la cour
- RM Gestion des ressources humaines et matérielles

5.2.3 Notes et références

[1] Le Monde Informatique du 13/04/2007

[2] « Generix finalise la cession de l'activité ERP GCE à Aurea », sur *boursier.com*, 5 mai 2014 (consulté le 31 décembre 2014)

[3] « Villeneuve-d'Ascq : quarante postes supprimés dans l'entreprise de logiciels Aurea », sur *lavoixdunord.fr*, 17 novembre 2014 (consulté le 31 décembre 2014)

5.2.4 Lien externe

- Site officiel
-  Portail des entreprises
-  Portail du logiciel
-  Portail du Nord-Pas-de-Calais

5.3 International Business Machines

 « IBM » redirige ici. Pour les autres significations, voir **IBM** (homonymie).

International Business Machines

Logo d'*International Business Machines Corporation*

International Business Machines Corporation, connue sous l'abréviation **IBM**, est une société multinationale américaine présente dans les domaines du matériel informatique, du logiciel et des services informatiques.

La société est née le 16 juin 1911 de la fusion de la *Computing Scale Company* et de la *Tabulating Machine Company* sous le nom de *Computing Tabulating Recording Company* (CTR). Celle-ci a changé de nom pour devenir *International Business Machines Corporation* le 14 février 1924. On lui prête le surnom de *Big Blue* en référence au bleu sombre, couleur longtemps associée à l'entreprise^[3]. Dans les années 1970 et les années 1980, IBM était la première capitalisation boursière au monde^[4].

5.3.1 Historique

1911-1929

Le 16 juin 1911, plusieurs entreprises ont été fusionnées par Charles Ranlett Flint pour former the Computing-Tabulating-Recording Company (C-T-R) ayant 1300 employés à son siège de New York. Flint recruta Thomas J. Watson, Senior. pour l'aider à diriger l'entreprise en 1914.

Le 14 février 1924, C-T-R fut renommé the International Business Machines Corporation (IBM), afin d'aligner son nom avec l'extension de ses activités.

1930-1979

En 1937, le gouvernement américain déploie l'équipement de tabulation IBM pour suivre les enregistrements de 26 millions de personnes bénéficiaires du *Social Security Act*.

En 1944, IBM met sur le marché *Harvard Mark I*, le premier ordinateur numérique, pouvant effectuer des séquences de calculs complexes.

En 1952, Thomas J. Watson, Junior. devient PDG de l'entreprise.

En 1957, IBM lance le *Fortran* et le disque dur.

En 1961, Thomas J. Watson, Jr. est élu président du conseil d'administration et Albert L. Williams devient PDG de l'entreprise.

En 1964, IBM dévoile son *OS/360*.

À partir de 1967, IBM affronte de nouveaux concurrents qui privilégient l'innovation à outrance.

Depuis 1980

En 1980, l'entreprise lance son projet de *IBM PC* en collaboration avec *Microsoft*.

En 1981, le premier swap de taux d'intérêt en plusieurs devises à recevoir une vaste publicité fut une opération arrangée par *Salomon Brothers* entre IBM et la Banque mondiale.

En 1991, IBM vend *Lexmark* et fonde *IBM Global Services*. Par ailleurs, la société reçoit la Médaille d'or du mérite des beaux-arts du Ministère de l'Éducation, de la Culture et des Sports espagnol pour avoir collaboré sur le projet d'informatisation des Archives Générales des Indes^[5].

En 1992, la première ébauche de smartphone, l'*IBM Simon*, est conçue par IBM.

En 1993, IBM enregistre la plus grosse perte de son histoire : 8 milliards \$^[6].

En 1994, IBM commercialise l'*IBM Simon*.

En 1995, IBM rachète *Lotus Software* pour 3,5 milliards de dollars.

En 1996, IBM rachète *Tivoli Systems* pour un montant de 750 millions de dollars.

En 2001, IBM acquiert la division bases de données d'*Informix Software* pour un montant de 1 milliard de dollars.

En 2002, IBM acquiert l'activité de conseil *PwC consulting* pour 3,5 milliards de dollars.

En 2003, IBM achète *Rational* pour 2,1 milliards de dollars.

En 2005, IBM vend sa division de fabrication d'ordinateurs personnels à *Lenovo* et acquiert l'entreprise *SPSS Inc.* pour 1,2 milliard de dollars et *Ascential Software Corporation*.

En 2006, IBM rachète *Filesnet* et *ISS*.

En 2008, IBM achète *Cognos* pour 5 milliards de dollars, et *Diligent Technologies* (stockage et déduplication)^[7].

En 2009, IBM achète *ILOG*, *SPSS*, *Guardium* et *Lombardi*

En 2010, IBM achète *Sterling Commerce* (en) à *AT&T* pour 1,4 milliard de dollars, *BigFix* (en), *OpenPages* (en), *Unica* et *Netezza* (en).

En 2011, IBM dévoile son programme d'intelligence artificielle *Watson* à *Jeopardy!* et rachète *i2 Limited* (en), *Algorithmics Inc.* (en)^[8] et *Cúram Software* (en)^[9].

En janvier 2014, IBM vend une partie de la branche serveur, à *Lenovo* pour 2,3 milliards de dollars, IBM gardant les activités serveur à haute valeur ajoutée^[10]. En octobre 2014, IBM accepte de payer 1,6 milliard de dollars à *GlobalFoundries*, pour se défaire de sa filiale déficitaire dans les semi-conducteurs^[11].

En fin d'année 2014, IBM annonce la signature de contrats d'infogérance importants avec *ABN Amro*, *WPP* et *Lufthansa*, le premier se montant à plusieurs milliards de dollars^[12] et les deux derniers ayant une valeur chacun de 1,25 milliard de dollars^[13].

5.3.2 Activités

IBM était autrefois très centrée sur la conception et la commercialisation de matériels informatiques et en particulier d'ordinateurs centraux (souvent appelés *mainframes*) auxquels son image est longtemps restée associée. IBM a en effet eu un rôle décisif dans leur développement et leur utilisation par la plupart des grandes organisations publiques ou privées.

1993 a été une année charnière qui a lancé la grande transformation dont la société actuelle est le résultat. Alors même que la communauté financière, inquiète des mauvais résultats, faisait pression pour que le groupe soit éclaté en de multiples sociétés indépendantes libres d'exploiter les différentes lignes de produits, le choix fut pris par le PDG récemment nommé, Louis V. Gerstner, de conserver l'intégrité du groupe. Le pari fait était que par une meilleure gestion du portefeuille d'offres de produits et services, IBM saurait mieux tirer parti de ses forces historiques (à commencer par ses puissantes capacités en recherche et développement) pour proposer des solutions intégrées et non plus seulement des produits isolés.

C'est aussi le moment choisi pour revoir toute l'organisation commerciale. La volonté de pouvoir répondre de manière globale et unifiée aux demandes émanant des grands clients internationaux.

C'est encore à ce moment qu'IBM se lance de manière très volontariste dans les activités de services avec en particulier les débuts de l'infogérance. Le développement de ces activités de services a notamment permis de recycler une partie des employés dont les postes avaient été supprimés à la suite de la réorganisation.

En 2009, la contribution au bénéfice des différentes activités se décompose comme suit :

- Matériels = 7 %
- Activités de financement = 9 %
- Logiciels = 42 %
- Services = 42 %

Les services représentent désormais un peu plus de la moitié du chiffre d'affaires. Ceci témoigne de la profonde transformation opérée par IBM depuis une dizaine d'années.

Depuis 2002 et l'acquisition de la branche conseil de PricewaterhouseCoopers, IBM est devenu la première entité de conseil dans le monde entier.

Matériel

L'histoire de la division Matériel d'IBM commence en 1886 lorsque Docteur Herman Hollerith mène les premiers tests de son système de tabulation en enregistrant et calculant des statistiques vitales destinées au département de santé de Baltimore au Maryland. Depuis cette date, IBM n'a eu de cesse d'innover dans le domaine du matériel : processeurs, serveurs, réseaux et solutions de calcul et de stockage. Aujourd'hui, la division Matériel d'IBM (appelée *Systems & Technology Group*) propose une offre complète de matériels professionnels, destinés à satisfaire tous les besoins informatiques, de la TPE à l'entreprise du CAC 40. La stratégie de la division est de construire pour ses clients une infrastructure dynamique, qui s'adapte à leurs besoins métiers. La mission de la division Matériels d'IBM couvre tous les aspects du cycle de vie des matériels informatiques, depuis la vente des composants et systèmes jusqu'au recyclage de ces systèmes en passant par l'intégration et le déploiement.

La gamme de produits de la division Matériel d'IBM se décompose en 6 grandes familles :

- **System z** (grands systèmes ou mainframes) : ligne de produits adaptée pour la continuité des affaires, l'optimisation de la charge de travail (workload) et l'intégration de processus d'entreprise.
- **Power** IBM Power Architecture est la famille de processeurs RISC construits par IBM. On les retrouve aussi bien dans les consoles de jeux que dans les plus grands super calculateurs. Power Systems a redéfini la façon dont l'infrastructure informatique se met au service des objectifs métier.
- **System x et BladeCenter** : Dotés des processeurs Intel et AMD Opteron hautes performances, les serveurs IBM System x permettent aux entreprises de toute taille de mettre en place des infrastructures flexibles et rentables.



Disques IBM2314 et perforateur de cartes 2540 (vers 1968)

- **Solutions de Stockage** : IBM offre un ensemble de solutions matériels (disques, bandes, réseaux) et logiciels pour maîtriser la forte croissance des données des entreprises.
- **Terminaux point de vente et bornes libre service** : IBM est aujourd'hui le leader mondial des terminaux points de vente vendus et de parcs installés dans le monde.
- **Offres réseaux** : IBM commercialise aussi du matériel de connexion réseau, adapté à l'évolution de l'infrastructure des entreprises.

Principales contributions :

- Construction et maintenance de *mainframes* et de gros serveurs (qui possèdent nettement plus de RAM et d'espace de stockage que les ordinateurs grand public).
- 47 % de la recherche et développement d'IBM est consacrée au matériel.
- Fabrication de disques durs. Technologie inventée par IBM en 1956. Cette activité a cependant été cédée à Hitachi en 2002 quand IBM s'est intéressée à millipede.
- Recherche en nanotechnologie (entre autres : millipede).
- IBM est l'un des deux architectes et fondateurs, et le promoteur principal, des processeurs Power utilisés dans la série Power Systems d'IBM (UNIX et i5), dans les anciens Macintosh d'Apple (gamme PowerPC) et en informatique embarquée. En 2005, Apple annonça son intention de migrer à des processeurs x86 au moment même où IBM annonçait son nouveau processeur, nommé CELL, conçu pour la console de jeu PlayStation 3. Auparavant la société avait déjà fourni des variantes de son architecture PowerPC pour les consoles de Microsoft (Xbox 360) et Nintendo (Game Cube et Wii).
- Concepteur et assembleur du *Roadrunner*, l'ordinateur le plus rapide au monde (en date du 11 juin 2008).
- Fabrication de semi-conducteurs (puces électroniques) avec notamment un service de fonderie pour des sociétés fabless de semi-conducteurs.

- En 2005, IBM Microelectronics pointe à la 19^e place des 20 plus grands fabricants de semi-conducteurs (sans tenir compte de son service de fonderie).^[réf. souhaitée]
- IBM a introduit en juin 2009 CloudBurst 1.1, premier système de Cloud Computing intégré prêt à l'emploi, à base de BladeCenter destiné aux environnements de développement et de test.

Logiciels



Un bâtiment IBM de R&D à Haïfa en Israël

Avec un chiffre d'affaires mondial de 22,089 milliards de \$ et 21 % de parts de marché en 2008, la branche logicielle d'IBM (IBM Software Group) est le premier fournisseur de solutions logicielles d'infrastructure.

La branche logicielle d'IBM dans le monde, c'est :

- 50 000 professionnels, dont 26 000 développeurs, 17 000 ingénieurs commerciaux et ingénieurs avant-vente.
- Une part de 40 % des 5 896 brevets déposés par IBM aux États-Unis en 2010^[14].
- Un total de 80 laboratoires de R&D et de plus de 40 centres d'innovation,
- Une représentation dans plus de 170 pays.
- Un écosystème de plus de 100 000 partenaires

Depuis sa création en 1995, IBM Software Group construit une gamme complète de logiciels d'infrastructure ouverts, évolutifs, interconnectables et adaptés aux problématiques sectorielles^[15].

Le portefeuille de technologies logicielles s'articule autour de cinq marques :

- **Information Management**, pour la gestion de l'information à la demande, propose des solutions intégrées dans les domaines du décisionnel, de l'ILM (Information Life cycle management) et du MDM (master data management).

- Lotus, première marque historique de IBM Software, propose des solutions de communication et de collaboration unifiées ainsi que des solutions de portail d'entreprise.
- Rational Rose répond aux problématiques de développements, de tests logiciels, de gouvernance de projets.
- Tivoli, pour l'administration des systèmes d'information, fournit des solutions intégrées pour améliorer la qualité de service et les performances de l'infrastructure et des applications.
- WebSphere, pour la gestion et l'intégration des applications, propose des solutions intégrées autour de l'ESB (Enterprise Service Bus) et du BPM (*Business Process Management* ou pilotage des processus métier) pour mettre en œuvre des architectures orientées services (SOA).

Par ailleurs IBM Software Group propose des solutions de PLM (Product Lifecycle Management) en partenariat avec Dassault Systèmes. Les solutions CATIA, ENOVIA, SMARTEAM permettent de concevoir, analyser et gérer les produits tout au long de leur cycle de vie, de leur conception à leur retrait du marché.

Principales contributions :

- Bases de données (Edgar F. Codd pour les bases de données relationnelles en 1970)
- Construite autour du noyau des bases de données (DB2), la branche Logiciels (Software Group en jargon interne) a été principalement constituée par croissance externe. Ses produits, peu connus du grand public (Tivoli, Websphere, Lotus, DB2, Rational) mais appréciés des professionnels, sont soit des « couches » intermédiaires (*middleware*) entre les logiciels applicatifs (Microsoft Office, SAP...) et les systèmes d'exploitation, soit des « suites professionnelles » visant le marché de la R&D (Rational Software).
- Le système d'exploitation de type UNIX AIX développé à partir de 1986 avec la contribution de Bull.

En 2012, l'activité logicielle a généré près de 11 milliards de dollars de bénéfices avant impôts, un résultat multiplié par trois en dix ans^[16].

Services

- Bureau d'IBM dans Rochester (Minnesota).
- IBM Avenida de América Building à Madrid en (Espagne).
- Bureau d'IBM à New York conçu par IM Pei.
- IBM Japan Makuhari Centre Technique, conçu par Yoshio.T.
- IBM Haifa Research Lab à Haifa en (Israël).

La branche services (*IBM Global Services* ou IGS) représente la moitié du chiffre d'affaires d'IBM en 2008, mais l'essentiel de la marge opérationnelle provient encore du matériel et du logiciel, sur lesquels IBM dispose de positions dominantes discrètes mais fructueuses.

IBM Global Services couvre deux grand domaines activités :

- Le conseil métier *Global Business Services* ou GBS qui intervient sur les couches organisationnelles et les systèmes d'information de l'entreprise pour en améliorer l'efficacité opérationnelle,
- les services pour les infrastructures informatiques *Global Technology Services* ou GTS qui regroupe un ensemble d'offres de services à forte valeur ajoutée permettant d'accompagner les entreprises dans leur transformation et de répondre à leurs nouveaux enjeux mondiaux.

GTS se différencie sur quatre segments d'activités principaux :

- Les services portant sur les infrastructures informatiques *Integrated Technology Services* ou ITS : ensemble de services d'infrastructure aidant les entreprises à optimiser et tirer pleinement parti de leurs infrastructures informatiques.

- Les services de maintenance *Maintenance and Technical Support Services* ou MTS : maintenance matériel et logiciel informatique avec actions préventives et correctives pour assurer une disponibilité optimale du matériel informatique par une approche multi-constructeur.
- Les services d'infogérance *Strategic Outsourcing* ou SO, qui propose des solutions d'externalisation de la fonction Informatique (*IT outsourcing*) comme l'hébergement de serveurs intranet et extranet, la gestion de la sécurité, du stockage et des sauvegardes, ou encore l'externalisation d'applications de messageries ou de progiciels de gestion intégrés.
- Les services de transformation des processus métier *Managed Business Process Services* ou MBPS : marché émergent d'externalisation des processus métiers de Ressources Humaines ou de back office, comme la comptabilité, la paie ou les achats de commodités.

La stratégie d'IBM dans les services est de s'implanter sur les marchés de services de masse en s'affranchissant de toute connotation technologique. IBM *Global Business Services* s'appuie sur la notoriété de la marque - qui dispose d'un fort capital de confiance - et sur la puissance de feu de son organisation pour faire valoir sa position de leader sur le marché des services informatiques^[17] : dotée du capital technologique et intellectuel pour répondre à ses propres besoins informatiques internes pour gérer ses 300 000 employés dans le monde, IBM peut et sait gérer les besoins informatiques de toute entreprise quelle que soit sa taille.

Plus de 42 % des 380 000 employés d'IBM sont désormais mobiles et gèrent quand, où et comment ils font leur travail^[18].

Communautés

En 2001, IBM a lancé la fondation Eclipse chargée du développement de technologies open source.

En 2007, IBM a mis en place le projet de modèle d'entreprise collaborative *BlueIQ* ayant comme public cible les 16 000 commerciaux n'ayant pas l'habitude de partager leurs informations et leurs contacts. En raison du succès rapide, le programme a été étendu à 400 000 collaborateurs IBM supplémentaires^[19]. L'équipe BlueIQ composée de 8 collaborateurs a identifié 20 tâches récurrentes à accomplir à l'aide des réseaux sociaux. Ensuite, elle a créé un communauté de 1 600 membres volontaires provenant de 50 pays pour promouvoir le réseau social d'entreprise^[20].

En avril 2009, IBM a développé le réseau social du site *developerWorks* avec *Lotus Connections*^[21]. En 2011, il réunit 600 000 profils, totalise 4 millions de visiteurs uniques par mois qui ont accès à une bibliothèque de 30000 articles, podcasts et tutoriels^[22]. D'après sa directrice Alice Chou, « My developerWorks permet aux entreprises, aux start-ups et aux partenaires de collaborer. De plus, la plateforme permet à IBM d'économiser d'environ 100 millions de dollars par an en matière d'assistance aux utilisateurs. » IBM souhaite créer un système d'attribution de récompenses par le développement de la notoriété afin de soutenir la communauté *developerWorks* ; cette dernière devant suivre les règles IBM de conduite des affaires^[23] et de *social computing (en)*^[24].

Au niveau mondial, IBM anime une communauté de 100 000 partenaires qui génèrent près de 35 % de son chiffre d'affaires^[25] ; 60 000 d'entre eux sont également éditeurs de logiciels^[26].

IBM a développé trois niveaux de partenariat pour commercialiser ses offres matérielles, logicielles et de services^[27] :

- *Business Partner Member* (engagement minimal et gratuit sous condition d'accepter les termes du contrat PartnerWorld)
- 7 800 *Business Partner Advanced*^[28] (engagement à entretenir avec IBM une relation commerciale fructueuse qui donne l'accès aux rapports de veille stratégique d'IBM, à l'aide au financement d'événements et aux feuilles de route pour le développement de produits)
- 7 300 *Business Partner Premier* (collaboration active avec IBM dans le but de proposer des solutions innovantes qui engendrent la désignation d'un responsable chargé de la gestion de leur relation avec IBM)

Depuis novembre 2010, IBM a mis en place trois nouveaux labels sectoriels pour ses partenaires^[29]. Par ailleurs, IBM organise son offre Logiciel en tant que service dans le cadre du Club Alliances^[30]. Loïc Simon estime que « 20 % des 500 plus grosses entreprises mondiales non-informatiques vont devenir des fournisseurs de services de cloud d'ici 2015 ».



Authorized Software Value Plus Enterprise Content Management

Logo des partenaires commerciaux IBM (charte graphique de novembre 2010)

Financement

Numéro 1 mondial du financement informatique, IBM Global Financing (IGF) bénéficie de 30 années d'expérience dans le financement de projets informatiques et solutions métiers. IGF est présent dans 50 pays et compte plus de 125 000 clients.

IGF propose :

- une palette d'offres de financement adaptées aux besoins des PME/PMI comme à ceux des grandes entreprises,
- des solutions de financement informatique incluant des matériels, des logiciels et / ou des prestations de services, fournis par IBM ou par d'autres prestataires,
- un savoir-faire dans la gestion du cycle de vie complet des matériels en location : reconditionnement, recyclage et effacement des données, mise au rebut du matériel selon les normes environnementales.

En 2009, IBM s'est associé à OSEO, établissement public qui soutient la croissance et l'innovation des entreprises. Ce partenariat permet à IBM de renforcer son aide aux PME en temps de crise en s'appuyant sur l'expertise d'OSEO en termes de financement des entreprises. Il s'inscrit dans le programme mondial de financement d'IBM qui a décidé en mai 2009 de débloquer 1,5 milliard d'euros pour stimuler les projets d'investissement informatique des entreprises en Europe. Le groupe IBM devient l'une des premières entreprises du secteur industriel français à être agréée par OSEO pour la garantie des financements octroyés.

Principaux apports

- Disque magnétique (RAMAC), 1956

- Langage Fortran, 1957
- Mémoire dynamique (DRAM), 1966
- Base de données relationnelle, 1970
- IBM 1750, 2750 and 3750 Switching Systems (en), 1970^[31]
- Code universel des produits, 1973
- Géométrie fractale, 1975
- Data Encryption Standard (DES), 1977
- Architecture RISC, 1980
- IBM PC basé sur un système d'exploitation Microsoft, 1981^[32]
- Microscope à effet tunnel, 1986, Nobel
- Supraconductivité à « moyenne » température, 1987, Nobel

En douze ans, IBM a déposé 29 021 brevets divers. En 2005, IBM a déposé 2 941 brevets.

Réseau IBM en France

La division française d'IBM a été créée dès 1914 et comptait 11 000 employés en 2011, dont 91 % de cadres à travers 19 implantations. IBM France évalue son écosystème de près de 40 000 personnes comprenant 1 500 sous-traitants et 3 000 PME^[33]. Depuis novembre 2009, le nouveau siège social d'IBM France accueille 4 000 collaborateurs dans un bâtiment HQE (Haute Qualité Environnementale) à Bois-Colombes, dans le quartier de Bécon-les-Bruyères, à quelques kilomètres de Paris.

Le Président d'IBM France est Alain Bénichou depuis janvier 2010^[34]. À son initiative, IBM France signe deux conventions de mécénat de compétences en 2011 : la première avec Jean-François Mattei, Président de la Croix-Rouge française, pour la mise en place du logiciel de gestion de l'aide alimentaire Aïda et la cocréation des Maisons Numériques de la Croix-Rouge^[35] ; la seconde convention avec Martin Hirsch pour accompagner l'Agence du Service Civique dans son développement^[36].

En 2011, les principaux grossistes spécialisés dans la vente de licence logiciel IBM pour la France sont Azlan, Arrow ECS et BestWare. En 2010, IBM France était la première SSII française en termes de chiffre d'affaires^[37].

En octobre 2014, IBM ouvre un *data center Cloud* IBM SoftLayer à Paris^[38].

Recherche et développement

Cette section est vide, insuffisamment détaillée ou incomplète. Votre aide est la bienvenue !

Depuis 18 ans, IBM est l'entreprise qui dépose le plus de brevets au monde^[39]. IBM déclare consacrer annuellement et de façon continue un budget de 6 milliards de dollars en R&D. En 2011, 6 186 brevets ont été déposés par IBM Research.

En juillet 2014 IBM annonce un investissement de 3 milliards de dollars pour repousser les limites technologiques rencontrées actuellement par les puces en silicium ^[40] : ordinateurs quantiques, nanotubes de carbone, nouvelles technologies de mémoire, ou puces photoniques sur silicium sont abordées parmi les nouvelles pistes de réflexion.

Brevets et logiciel libre

En octobre 2000, Louis Gerstner annonce qu'IBM va investir 1 milliard de dollars dans Linux^[41]. IBM a délivré 500 brevets pour la communauté open source en 2005^[42].

Arrivé au sixième rang des entreprises contributrices au noyau Linux, IBM a annoncé en septembre 2013 son projet d'investir à nouveau 1 milliard de dollars dans les technologies Linux et open source pour ses serveurs Power Systems^[43] sur les quatre à cinq prochaines années^[44].

De 1993 à 2012, IBM a enregistré plus de 67 000 brevets déposés aux États-Unis. IBM y reste le premier « inventeur » depuis 20 ans avec 6 478 brevets déposés en 2012^[45].



Siège social d'IBM France à Bois-Colombes

5.3.3 Culture d'entreprise

La compagnie a souvent été décrite comme possédant une culture fortement tournée vers les affaires. Ses cadres et directeurs généraux ont traditionnellement été sélectionnés parmi ses représentants. Ce recrutement aurait, par ailleurs, souvent été motivé par l'objectif de concrétiser des perspectives de ventes à des clients importants.

IBM est probablement l'une des structures qui a développé le plus fortement et précocement le principe de culture

d'entreprise^[46]. Durant la majeure partie du XX^e siècle, l'uniforme, constitué d'un costume bleu, d'une chemise blanche et d'une cravate foncée, y était de rigueur pour les commerciaux^[47]. Ces codes ont commencé à s'assouplir à la fin des années 1970, puis l'ont été considérablement dans les années 1990. À ses débuts, IBM faisait aussi chanter à ses salariés des hymnes à la gloire de l'entreprise^{[48],[49]}. De nos jours, l'habillement et l'attitude de ses employés ne diffèrent pas particulièrement de ceux d'autres grandes entreprises.

En 2003, la culture d'entreprise d'IBM a évolué avec le projet de réécrire ses « valeurs d'entreprise » pendant 3 jours avec 50 000 employés ; son résultat a été formulé en trois principes : « être dédié au succès de chaque client », « se focaliser sur les innovations qui comptent – pour l'entreprise et pour le monde », « donner confiance et responsabilité personnelle dans toutes les relations »^[50].

5.3.4 Présidence de la société

IBM a été dirigée depuis sa création par neuf PDG qui ont chacun durablement marqué la Compagnie.

George Winthrop Fairchild : le père fondateur

Président de l'*International Time Recording Company* depuis 1900, George Winthrop Fairchild la fit fusionner en 1911 avec la *Tabulating Machine Company* fondée par Herman Hollerith pour former la *Computing-Tabulating-Recording Company* ; il la renomma en **International Business Machines** en 1924, peu avant sa mort.

Thomas J. Watson Sr : la mécanographie

Sous la présidence de Thomas J. Watson Sr jusqu'en 1956, la société se développa dans les années 1930, grâce aux brevets de mécanographie sur la carte perforée Hollerith.

Thomas J. Watson Jr : les *mainframes*

En 1954, IBM lance sur le marché le modèle 650, premier ordinateur *mainframe*^[51] produit en grande série : il coûtait un demi-million de dollars, occupait plusieurs mètres cubes et était doté d'une mémoire vive de 2 000 « mots » (2 kilooctets). IBM France forgera pour l'occasion le néologisme ordinateur.

T. Vincent Learson

T. Vincent Learson (1912-1996) prit le relais lors du départ de Watson Jr. Ce dernier devint président du comité de direction avant d'être nommé ambassadeur des États-Unis en Union soviétique (1979).

Frank T. Cary

Rapidement, c'est Frank T. Cary qui est élu PDG (1973). C'est à lui qu'échoit la double tâche :

- de développer l'informatique distribuée de plus en plus réclamée par les clients (systèmes 32, 34, 36, 8100...)
- de faire passer aussi IBM à la phase de la micro-informatique. Après quelques machines de succès divers (5100, Système 23 Datamaster, Visiotexte...), il lance dans le plus grand secret le projet *Acorn*, qui deviendra en trois ans le PC.

Il fait *relooker* légèrement le logo, et commence à rompre avec la traditionnelle politique de location seule : désormais, IBM vend aussi ses machines, ce qui provoque un afflux d'argent frais. Pendant quelque temps, à l'inverse de ce qui se passe habituellement dans l'industrie, IBM n'a plus de *frais financiers*, mais des *bénéfices financiers*.

Le procès *antitrust* qui pèse sur IBM laissant craindre une décision de démantèlement de la compagnie, il prend les devants en la réorganisant en deux grands pôles : *mainframes* et *systèmes de grande diffusion*. Le démantèlement n'aura en fin de compte pas lieu.

John R. Opel

Cette section est vide, insuffisamment détaillée ou incomplète. **Votre aide est la bienvenue !**

John F. Akers (1985 - 1993)

John Akers gère le succès du PC et du PC/AT, et introduit même le PC/RT, mais - impressionné par le succès du Macintosh qui constitue alors un système fermé - il décide de *fermer* à son tour le standard PC en lançant en 1987 le PS/2 : nouveau BIOS incompatible, bus assujéti à des droits d'usage, apparence déposée, OS/2 non finalisé encore : *cette décision se révélera la plus catastrophique jamais prise par IBM*, qui - sans tout de suite le comprendre - abandonne ainsi *de facto* la continuité du standard à Compaq et aux clones. La page où IBM menait la danse avec 21 % du marché de l'ordinateur personnel est désormais tournée.

- Un jugement sévère sur John Akers (1991)

En revanche, la filière du Système 38 devenu AS/400 (machine de milieu de gamme introduite en février 1988) est gérée sans la moindre faute : cette machine au début secondaire, qui succède au système 38 et s'inspire donc du projet FS, sera à la fin des années Akers devenue un formidable succès avec plus de 600 000 exemplaires, dont la moitié à des entreprises ne possédant pas de service informatique en propre.

Cela ne compense pas le manque à gagner sur le PC et les *mainframes*, et IBM connaît alors le premier déficit financier de son histoire.

Louis V. Gerstner, Jr. (1993 - 2002) : des produits vers les services

- L'entreprise devient fortement déficitaire en 1991. Pour la première fois de son histoire, elle confie son destin non plus à un de ses membres, mais à un de ses anciens clients, Louis Gerstner (« Lou »). Sa décision fondamentale a été de ne pas céder aux demandes des analystes et marchés financiers qui réclamaient qu'IBM soit découpé en multiples sociétés indépendantes. Il a au contraire affirmé sa conviction que la valeur ajoutée d'un portefeuille intégré d'activités (matériels, logiciels et services) est supérieure à la somme de ces activités prises séparément. Celui-ci recentre la société, lance un emprunt à *cent ans* (totalement souscrit par le public en quelques jours), en modernise l'infrastructure (qui restait très centrée sur les *mainframes*) pour y généraliser Lotus Notes (plutôt que d'en acheter des licences, il acheta directement la société !), et la rend à nouveau profitable - mais amputée d'une bonne partie de son patrimoine immobilier. Le plan produit a également été drastiquement nettoyé : des milliers de produits sont abandonnés ou revendus (Lexmark est l'ancienne division imprimantes d'IBM), et le développement du système d'exploitation OS/2, concurrent de Windows, est arrêté.
- IBM a développé en premier (2001) le concept d'*e-business on demand* qui peut se définir comme étant la mobilisation des ressources informatiques en fonction de l'intensité de l'activité de l'entreprise cliente. La facturation se fait alors en fonction exacte de la consommation, comme pour de l'électricité. Ce concept, lié à celui de *grid computing*, cherche plus à créer le besoin qu'à répondre à une attente précise auprès des grands groupes. En bonne santé financière et donc capable de supporter le financement de ce genre d'opération, IBM met ainsi la pression sur ses concurrents, selon une stratégie qui n'est pas sans rappeler le programme militaire *Strategic Defense Initiative* (SDI), surnommé « Guerre des Étoiles », de l'administration Reagan et qui avait pour but de saturer par réaction la recherche militaire des Soviétiques et donc de lui retirer toute capacité d'initiative.

Lou Gerstner a rejoint Thomas Watson au rang des figures mythiques d'IBM. Il est celui qui a sauvé IBM en maintenant son intégrité et a transformé la compagnie en profondeur. Il en racontera l'histoire dans son livre : *J'ai fait danser un éléphant (Who Says Elephants Can't Dance ? Inside IBM's Historic Turnaround)*.

Samuel J. Palmisano (2002 - 2011)

Sam Palmisano poursuit le développement d'IBM vers les services et son désengagement des lignes de produit technologiques.

- en 2002, après avoir jeté les bases de la technologie de stockage de demain *Millipede*, IBM invoque des problèmes de qualité ponctuels sur sa ligne de disques Deskstar pour céder cette activité à *Hitachi*. En fait, si *Millipede* aboutit, cela signifie très exactement la mort du concept même de disque dur, et la compagnie prend donc là un pari ambitieux.
- Signe fort : IBM s'engage de plus en plus fortement sur *Linux* et le fait savoir. Le fait de ne plus développer de système « propriétaire » non seulement diminue les coûts, mais entend marquer la nouvelle politique d'ouverture.
- en octobre 2002, IBM achète la branche *consulting* (concept anglo-saxon intégrant conseil en management classique et SSII) de PricewaterhouseCoopers et met ainsi un point d'orgue à sa stratégie de diversification dans les services : avec cette acquisition, IBM devient numéro un mondial en effectifs, devant Accenture.
- en 2004, IBM a ouvert les spécifications de l'architecture *PowerPC* et invité tous les fondateurs à y prendre part ; le but est de contester à *Intel* le *leadership* dans le domaine de la micro-informatique avec son architecture *x86*.
- le 8 décembre 2004, le fabricant de PC chinois *Lenovo* rachète pour 1,25 milliard de dollars la branche PC d'IBM et devient ainsi le troisième constructeur d'ordinateurs individuels au monde (après *Dell* et *Hewlett-Packard*). Après les disques durs, c'est un autre fleuron matériel de *Big Blue* qui est vendu. Officiellement, la branche n'était pas rentable, les pertes opérationnelles représentant quelques pour cent du chiffre d'affaires. Depuis 1991, la stratégie apparente d'IBM en matière de cession d'activité est de vendre quand l'avantage technologique est en passe de ne plus être suffisant pour justifier des prix supérieurs à ceux des concurrents.
- en avril 2005, IBM annonce des résultats en deçà des objectifs, ce qui provoque un coup de tonnerre sur les places financières et amène de nombreuses questions sur la pertinence de la stratégie « plus de services, moins de technologie » de Palmisano.
- Fin mai 2005, IBM perd le marché des ordinateurs *Apple* : à partir de 2006, ceux-ci seront équipés de processeurs *Intel*.
- Parallèlement, lors de l'E3, il devient officiel que c'est IBM qui équipera *Microsoft*, *Nintendo* et *Sony* en processeurs *Cell* (déclinaison du *POWER4* équipée de plusieurs processeurs multimédia intégrés) pour l'intégralité de leur nouvelle génération de consoles.
- En 2009, le projet *Blue Insight* migre 140 000 utilisateurs vers le nuage privé d'IBM^[52] : IBM estime avoir baissé ses coûts de configuration ou de suivi d'exploitation de 50 % et ses défauts logiciels de 30 %.

Les revenus du CIO d'IBM sont de 18,77 millions de dollars en 2006, 20,91 millions en 2007 et 20,97 millions en 2008^[53].

En octobre 2011, Sam Palmisano annonce son remplacement au 1^{er} janvier 2012 par *Gini Rometty*^[54].

Gini Rometty (2012 -)

Gini Rometty est la première femme à diriger IBM.

5.3.5 Logo

- Ce logo a été utilisé de 1924 à 1946. Il représente le globe, ceinturé du mot “international”^[1].
- Ce logo a été utilisé de 1947 à 1956. Le globe a été remplacé par les simples lettres “IBM” dans une police nommée “Beton Bold”^[2]
- Ce logo a été utilisé de 1956 à 1972. IBM a dit que les lettres avaient une apparence plus équilibrée et plus robuste^[3].
- Le logo en barres a été utilisé la première fois en 1967 et a remplacé le logo plein en 1972. Les barres horizontales suggèrent « la vitesse et le dynamisme ».

1. "IBM Archives : International Business Machines (1924-1946)." Retrieved January 16, 2007.

2. "IBM Archives : IBM in transition (1947-1956)." Retrieved January 16, 2007.

3. "IBM Archives : IBM continuity (1956-1972)." Retrieved January 16, 2007.

Le logo actuel (les 2 versions en 8 et 13 barres), ainsi que le précédent ont été créés par le graphiste Paul Rand^{[55],[56]}. Dans les années 1980, les caractéristiques du logo d'Apple semblaient se moquer gentiment de celui d'IBM, présenté comme son principal « concurrent » à l'époque.

Le bleu officiel du logo est CMYK : C75 M43 Y0 K0, RGB : R75 G107 B175, WEB : #006699^[57].

Un logo *informel* d'IBM conçu par le designer Paul Rand circule dans la Compagnie depuis les années 1990 sous forme de fond d'écran et de pin's : un œil, une abeille, et la lettre M ("eye", "bee", M).

5.3.6 Principaux actionnaires

En décembre 2011^[58] :

- State Street Corporation 5,50 %
- Berkshire Hathaway 4,87 %
- The Vanguard Group 4,22 %
- BlackRock 2,48 %
- State Farm Mutual Automobile Insurance Co. 1,41 %
- Northern Trust 1,40 %
- Bank of New York Mellon Corporation 1,36 %
- Wellington Management Company 1,24 %

5.3.7 Résultats financiers

En 2004, IBM a réalisé, d'après son rapport annuel, un chiffre d'affaires d'un peu plus de 96 milliards de dollars. L'entreprise disposait alors d'une force de travail de 330 000 hommes répartis dans 75 pays.

D'après le même rapport, la Russie, l'Inde, la Chine et le Brésil ont constitué ses principaux moteurs de croissance.

En 2005, l'entreprise cède sa division PCD (*Personal Computing Division*), chargée de l'informatique personnelle, à la société chinoise Lenovo, et conforte ainsi sa stratégie orientée vers l'intégration et les services.

En 2006, IBM dans son rapport annuel annonce 91,4 milliards de dollars de Chiffre d'affaires pour un bénéfice de 9,4 milliards de dollars. Les systèmes (*hardware*) représentent 23 % des bénéfices avant impôts, les services 37 % et les logiciels 40 %. IBM est présent dans 170 pays dont 25 % dans la zone Europe/Moyen-Orient/Afrique, 30 % dans la zone Asie/Pacifique et 45 % dans la zone Amériques. Les augmentations de bénéfices les plus marquantes concernent la Chine (+16 %), le Brésil (+19 %), la Russie (+21 %) et l'Inde (+38 %).

De 2002 à 2007, les investissements d'IBM ont débouché sur des économies cumulées de 4,1 milliards de dollars dans le cadre de la transformation de ses propres centres informatiques.

En 2009, le chiffre d'affaires était de 95,76 milliards de dollars et les bénéfices de 13,43 milliards de dollars^[59].

IBM prévoit de doubler sa puissance de calcul d'ici 2010 sans accroître sa consommation d'énergie ni son empreinte carbone^[60]. Cette multinationale américaine fait partie des entreprises cotées au Dow Jones Industrial Average. Elle est également, ce qui est moins connu, présente dans le NASDAQ.

Fin 2011, et pour la première fois depuis quinze ans, IBM dépasse Microsoft en capitalisation boursière^[61] avec 161 milliards d'euros, validant sa stratégie de réorientation vers les services entamée dès 2005.

Mais, depuis, IBM ne cesse de reculer. Fin 2014, l'entreprise a enregistré son 11^e trimestre consécutif dans le rouge. Selon *Forbes*, IBM s'apprêterait à licencier près de 112 000 de ses salariés, principalement aux États-Unis^{[62],[63]}, une rumeur démentie par IBM^[64] qui parle d'une diminution d'effectif beaucoup moins importante. Une autre source indique que le nombre de licenciements pourrait être de 11 à 12000 en 2015^[65].

5.3.8 Critiques

IBM et le régime nazi

Dès 1934 la filiale allemande d'IBM, Dehomag (pour « *DEutsche HOLLerith MASchinen Gesellschaft* ») fournit au régime nazi des machines mécanographiques de poinçonnage de cartes perforées qui servent au réarmement, à la gestion de la force de travail des prisonniers politiques et aux nombreux recensements de la population allemande dès 1933^[66], ce qui fit de Dehomag la filiale d'IBM la plus profitable à la fin des années 1930. En parallèle, ces mêmes machines servirent au décompte de la population juive dans les ghettos et les camps de concentration durant la Seconde Guerre mondiale.

Ce sont aussi ces recensements efficaces, contenant des données ethniques et religieuses, qui permirent aux nazis de se saisir rapidement et presque totalement des populations de Juifs et de Roms en Allemagne et, avec une efficacité plus variable, dans les autres pays sous domination allemande.

Edwin Black, dans son livre intitulé *IBM et l'holocauste* publié d'abord en anglais en 2001 puis en français la même année, démontre que le système informatique vendu par la corporation IBM (surtout la machine Hollerith) et les cartes perforées imprimées aux États-Unis expliquent en grande partie l'efficacité du recensement des Juifs et des Roms sous le Troisième Reich ce qui donna à ces machines IBM, ainsi qu'aux experts qui les faisaient fonctionner, un rôle essentiel dans la Shoah.

Parmi les différents éléments de la démonstration figure l'écart significatif entre la mortalité des Juifs de France et de Hollande, 25 % contre 73 % : pour la France on observe une différence significative entre le nombre de Juifs français effectivement déportés, environ 85 000, et le quota initial fixé à 100 000, sur un total situé entre 300 000 et 350 000, avec un effet inverse en Hollande. En effet, René Carmille, chef du service national de statistiques, et en fait un résistant, manipula les fiches des machines Hollerith de façon à les rendre inutilisables (ce qui le conduisit lui-même à Dachau où il mourut en 1944), alors qu'en Hollande le système, servi par des fonctionnaires zélés tel Jacobus Lambertus Lentz, fonctionna parfaitement et permit même aux nazis de revoir à la hausse le quota initial de Juifs hollandais déportés.

Par ailleurs, ce livre nous apprend aussi que pendant un temps le tatouage inscrit sur le bras des détenus à Auschwitz a correspondu à leur numéro d'identification dans le système informatique^[67].

En tout état de cause, le dirigeant de IBM, Thomas J. Watson, se révéla indifférent à l'aspect éthique des activités d'IBM sous le régime nazi, bien qu'il fût au courant des politiques ouvertement racistes de ce dernier.

À la suite de la parution très médiatisée du livre, IBM s'engagea à donner des explications, qui prirent la forme de démentis^[68] s'appuyant sur la critique d'historiens accusant Edwin Black d'avoir recours à des méthodes d'investigation approximatives. Les travaux d'Edwin Black sur les liens, certes non causaux, de IBM avec la Shoah ainsi que le Porajmos sont pourtant désormais reconnus ; bien qu'encore discutés sur certains points.

Collaboration avec la NSA

Le 16 décembre 2013, le fonds de pension de la Louisiane, qui détient des actions dans l'entreprise, attaque IBM en justice au motif d'avoir collaboré avec la NSA dans le cadre du programme de surveillance PRISM^[69]. Le Louisiana Sheriffs' Pension & Relief Fund estime le manque à gagner d'IBM en Chine à 12 milliards depuis qu'Edward Snowden a rendu l'affaire publique. Une baisse de 40% des ventes de matériel sur le marché chinois a été constatée lors du trimestre précédent, une baisse soudaine qui semble liée à la collaboration aux écoutes de la NSA. Il n'est toutefois pas encore prouvé que cette affaire soit la seule cause d'une telle chute des ventes.

5.3.9 Notes et références

[1] IBM voit son chiffre d'affaires fléchir de 6% en 2014, Challenges, 20.01.2015

[2] Résultats décevants pour IBM, qui veut céder son activité de serveurs, Le Monde, 22.01.2015

[3] Les ordinateurs et périphériques IBM étaient disponibles de 1960 à 1990 en 6 couleurs : bleu, vert sombre, rouge, jaune, gris et noir. Pour des raisons non élucidées – peut-être le fait que cette couleur était majoritaire sur les dépliant marketing – plus de deux tiers des clients choisissaient le bleu.

[4] « Classement des capitalisations boursières européennes et américaines depuis 1975 » sur Vernimmen.net

- [5] (es) Juan Carlos I^{er} et Jordi Sole Tura, « 1834/1991 de 27 de diciembre por el que se concede la Medalla al Mérito en las Bellas Artes, en su categoría de Oro, a las personas que se citan », *Boletín de Estado*, Madrid, n° 311, 28 décembre 1991, p. 41789 (lire en ligne).
- [6] « Life science : Fade or flourish ? », Guy Lefever, Michele Pesanello, Heather Fraser & Lee Taurman, IBM Institute for business value, 2011, p. 2
- [7] <http://www-03.ibm.com/press/us/en/pressrelease/23929.wss>
- [8] IBM achète I2 et Algorithmics
- [9] IBM étend son initiative Smarter Cities avec l'acquisition de Cúram Software
- [10] Lenovo to buy IBM's low-end server unit for \$2.3 billion, Paul Carsten et Soham Chatterjee, Reuters, 23 janvier 2014
- [11] IBM to pay Globalfoundries \$1.5 billion to take chip unit off its hands : WSJ, Reuters, 20 octobre 2014
- [12] « IBM passe un gros contrat avec ABN Amro », sur *Le Figaro*, 1^{er} décembre 2014
- [13] IBM signs \$1.25 billion WPP cloud deal and says more coming, Eric Auchard, Reuters, 2 décembre 2014
- [14] IBM Software Smart Book, 2011, page 4.
- [15] Des solutions logicielles adaptées à vos besoins
- [16] IBM Connections 4.5 : analyser le collaboratif par Maryse Gros le 30/01/2013.
- [17] Classement Fortune 500
- [18] Corporate responsibility report, 2008, page 4 et 16.
- [19] (en) Creating a new kind of workplace at IBM
- [20] Erin Traudt et Richard Vancil, Livre blanc d'IDC : *Comment entrer dans l'ère du Social Business : l'expérience d'IBM*, janvier 2011, page 10.
- [21] Site d'IBM developerWorks : la ressource des développeurs informatiques
- [22] Erin Traudt et Richard Vancil, Livre blanc d'IDC : *Comment entrer dans l'ère du Social Business : l'expérience d'IBM*, janvier 2011, page 5.
- [23] (en) Règles IBM de conduite des affaires
- [24] (en) IBM Social Computing Guidelines
- [25] (en) Citation de Rich Hume, General Manager, Global Business Partners
- [26] IBM Software Smart Book, 2011, page 5.
- [27] IBM PartnerWorld, le réseau mondial des partenaires commerciaux
- [28] [PDF] (en) Partner Program Review : IBM PartnerWorld in 2011
- [29] Communiqué de presse sur <http://www-03.ibm.com/press/fr/fr/pressrelease/33028.wss>
- [30] Forum SaaS et Cloud IBM
- [31] autocommutateur développé à IBM La Gaude, France
- [32] Le 12 août 1981, IBM lançait une nouvelle machine, « l'IBM 5150 », une innovation qui allait donner un nouvel essor au marché des « ordinateurs personnels » vers le grand public.
- [33] [PDF] IBM en France (juillet 2010)
- [34] Le président d'IBM France dévoile la feuille de route du groupe d'ici 2015
- [35] (fr) « Partenariat entre IBM France et la Croix-Rouge française », sur croix.rouge.fr, 21 avril 2011
- [36] (fr) « Convention de partenariat avec IBM », sur service6civique.gouv.fr, 21 avril 2011
- [37] En 2010, IBM a réalisé en France 2,463 milliards de dollars de revenus sur le terrain des services informatiques
- [38] Maryse Gros, « IBM ouvre le 1er datacenter de SoftLayer à Paris », sur *Le Monde Informatique*, 23 octobre 2014

- [39] IBM étend son emprise sur l'analytique
- [40] IBM met 3 milliards de dollars sur la table pour explorer "l'après-silicium", *Journal du Net*, 10 juillet 2014
- [41] (en) In 2000, IBM announced it would invest \$1 billion in Linux. The commitment caught the attention of CEOs and CIOs all over the world. Today, Linux is the fastest-growing operating system in the world.
- [42] (fr) Open for business (2007), Jaap Bloem & Menno van Doorn (trad. Audrey Vuillermier), éd. VINT, 2007 (ISBN 978-90-75414-20-2), p. 38
- [43] IBM va investir un milliard de dollars dans Linux et l'open source pour Power Systems
- [44] (en) IBM Again Pledges \$1 Billion to a Linux Effort
- [45] IBM N°1 en dépôt de brevets aux Etats-Unis pour la 20ème année consécutive
- [46] Les étapes d'une refondation
- [47] Big Lou - L'homme qui réinvente IBM est apparu sans cravate et en chemise bleu lavande (une entorse à la sacro-sainte tenue de l'IBMer, chemise blanche-costume sombre)
- [48] http://www-03.ibm.com/ibm/history/exhibits/music/music_CH1.html
- [49] http://www-03.ibm.com/ibm/history/exhibits/music/music_clips.html
- [50] (en) three principles : Dedication to every client's success ; Innovation that matters, for our company and for the world ; Trust and personal responsibility in all relationships
- [51] Le *Dictionnaire Google* donne comme définition : 1. unité centrale 2. processeur central. Il y a un article Mainframe en italien. Le site *Grand dictionnaire* donne comme définition : « Ordinateur puissant offrant des possibilités et des ressources étendues, et auquel d'autres ordinateurs ou de nombreux terminaux peuvent être reliés. » , ce qui correspond à Ordinateur central.
- [52] Le système est bâti sur les serveurs IBM de la gamme System Z et utilise la solution de business décisionnel Cognos rachetée pour 5 milliards de dollars en 2007.
- [53] Olivier Rafal, « Le PDG d'IBM a maintenu sa rémunération en 2008 », *Le Monde informatique*, 10 mars 2009 , consulté le 11 mars 2009
- [54] "Ginni" Rometty, 54 ans, est entrée en 1981 chez IBM et a la charge des ventes, du marketing et de la stratégie.
- [55] "IBM Archives : IBM international recognition (1972-)." Retrieved January 16, 2007.
- [56] IBM Logo History.
- [57] ftp://ftp.software.ibm.com/ftp/lotusweb/certification/C4_Guidelines_V2.pdf
- [58] <http://finance.yahoo.com/q/mh?s=IBM+Major+Holders>
- [59] http://www.forbes.com/lists/2010/18/global-2000-10_The-Global-2000_Prof.html
- [60] (en) Projet Big Green, page 13.
- [61] IBM dépasse Microsoft en Bourse Le Monde - 03/10/2011
- [62] Près de 112 000 employés d'IBM licenciés cette semaine ?, MacPlus, 26 janvier 2015
- [63] Next Week's Bloodbath At IBM Won't Fix The Real Problem, Forbes, 22 janvier 2015
- [64] IBM Dismisses Report of Massive Layoffs, WSJ
- [65] Sources Say IBM Planning On Laying Off 12,000 Over Next Year, TechCrunch, 26 janvier 2015
- [66] (en) IBM et l'holocauste
- [67] (en) IBM and Auschwitz
- [68] (en)<http://www-03.ibm.com/press/us/en/pressrelease/828.wss>
- [69] IBM poursuivi pour avoir collaboré avec la NSA, silicon.fr, 16 décembre 2013

5.3.10 Voir aussi

Articles connexes

- [Blue Gene](#)
- [Deep Blue](#)
- [Herman Hollerith](#)
- [Liste de produits IBM](#)
- [Projet FS](#)
- [IBM Airline Control Program](#)
- [IBM Research](#)
- [Joshua Blue](#)

Liens externes

- (fr) [Portail d'IBM France](#)
- (fr) [Grandeur et déclin d'IBM - Par Jean-Jacques Duby qui fut longtemps directeur scientifique de la filiale française](#)
- (fr) [Article dans 01net sur IBM et les marchés financiers](#)
- (en) [Autobiographie de l'entreprise](#)
- (en) [Rapports annuels](#)
- (en) [IBM Redbooks - livres, articles et manuels sur des produits dont IBM](#)
-  [Portail des entreprises](#)
-  [Portail de l'informatique](#)
-  [Portail des États-Unis](#)

5.4 Seeburger

Seeburger est un fournisseur de solutions d'intégration conçues pour optimiser les transactions de l'entreprise étendue par l'automatisation des échanges entre tous les partenaires, quelles que soient leur taille et leurs ressources techniques.

Le groupe a été fondé en Allemagne en 1986 pour fournir des solutions d'intégration à l'industrie automobile. Aujourd'hui, il sert plus de 8.800 clients dans plus de 50 pays et plus de 15 secteurs d'activité avec sa B2B Gateway (passerelle B2B) et d'autres produits et services complémentaires. Seeburger est présent dans le monde entier, en particulier en Europe, en zone Asie-Pacifique et en Amérique du Nord.

Les solutions EDI de SEEBURGER peuvent aussi bien être installées directement chez les utilisateurs qu'être exploitées en tant que Service hébergé dans le centre de données SEEBURGER ou – sur demande – être utilisées comme Service Cloud. Des formes mixtes, comme les Services à Distance sont également possibles sans problème.

SEEBURGER propose une Suite de Business Integration (BIS) complète qui sert de plateforme centrale et d'échange de données pour tous les processus commerciaux avec des partenaires externes. La solution prend en charge toutes les passerelles telles que EDI, papier, fax, web vers les partenaires commerciaux, convertit les données dans les formats nécessaires, pilote les processus à accomplir en fonction de règles configurables, couple tous les principaux systèmes ERP et surveille les processus. En outre, pour les utilisateurs de SAP, des solutions permettant la mise en oeuvre

« d'utilisations optimisées » sont mises à disposition pour le déroulement des cycles de Commande/Encaissement (Order-to-Cash) et d'Achat/Règlement (Purchase-to-Pay). Pour les industries et pour les fournisseurs d'énergie, des solutions de surveillance spéciales sont proposées dans SAP. Les données non structurées et sensibles de toute sorte peuvent également être transmises de manière sûre et traçable via la Suite de Business Intégration BIS, sachant qu'ici des exigences de « Gouvernances » et de conformité IT sont couvertes.

Grâce à une conception homogène réalisée à 100 % par SEEBURGER, à son conseil, à la mise en œuvre de logiciels d'intégration éprouvés, et à son catalogue étendu de solutions disponibles – composé des données des partenaires et de leurs processus commerciaux (mapping des messages) – leurs utilisateurs obtiennent très vite une automatisation intégrale des processus commerciaux pour tous leurs partenaires, petits ou grands.

Ainsi, leurs utilisateurs réalisent d'importantes économies grâce à la disparition d'étapes de traitements manuels, à l'absence d'erreurs dans les données et processus, à une transparence totale, à une traçabilité, et à une accélération des processus.

Les utilisateurs de SEEBURGER profitent de 26 ans d'expertise dans le secteur B2B et de leur savoir-faire en termes de processus mis en œuvre pour des projets réalisés chez plus de 8800 entreprises comme Beiersdorf, Bosch, Coop, Danfoss, EnBW, E.ON IT, Heidelberger Druckmaschinen, Intersport, Lidl, Osram, Siemens, s.Oliver, Schiesser, SupplyOn, RWE, Volkswagen, etc... La société SEEBURGER AG a été créée en 1986 à Bretten et exploite 19 filiales dans le monde, en Europe, en Asie et Amérique du Nord, et agit en tant que partenaire SAP de longue date.

SEEBURGER est membre de plusieurs institutions et organisations industrielles telles que BVL, W3C, OASIS, ebXML, AIAG, ODETTE, GALIA, VDA/ITA, VDMA, EAN, EDIFICE, GS1, EDNA, EPCglobal, RosettaNet, ERCIM, etc...

5.4.1 Historique

- 1986 Lancement commercial de la solution EDI la plus vendue dans le secteur automobile.
- 1993 Extension de la solution EDI à d'autres secteurs d'activité.
- 1995 Première entreprise au monde à recevoir la certification SAP R/3 Rel. 3.x. (1998 pour SAP R/3 Rel. 4.x.).
- 1996 Prix européen de l'innovation logicielle pour le produit EDWIN à l'ESIP.
- 1997 Début de l'internationalisation afin d'offrir un support client mondial.
- 1999 Seeburger est le leader européen des échanges de données informatisés (EDI/EDIFACT).
- 2001 Positionnement réussi comme fournisseur de solutions d'intégration Business-to-Business.
- 2003 Renforcement du partenariat avec SAP.
- 2004 Extension de la gamme de services en externalisation et infogérance.
- 2005 Avec l'acquisition de la société Freeformation GmbH, Trèves (Document Automation), SEEBURGER complète son offre dans le domaine des passerelles et du traitement des processus.
- 2006 SEEBURGER améliore et étend son portefeuille dans les domaines SCM/RFID et PDM/PLM. SEEBURGER se positionne pour la première fois en leader dans le Rapport du « Gartner Magic Quadrant » pour les fournisseurs de passerelles B2B.
- 2007 SEEBURGER et BOSCH remportent l'elog@istics Award 2007 de l'AKJ Automotive pour le premier projet mondial de supports de petites charges RFID. La distinction du China Award 2007 « Best IT Supply Chain Solution China » pour la solution Just-in-time/Just-in-Sequence dans le groupe automobile Beijing Benz-Daimler Chrysler (BBDC) confirme un succès considérable sur le marché d'avenir que constitue l'Asie.
- 2008 Gartner confirme à nouveau l'importance du Serveur Business Integration dans le quadrant des « leaders » dans le domaine des fournisseurs de passerelles B2B. Le nombre d'adaptateurs SEEBURGER certifiés SAP s'étend à 28 : 14 adaptateurs EDI reçoivent la certification pour une utilisation avec SAP NetWeaver PI ; 14 autres sont à nouveau certifiés pour une utilisation avec SAP NetWeaver XI.

- 2010 SEEBURGER étend son offre stratégique de produits et présente, avec la Suite Business Integration (BIS), une plateforme d'intégration complète de nouvelle génération pour l'intégration des partenaires commerciaux, allant de l'échange de données pur via des solutions EDI/B2B et MFT aux processus intégrés et à la surveillance dans SAP. Des analystes renommés confirment le rôle de leader qu'a pris SEEBURGER sur son secteur : SEEBURGER est considéré comme le leader des solutions d'intégration complètes par l'entreprise d'analyse de marché indépendante Forrester dans son CIS-Report et est qualifié de visionnaire dans le Gartner Magic Quadrant en termes de projets d'intégration d'applications de systèmes. SEEBURGER poursuit son expansion internationale et ouvre une filiale en Belgique et des représentations à Moscou, en République Tchèque, et en Turquie.
- 2011 25 ans de SEEBURGER – Integrating Best-Run Businesses. Plus de 8 500 clients satisfaits, de la PME au consortium international, sont les témoins de notre réussite et de la croissance continue de SEEBURGER depuis 1986.

5.4.2 Produits

- Support et Consulting EDI
- Industries
- Services Hébergés / Cloud
- Produits et Solutions
- Solutions SAP

5.4.3 Liens externes

- www.seeburger.com
- www.seeburger.fr
- www.seeburger.de
-  Portail des entreprises
-  Portail de l'Allemagne

5.5 Tibco Software

 Pour les articles homonymes, voir [Tibco](#).

Tibco Software (NASDAQ : TIBX) est une entreprise informatique américaine de logiciels, fournisseur de logiciels d'infrastructure pour les organisations dans un contexte de **cloud computing**. Tibco se propose de gérer l'information, les prises de décisions, les process et les applications en temps réel pour plus de 4 000 clients^[1].

Tibco a été fondée en 1997 par Vivek Ranadivé^[2] ; le siège est situé Palo Alto en Californie, et la société dispose d'implantations en Amérique du Nord, Europe, Asie, Proche-Orient et Amérique du Sud^[3].

Ses principaux concurrents sont **IBM** et **Oracle**.

5.5.1 Principaux produits

- ActiveMatrix^[4]
- BusinessEvents

- Collaborative Information Manager^[5]
- Silver
- Spotfire ^[6]
- tibbr^[7]
- FTL^[8]
- ActiveSpaces Datagrid
- TopLink

5.5.2 Principales acquisitions

- En 2002, TIBCO rachète Talarian.
- En 2004, Staffware.
- En 2005, Objectstar.
- En 2007, Spotfire.
- En 2009, DataSynapse.
- En 2010, Netrics, Proginet (transfert de fichiers), OpenSpirit, Kabira, Loyalty Lab Inc.
- En 2011, Nimbus
- En 2012, LogLogic, gestion de logs
- En 2013, Maporama
- En 2014, Jaspersoft éditeur de l'outil JasperReports^[9]

5.5.3 Notes et références

[1] <http://www.fr.tibco.com/company/default.jsp>

[2] (en) « Interview de Vivek Ranadivé, fondateur » (consulté le 31/08/2012)

[3] Implantations dans le monde

[4] <http://www.itbusinessedge.com/news/tibco-takes-bpm-higher-level>

[5] http://www.information-management.com/issues/2007_53/10002159-1.html

[6] <http://www.informationweek.com/hardware/grid-cluster/tibco-brings-analytics-to-spotfire-31/223400043?pgno=1>

[7] <http://www.zdnet.com/blog/howlett/tibco-launches-tibbr-enough-to-make-enterprise-2-0-viable/2800>

[8] <http://www.informationweek.com/software/information-management/tibco-ftl-speeds-capital-market-messagin/229400462>

[9] TIBCO Software rachète Jaspersoft pour 185 millions de dollars

5.5.4 Liens externes

- Site officiel
- (en) <http://www.dailyfinance.com/2012/08/27/is-tibco-software-hiding-weakness/>
- (en) Cet article est partiellement ou en totalité issu de l'article de Wikipédia en anglais intitulé « Tibco Software » (voir la liste des auteurs).
-  Portail des entreprises
-  Portail de San Francisco et sa région
-  Portail de l'informatique

Chapitre 6

Compléments

6.1 Théorie de l'information

 Pour l'article homonyme, voir La Théorie de l'information.

La **théorie de l'information**, sans précision, est le nom usuel désignant la **théorie de l'information de Shannon**, qui est une théorie **probabiliste** permettant de quantifier le contenu moyen en **information** d'un ensemble de **messages**, dont le **codage informatique** satisfait une **distribution statistique** précise. Ce domaine trouve son origine scientifique avec **Claude Shannon** qui en est le père fondateur avec son article *A Mathematical Theory of Communications* publié en 1948.

Parmi les branches importantes de la théorie de l'information de Shannon, on peut citer :

- le codage de l'information,
- la mesure quantitative de **redondance** d'un texte,
- la compression de données,
- la cryptographie.

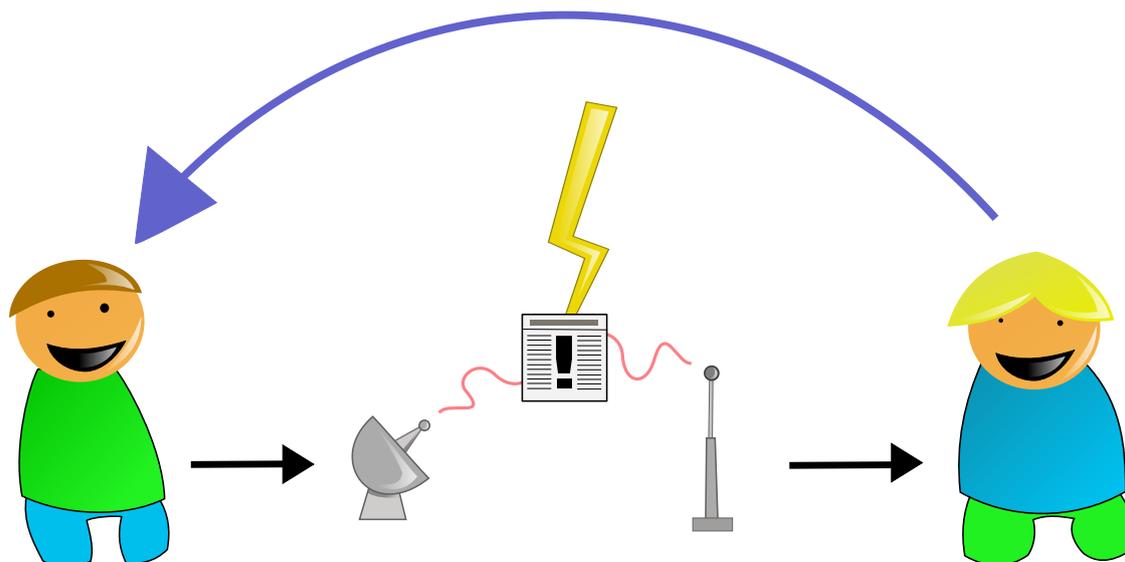
Dans un sens plus général, une théorie de l'information est une théorie visant à quantifier et qualifier la notion de contenu en **information** présent dans un ensemble de données. À ce titre, il existe une autre théorie de l'information : la théorie algorithmique de l'information, créée par Kolmogorov, Solomonov et Chaitin au début des années 1960.

6.1.1 L'information selon Shannon, un concept de la physique mathématique

L'information est un concept physique nouveau qui a surgi dans un champ technologique. Le concept théorique d'information a été introduit à partir de recherches théoriques sur les systèmes de **télécommunication**. L'origine de ces recherches remonte aux études entreprises dès la fin du XIX^e siècle, en physique et en mathématique par **Boltzmann** et **Markov** sur la notion de **probabilité** d'un événement et les possibilités de mesure de cette probabilité. Plus récemment, avant la **Seconde Guerre mondiale**, les contributions les plus importantes sont dues à la collaboration des mathématiciens et des ingénieurs des **télécommunications**, qui ont été amenés à envisager les propriétés théoriques de tout système de signaux utilisé par les êtres, vivants ou techniques, à des fins de communication.

À la suite des travaux de **Hartley** (1928), **Shannon** (1948) détermine l'information comme grandeur mesurable, sinon observable —car nul n'a jamais vu l'information— et celle-ci devient la poutre maîtresse de la théorie de la communication qu'il élabore avec **Warren Weaver**^[1].

Cette théorie est née de préoccupations techniques pratiques. La société Bell cherche à transmettre les messages de la façon à la fois la plus économique et la plus fiable. Aussi le cadre originel de la théorie est celui d'un système de communications où un émetteur transmet un message à un récepteur à travers un canal matériel/énergétique donné. Émetteur et récepteur ont par hypothèse un répertoire commun, un code qui contient les catégories de signaux utilisables. Ainsi le message codé est transmis, de l'émetteur au récepteur à travers le canal, sous forme de signes ou signaux portés par de la matière/énergie.



Modèle de la communication de Shannon et Weaver

Ainsi, le concept d'information a été l'objet d'une théorie que la postérité a choisi d'appeler "théorie de l'information" alors qu'il s'agissait, à proprement parler, d'une théorie mathématique de la communication de l'information ; or cette expression est exactement celle de Shannon et Weaver ! Cette source de confusion est régulièrement rappelée dans la littérature. On dit, en pareil cas, que l'expression abrégée a été retenue par l'usage ; l'emploi du sigle TMCI clarifierait pourtant bien la situation.

Cette théorie mathématique appliquée aux techniques de la télécommunication a été élaborée plus spécialement par Claude Shannon, ingénieur à la Compagnie des Téléphones Bell et reste jusqu'à nos jours la base du concept dit scientifique d'information. Cependant, cette théorie ne pourrait s'appuyer ni sur la forme matérielle/énergétique, ni sur le contenu cognitif des messages émis : leur contenu sémantique est laissé de côté, de même que leur contenant physique, pour ne s'intéresser qu'aux aspects mathématiques et communicationnels.

Dans sa conception originale, la théorie de l'information de Shannon s'est limitée à analyser les moyens à mettre en œuvre dans les techniques de télécommunication pour transmettre l'information le plus rapidement possible et avec le maximum de sécurité. Elle s'est donc efforcée de développer des méthodes susceptibles de minimiser la probabilité d'erreur dans la reconnaissance du message. Une notion fondamentale sera nécessaire pour développer ces méthodes : la mesure de l'information, au sens mathématique du terme.

Pour Shannon, l'information présente un caractère essentiellement aléatoire. Un événement aléatoire est par définition incertain. Cette incertitude est prise comme mesure de l'information. Une information sera donc uniquement définie par sa probabilité ($I = -\log p$). Donc l'information est la mesure de l'incertitude calculée à partir de la probabilité de l'événement. Shannon a donc confondu la notion d'information et de mesure d'incertitude. Il faut remarquer que dans cette définition l'information est bien synonyme de mesure d'incertitude. Dans cet ordre d'idée, plus une information est incertaine, plus elle est intéressante, et un événement certain ne contient aucune information. En théorie de l'information de Shannon, il s'agit donc de raisonner en probabilité et non en logique pure.

L'information de Shannon se mesure en unités binaires dites bits. Le bit peut être défini comme un événement qui dénoue l'incertitude d'un récepteur placé devant une alternative dont les deux issues sont pour lui équiprobables. Plus les éventualités que peut envisager ce récepteur sont nombreuses, plus le message comporte d'événements informatifs, plus s'accroît la quantité de bits transmis. Il est clair que nul récepteur ne mesure en bits l'information obtenue dans un message. C'est seulement le constructeur d'un canal de télécommunication qui a besoin de la théorie, et mesure l'information en bits pour rendre la transmission de message la plus économique et la plus fiable.

La notion d'information d'après Shannon est nécessairement associée à la notion de « redondance » et à celle de « bruit ». Par exemple, en linguistique l'information n'est ni dans le mot, ni dans la syllabe, ni dans la lettre. Il y a des lettres voire des syllabes qui sont inutiles à la transmission de l'information que contient le mot : il y a dans une phrase, des mots inutiles à la transmission de l'information. La théorie de Shannon appelle redondance tout ce qui dans le message apparaît comme en surplus. Aussi est-il économique de ne pas transmettre la redondance.

L'information chemine à travers un canal matériel/énergétique : fil téléphonique, onde radio, etc. Or, dans son che-

minement, l'information rencontre du bruit. Le bruit est constitué par les perturbations aléatoires de toutes sortes qui surgissent dans le canal de transmission et tendent à brouiller le message. Le problème de la dégradation de l'information par le bruit est donc un problème inhérent à sa communication. Ici, l'idée de redondance présente une face nouvelle ; alors qu'elle apparaît comme un surplus inutile sous l'angle économique, elle devient, sous l'angle de la fiabilité de la transmission un fortifiant contre le bruit, un préventif contre les risques d'ambiguïté et d'erreur à la réception.

6.1.2 Le statut physique de la théorie de l'information

Très vite de multiples applications de la théorie de l'information de Shannon sont apparues dans le domaine des sciences humaines^[2] : les modèles mathématiques élaborés ont permis de préciser certains concepts utilisés couramment dans les analyses linguistiques structurales, en même temps qu'ils faisaient apparaître les limites inhérentes à ce type d'analyse et provoquaient des recherches nouvelles (en traduction automatique et en psycho-linguistique). Tandis que se développait un champ scientifique nouveau : la *cybernétique*^[3].

Cependant, une caractéristique majeure de la théorie shannonienne est de donner à la notion d'information (telle que définie par cette théorie) un statut physique à part entière. Effectivement, l'information acquiert les caractères fondamentaux de toute réalité physique organisée : abandonnée à elle-même, elle ne peut évoluer que dans le sens de sa désorganisation, c'est-à-dire l'accroissement d'entropie ; de fait, l'information subit, dans ses transformations (codage, transmission, décodage, etc.), l'effet irréversible et croissant de la dégradation. Par conséquent Shannon définit comme entropie d'information la mesure H ($H = -K \log p$). De façon étonnante, l'équation par laquelle Shannon définit l'entropie de l'information coïncide, mais de signe inverse, avec l'équation de Boltzmann-Gibbs définissant l'entropie S en thermodynamique ($S = K \log p$). Cet épisode important a été abondamment commenté^[4].

Certains, comme Couffignal^[5], ont soutenu que la coïncidence est sans signification : l'application de la fonction de Shannon à la thermodynamique et à l'information serait un hasard de rencontre de l'application d'une même formule mathématique, sans plus. Certes, il peut y avoir rencontre de deux équations de probabilité provenant d'univers différents.

À l'inverse, Brillouin avait prétendu établir une relation logique entre le H de Shannon et le S de Boltzmann, ce que retiennent la plupart des chercheurs qui appliquent la théorie aux disciplines non mathématiques, la biologie en particulier. Selon ce point de vue, il est possible d'inscrire l'information shannonienne dans la physique. En effet, il existe une dualité dans le concept d'information reliant l'information à la matière/énergie véhiculant cette information. L'information shannonienne s'enracine ainsi dans la physique d'une part, dans les mathématiques d'autre part, mais sans qu'on puisse la réduire aux maîtres-concepts de la physique classique : masse et énergie. Comme le dit Wiener : « l'information n'est ni la masse, ni l'énergie, l'information est l'information », ce qui laisse la porte ouverte à des conceptions diverses, à commencer par celle d'un troisième constituant de l'univers^[6], après la matière et l'énergie précisément !

6.1.3 Développement de la théorie mathématique de l'information

La théorie mathématique de l'Information résulte initialement des travaux de Ronald Aylmer Fisher. Celui-ci, statisticien, définit formellement l'information comme égale à la valeur moyenne du carré de la dérivée du logarithme de la loi de probabilité étudiée.

$$\mathcal{I}(\theta) = E \left\{ \left[\frac{\partial}{\partial \theta} \ln f(X; \theta) \right]^2 \middle| \theta \right\}$$

À partir de l'inégalité de Cramer, on déduit que la valeur d'une telle information est proportionnelle à la faible variabilité des conclusions résultantes. En termes simples, moins une observation est probable, plus son observation est porteuse d'information. Par exemple, lorsque le journaliste commence le journal télévisé par la phrase « Bonsoir », ce mot, qui présente une forte probabilité, n'apporte que peu d'information. En revanche, si la première phrase est, par exemple « La France a peur », sa faible probabilité fera que l'auditeur apprendra qu'il s'est passé quelque chose, et, partant, sera plus à l'écoute.

D'autres modèles mathématiques ont complété et étendu de façon formelle la définition de l'information.

Claude Shannon et Warren Weaver renforcent le paradigme. Ils sont ingénieurs en télécommunication et se préoccupent de mesurer l'information pour en déduire les fondamentaux de la Communication (et non une théorie de l'information). Dans *Théorie Mathématique de la Communication* en 1948, ils modélisent l'information pour étudier les lois correspondantes : bruit, entropie et chaos, par analogie générale aux lois d'énergétique et de thermodynamique.

Leurs travaux complétant ceux d'Alan Turing, de Norbert Wiener et de John von Neumann (pour ne citer que les principaux) constituent le socle initial de la théorie du signal et des « Sciences de l'Information ».

Pour une source X comportant n symboles, un symbole i ayant une probabilité p_i d'apparaître, l'entropie H de la source X est définie comme :

$$H(X) = - \sum_i^n p_i \log_2(p_i)$$

C'est au départ le logarithme naturel qui est utilisé. On le remplacera pour commodité par le logarithme à base 2, correspondant à une information qui est le bit. Les considérations d'entropie maximale (MAXENT) permettront à l'inférence bayésienne de définir de façon rationnelle ses distributions *a priori*.

L'informatique constituera une déclinaison technique automatisant les traitements (dont la transmission et le transport) d'information. L'appellation « Technologies de l'Information et de la Communication » recouvre les différents aspects (systèmes de traitements, réseaux, etc.) de l'informatique au sens large.

Les sciences de l'information dégagent du sens depuis des données en s'appuyant sur des questions de corrélation, d'entropie et d'apprentissage (voir Data mining). Les technologies de l'information, quant à elles, s'occupent de la façon de concevoir, implémenter et déployer des solutions pour répondre à des besoins identifiés.

Adrian Mc Donough dans *Information economics* définit l'information comme la rencontre d'une donnée (data) et d'un problème. La connaissance (*knowledge*) est une information potentielle. Le rendement informationnel d'un système de traitement de l'information est le quotient entre le nombre de bits du réservoir de données et celui de l'information extraite. Les data sont le *cost side* du système, l'information, le *value side*. Il en résulte que lorsqu'un informaticien calcule la productivité de son système par le rapport entre la quantité de données produites et le coût financier, il commet une erreur, car les deux termes de l'équation négligent la quantité d'information réellement produite. Cette remarque prend tout son sens à la lumière du grand principe de Russel Ackoff qui postule qu'au-delà d'une certaine masse de données, la quantité d'information baisse et qu'à la limite elle devient nulle. Ceci correspond à l'adage « trop d'information détruit l'information ». Ce constat est aggravé lorsque le récepteur du système est un processeur humain, et pis encore, le conscient d'un agent humain. En effet, l'information est tributaire de la sélection opérée par l'attention, et par l'intervention de données affectives, émotionnelles, et structurelles absentes de l'ordinateur. L'information se transforme alors en sens, puis en motivation. Une information qui ne produit aucun sens est nulle et non avenue pour le récepteur humain, même si elle est acceptable pour un robot. Une information chargée de sens mais non irriguée par une énergie psychologique (drive, cathexis, libido, ep, etc.) est morte. On constate donc que dans la chaîne qui mène de la donnée à l'action (données → information → connaissance → sens → motivation), seules les deux premières transformations sont prises en compte par la théorie de l'information classique et par la sémiologie. Kevin Bronstein remarque que l'automate ne définit l'information que par deux valeurs : le nombre de bits, la structure et l'organisation des sèmes, alors que le psychisme fait intervenir des facteurs dynamiques tels que passion, motivation, désir, répulsion, etc. qui donnent vie à l'information psychologique.

6.1.4 Exemples d'information

Une information désigne, parmi un ensemble d'événements, un ou plusieurs événements possibles.

En théorie, l'information diminue l'incertitude. En théorie de la décision, on considère même qu'il ne faut appeler « information » que ce qui est « susceptible d'avoir un effet sur nos décisions » (peu de choses dans un journal sont à ce compte des informations...)

En pratique, l'excès d'information, tel qu'il se présente dans les systèmes de messagerie électronique, peut aboutir à une saturation, et empêcher la prise de décision.

Premier exemple

Soit une source pouvant produire des tensions entières de 1 à 10 volts et un récepteur qui va mesurer cette tension. Avant l'envoi du courant électrique par la source, le récepteur n'a aucune idée de la tension qui sera délivrée par la source. En revanche, une fois le courant émis et reçu, l'incertitude sur le courant émis diminue. La théorie de l'information considère que le récepteur possède une incertitude de 10 états.

Second exemple

Une bibliothèque possède un grand nombre d'ouvrages, des revues, des livres et des dictionnaires. Nous cherchons un cours complet sur la théorie de l'information. Tout d'abord, il est logique que nous ne trouverons pas ce dossier dans des ouvrages d'arts ou de littérature ; nous venons donc d'obtenir une information qui diminuera notre temps de recherche. Nous avons précisé que nous voulions aussi un cours complet, nous ne le trouverons donc ni dans une revue, ni dans un dictionnaire. Nous avons obtenu une information supplémentaire (nous cherchons un livre), qui réduira encore le temps de notre recherche.

Information imparfaite

Soit un réalisateur dont j'aime deux films sur trois. Un critique que je connais bien éreinte son dernier film et je sais que je partage en moyenne les analyses de ce critique quatre fois sur cinq. Cette critique me dissuadera-t-elle d'aller voir le film ? C'est là la question centrale de l'inférence bayésienne, qui se quantifie aussi en bits.

6.1.5 Contenu d'information et contexte

Il faut moins de bits pour écrire « chien » que « mammifère ». Pourtant l'indication « Médor est un chien » contient bien plus d'information que l'indication « Médor est un mammifère » : le contenu d'information sémantique d'un message dépend du contexte. En fait, c'est le couple message + contexte qui constitue le véritable porteur d'information, et jamais le message seul (voir [paradoxe du compresseur](#)).

6.1.6 Mesure de la quantité d'information

Quantité d'information : cas élémentaire

Considérons N boîtes numérotées de 1 à N . Un individu A a caché au hasard un objet dans une de ces boîtes. Un individu B doit trouver le numéro de la boîte où est caché l'objet. Pour cela, il a le droit de poser des questions à l'individu A auxquelles celui-ci doit répondre sans mentir par OUI ou NON. Mais chaque question posée représente un coût à payer par l'individu B (par exemple un euro). Un individu C sait dans quelle boîte est caché l'objet. Il a la possibilité de vendre cette information à l'individu B. B n'acceptera ce marché que si le prix de C est inférieur ou égal au coût moyen que B devrait dépenser pour trouver la boîte en posant des questions à A. L'information détenue par C a donc un certain prix. Ce prix représente la quantité d'information représentée par la connaissance de la bonne boîte : c'est le nombre moyen de questions à poser pour identifier cette boîte. Nous la noterons I .

EXEMPLE :

Si $N = 1$, $I = 0$. Il n'y a qu'une seule boîte. Aucune question n'est nécessaire.

Si $N = 2$, $I = 1$. On demande si la bonne boîte est la boîte n° 1. La réponse OUI ou NON détermine alors sans ambiguïté quelle est la boîte cherchée.

Si $N = 4$, $I = 2$. On demande si la boîte porte le n° 1 ou 2. La réponse permet alors d'éliminer deux des boîtes et il suffit d'une dernière question pour trouver quelle est la bonne boîte parmi les deux restantes.

Si $N = 2^k$, $I = k$. On écrit les numéros des boîtes en base 2. Les numéros ont au plus k chiffres binaires, et pour chacun des rangs de ces chiffres, on demande si la boîte cherchée possède le chiffre 0 ou le chiffre 1. En k questions, on a déterminé tous les chiffres binaires de la bonne boîte. Cela revient également à poser k questions, chaque question ayant pour but de diviser successivement le nombre de boîtes considérées par 2 (méthode de dichotomie).

On est donc amené à poser $I = \log_2(N)$, mais cette configuration ne se produit que dans le cas de N événements équiprobables.

Quantité d'information relative à un événement

Supposons maintenant que les boîtes soient colorées, et qu'il y ait n boîtes rouges. Supposons également que C sache que la boîte où est caché l'objet est rouge. Quel est le prix de cette information ? Sans cette information, le prix à payer est $\log_2(N)$. Muni de cette information, le prix à payer n'est plus que $\log_2(n)$. Le prix de l'information « la boîte cherchée est rouge » est donc $\log_2(N) - \log_2(n) = \log_2(N/n)$.

On définit ainsi la quantité d'information comme une fonction croissante de $\frac{N}{n}$ avec :

- N le nombre d'évènements possibles
- n le nombre d'éléments du sous-ensemble délimité par l'information

Afin de mesurer cette quantité d'information, on pose : $I = \log_2 \left(\frac{N}{n} \right)$

I est exprimé en bit (ou « logon », unité introduite par Shannon^[citation nécessaire], de laquelle, dans les faits, bit est devenu un synonyme), ou bien en « nat » si on utilise le logarithme naturel à la place du logarithme de base 2.

Cette définition se justifie, car l'on veut les propriétés suivantes :

1. l'information est comprise entre 0 et ∞ ;
2. un évènement avec peu de probabilité représente beaucoup d'information (exemple : « Il neige en janvier » contient beaucoup moins d'information que « Il neige en août » pour peu que l'on soit dans l'hémisphère nord) ;
3. l'information doit être additive.

Remarque : lorsqu'on dispose de plusieurs informations, la quantité d'information globale n'est pas la somme des quantités d'information. Ceci est dû à la présence du logarithme. Voir aussi : information mutuelle, information commune à deux messages, qui, dans l'idée, explique cette « sous-additivité » de l'information.

Entropie, formule de Shannon

Article détaillé : entropie de Shannon.

Supposons maintenant que les boîtes soient de diverses couleurs : n_1 boîtes de couleur C_1 , n_2 boîtes de couleur C_2 , ..., nk boîtes de couleurs C_k , avec $n_1 + n_2 + \dots + nk = N$. La personne C sait de quelle couleur est la boîte recherchée. Quel est le prix de cette information ?

L'information « la boîte est de couleur C_1 » vaut $\log N/n_1$, et cette éventualité a une probabilité n_1/N . L'information « la boîte est de couleur C_2 » vaut $\log N/n_2$, et cette éventualité a une probabilité n_2/N ...

Le prix moyen de l'information est donc $n_1/N \log N/n_1 + n_2/N \log N/n_2 + \dots + nk/N \log N/nk$. Plus généralement, si on considère k évènements disjoints de probabilités respectives p_1, p_2, \dots, p_k avec $p_1 + p_2 + \dots + p_k = 1$, alors la quantité d'information correspondant à cette distribution de probabilité est $p_1 \log 1/p_1 + \dots + p_k \log 1/p_k$. Cette quantité s'appelle entropie de la distribution de probabilité.

L'entropie permet donc de mesurer la quantité d'information moyenne d'un ensemble d'évènements (en particulier de messages) et de mesurer son incertitude. On la note H :

$$H(I) = - \sum_{i \in I} p_i \log_2 p_i$$

avec $p_i = \frac{n_i}{N}$ la probabilité associée à l'apparition de l'évènement i .

Codage de l'information

On considère une suite de symboles. Chaque symbole peut prendre deux valeurs s_1 et s_2 avec des probabilités respectivement $p_1 = 0,8$ et $p_2 = 0,2$. La quantité d'information contenue dans un symbole est :

$$p_1 \times \log_2 \frac{1}{p_1} + p_2 \times \log_2 \frac{1}{p_2} \approx 0,7219$$

Si chaque symbole est indépendant du suivant, alors un message de N symboles contient en moyenne une quantité d'information égale à $0,72N$. Si le symbole s_1 est codé 0 et le symbole s_2 est codé 1, alors le message a une longueur de N , ce qui est une perte par rapport à la quantité d'information qu'il porte. Les théorèmes de Shannon énoncent qu'il est impossible de trouver un code dont la longueur moyenne soit inférieure à $0,72N$, mais qu'il est possible de coder le message de façon que le message codé ait en moyenne une longueur aussi proche que l'on veut de $0,72N$ lorsque N augmente.

Par exemple, on regroupe les symboles trois par trois et on les code comme suit :

Le message $s_1s_1s_1s_1s_2s_2s_2s_1$ sera codé 010011110.

La longueur moyenne du code d'un message de N symboles est : $\frac{N}{3}(0,512 + 3 \times 0,128 \times 3 + 3 \times 0,032 \times 5 + 0,008 \times 5) = 0,728N$

Article détaillé : théorie des codes.

6.1.7 Limites de cette théorie

L'une des caractéristiques fondamentales de cette théorie est l'exclusion de la *sémantique*. La théorie de l'information est indifférente à la *signification* des messages. Le *sens* d'un message peut pourtant être considéré comme essentiel dans la caractérisation de l'information. Mais le point de vue de la théorie de l'information se limite à celui d'un messager dont la fonction est de transférer un objet^[7].

La théorie de l'information de Shannon est toujours relative à un ensemble de données, une famille de chaînes de caractères, caractérisée par une loi de distribution bien précise. Elle donne donc un contenu en information *en moyenne*, ce qui en fait une théorie *probabiliste*, particulièrement bien adaptée au contexte de la transmission de donnée, et dans ce cadre cette théorie a produit des résultats importants. En revanche, elle n'est pas en mesure de quantifier le contenu en information d'une chaîne prise isolément, un brin d'ADN par exemple, alors que la *théorie algorithmique de l'information* en est capable jusqu'à un certain point. Mais cette dernière théorie possède également ses propres limitations. C'est pourquoi il ne faut pas considérer que la notion d'information est entièrement cernée par la théorie de l'information de Shannon, ou la théorie algorithmique de l'information, mais que cette notion a besoin d'une variété de modélisations formelles pour s'exprimer^[8].

L'*information de Fisher* semble ainsi parfois avantageusement remplacer l'information de Shannon dans la mesure où elle est une quantification locale et non globale de l'information contenue dans une distribution^[9]. Cela dit, les deux notions sont liées^[10] et peuvent dans diverses applications mener aux mêmes résultats.

6.1.8 Notes et références

- [1] C.E. Shannon et W. Weaver : The mathematical Theory of communication, University of Illinois, Urbana III, 1949.
- [2] Voir par exemple, R. Escarpit : Théorie générale de l'information et de la communication, Hachette 1980.
- [3] La cybernétique naquit en happant le concept scientifique d'information naissant pour l'intégrer dans la théorie des machines. La cybernétique est une discipline entamée par Norbert Wiener en 1948. Elle est définie comme étant l'étude de la théorie de la commande et de la communication tant dans la machine que dans les êtres vivants. En effet, dans la conception de Wiener, la cybernétique n'est pas seulement une science des machines mais s'intéresse à tout système, vivant ou non, capable d'auto-contrôle et de communication. Ainsi Wiener fonda la cybernétique en liant la commande à la communication de l'information.
- [4] Jérôme Segal. *Le zéro et le un. Histoire de la notion scientifique d'information au 20ème siècle*. Éditions Syllepse, 2003, 890 p. ISBN 2 84797 046 0. James Gleick. *The information. A history, a theory, a flood*. Fourth Estate, 2011, 527 p. ISBN 978 0 00 722574 3.
- [5] L. Couffignal : La cybernétique, PUF, Que sais-je ?, Paris 1968.
- [6] Doyle, R. (Bob) et "The Information Philosopher" : www.informationphilosopher.com (I-Phi ou Pi). Stonier, T. *Information and the internal structure of the universe*. Springer Verlag, 1990. *Beyond information. The natural history of intelligence*. Springer Verlag, 1992.
- [7] Université de Tours, vue d'ensemble de la théorie de l'information, page 5
- [8] Jean-Paul Delahaye, *Information, complexité et hasard*, 1999, Hermes. Chapitre 1
- [9] B.R.Frieden, *Science from Fisher Information*, 2004, Cambridge University Press
- [10] C.R.Rao, *Differential Metrics in Probability Spaces in Differential Geometry in Statistical Inference*, Chapitre 5, Institute of Mathematical Statistics

6.1.9 Voir aussi

Bibliographie

- Léon Brillouin *Science et théorie de l'information*, J. Gabay, 2000 (ISBN 2876470365)
- Léon Brillouin *Science and information theory* (typographie plus lisible, mais version en anglais)
- (en) [PDF] C. E. Shannon « A Mathematical Theory of Communication », sur *L'Institut d'électronique et d'informatique Gaspard-Monge* (Reprinted with corrections from The Bell System Technical Journal, Vol. 27, p. 379–423, 623–656, July, October, 1948.)
- Thomas M. Cover, Joy A. Thomas, *Elements of Information Theory*, Wiley-Interscience, 2006 (ISBN 978-0-471-24195-9) [détail des éditions]
- (en) David MacKay, *Information Theory, Inference, and Learning Algorithms*, Cambridge University Press, 2003 (ISBN 0-521-64298-1) [détail des éditions]

Articles connexes

- Théorie algorithmique de l'information
- Autorégulation
- Codage de l'information
- Compression de données
- Projet:Sciences de l'information et des bibliothèques
- Sciences de l'information et de la communication
- Sciences de l'information et des bibliothèques
- Technologies de l'information et de la communication
- Théorème de Cox-Jaynes
- Traitement de l'information
- Fuite d'information
- Transport de l'information

Lien externe

- « Un cours de théorie de l'information par Louis Wehenkel », sur *Université de Liège*
-  Portail de l'informatique
-  Portail des mathématiques
-  Portail de l'informatique théorique

6.2 Liste de normes ISO par domaines

Fin 2009, l'Organisation internationale de normalisation (ISO) compte environ 18 000 normes actives^[1].

Cette liste en donne des exemples, classés par domaines. Certaines normes peuvent apparaître dans plusieurs domaines.

6.2.1 Normes fondamentales

Grandeurs et unités

- ISO 31 : Grandeurs et unités. Cette norme se compose de 14 parties :
 - Espace, tempsetc.
- ISO 1000 : Unités du Système International et recommandations pour l'emploi de leurs multiples et de certaines autres unités

Format des papiers d'écritures

- ISO 216 : Formats des papiers d'écriture (ISO/DIS 216)

Langues

- ISO 639 : Codes pour la représentation des noms de langue
 - ISO 639-1 : Code alpha-2
 - ISO 639-2 : Code alpha-3
 - ISO 639-3 : Code alpha-3 pour traitement exhaustif des langues (ISO/DIS 639-3)
 - ISO 639-4 : Guide d'implémentation et principes généraux des codes de langue (ISO/CD 639-4)
 - ISO 639-5 : Code alpha-3 pour les familles de langues et groupes de langues (ISO/CD 639-5)

Livre et bibliographie

- ISO 690 : Références bibliographiques
- ISO 2108 : Numéro international normalisé du livre (ISBN)
- ISO 3297 : Numéro ISSN

Caractères et codes de représentation

- ISO 233 : Translittération des caractères arabes en caractères latins
- ISO 233-2 : Translittération des caractères arabes en caractères latins — Partie 2 : Langue arabe — Translittération simplifiée
- ISO 646 : Codes de caractères graphiques
- ISO 2022 : Structure de code de caractères et techniques d'extension (ISO/CEI 2022)
- ISO 3166 : Codes pour la représentation des noms de pays et de leurs subdivisions
- ISO 4217 : Codes pour la représentation des monnaies et types de fonds
- ISO 5218 : Codes de représentation des sexes humains
- ISO 6709 : Représentation normalisée des latitude, longitude et altitude pour la localisation des points géographiques
- ISO 8601 : Représentation des date (métadonnée) et heure
- ISO 8859 : Codage des alphabets qui incluent ASCII
- ISO 9984 : Conversion des caractères géorgiens en caractères latins
- ISO 10646 : Jeu universel de caractères
- ISO 15924 : Codes des écritures et alphabets

6.2.2 Normes de spécifications

Elles indiquent les caractéristiques, les seuils de performance d'un produit ou d'un service.

- EN 2076-2 : Série aérospatiale - Lingots et pièces moulées en alliages d'aluminium et de magnésium - Spécification technique - Partie 2 - Lingots pour refusion.

6.2.3 Normes d'analyse et d'essais

Elles indiquent les méthodes et moyens pour la réalisation d'un essai sur un produit.

- ISO 6506-1 : Matériaux métalliques - Essai de dureté Brinell - Partie 1 : Méthode d'essai).

6.2.4 Qualité

Article détaillé : Série des normes ISO 9000.

- ISO 9000 : Systèmes de management de la qualité - principes essentiels et vocabulaire
- ISO 9001 : Systèmes de management de la qualité - Exigences
- ISO 9004 : Systèmes de management de la qualité - Lignes directrices pour l'amélioration des performances

Nota : les normes ISO 9002 et ISO 9003 ont été annulées lors de la publication de la version 2000 de la norme ISO 9001.

Qualité appliquée au marché de l'automobile

- ISO/TS 16949 : Référentiel commun basé sur la norme ISO 9001 avec des exigences propres au marché de l'automobile.

Qualité appliquée aux laboratoires d'analyse

- ISO 15189 : Qualité et compétence des laboratoires de biologie médicale.
- ISO 17025 : Qualité et compétence des laboratoires d'étalonnages et d'essais

Qualité appliquée aux projets

- ISO 10006 : Management de la qualité appliqué aux projets.

6.2.5 Banques et services financiers

- ISO 3166 : Codes des pays
- ISO 4217 : Codes des monnaies du monde
- ISO 6166 : Codes des valeurs mobilières
- ISO 9362 : Codes d'identification des banques
- ISO 10383 : Codes d'identification des marchés boursiers règlementés et non règlementés
- ISO 13616 : Numéros de compte bancaire international (IBAN)

6.2.6 Technologies de l'information et de la communication

Cette section ne donne que les normes parmi les plus caractéristiques. Pour une liste plus complète, voir : JTC1

Données et records management

- ISO 9660 : Système de fichiers pour CD-ROM.
- ISO 11179 : Représentation des métadonnées des organisations dans un registre de métadonnées.
- ISO 15489 : Records management, sur la gestion des documents d'archive et les métadonnées.
- ISO 15836 : Information et documentation - L'ensemble des éléments de métadonnées Dublin Core
- ISO 22310 :2006 : Information et documentation -- Lignes directrices pour les rédacteurs de normes pour les exigences de "records management" dans les normes. En anglais seulement.
- ISO 23081-1 :2006 : Information et documentation - Processus de gestion des enregistrements - Métadonnées pour les enregistrements - Partie 1 : Principes (2006). En anglais seulement.

Terminologies

- ISO 12620 : catégories de données terminologiques utilisées pour représenter les unités d'information d'un langage de représentation de données terminologiques.
- ISO 16642 : Computer applications in terminology -- Terminological markup framework (en anglais seulement)

Échanges d'informations et de données (voir aussi environnement cycle de vie)

- ISO 646 : Jeu ISO de caractères codés à 7 éléments pour l'échange d'information
- ISO 2709 : Format pour l'échange d'information (ISO/DIS 2709)
- ISO 10303 : STEP Norme pour l'échange de modèles de données de produits
- ISO 14048 : Formats de documentation de données
- ISO/CEI 19503 : XML Metadata Interchange, standard d'échange de données basé sur XML (en anglais seulement)
- ISO 21127 : Information et documentation -- Une ontologie de référence pour l'échange d'informations du patrimoine culturel (en anglais et en français)

Langages de programmation

- ISO 1539 : Langage de programmation Fortran
- ISO 8652 : Langage de programmation Ada
- ISO 9899 : Langage de programmation C
- ISO 10279 : Langage de programmation BASIC

Réseaux

- ISO 7498 : Modèle réseau OSI en 7 couches
- ISO 11898 : Gestionnaire de réseau de communication (CAN)

Exigences Logicielles, évaluation des procédés

- ISO 9126 : Gestion des exigences logicielles
- ISO 15504 : Technologies de l'information -- Évaluation des procédés --

Sécurité des Systèmes d'Information

- ISO 13335 : Concepts et modèles pour la gestion de la sécurité des TIC (1996),
- ISO 15408 : Critères communs pour l'évaluation de la sécurité des Technologies de l'Information,
- ISO/CEI 17799 : Code de bonnes pratiques pour la gestion de la sécurité d'information (ancienne référence de la norme ISO/CEI 27002).
- ISO/CEI 27001 : Description des exigences pour la mise en place d'un Système de Management de la Sécurité de l'Information (SMSI).
- ISO/CEI 27002 : Nouveau nom pour la norme ISO/CEI 17799.
- ISO/CEI 27005 : Proposition d'une méthode d'appréciation des risques (cette phase est obligatoire dans le cadre d'une certification ISO/CEI 27001).
- ISO/CEI 27006 : Contient des informations sur le profil propre de l'auditeur 27001.

Voir aussi la section Sécurité de l'information.

Archivage électronique

Plusieurs normes indiquent la marche à suivre pour assurer l'archivage des documents sur des supports autres que le papier :

- NF Z 42-013, publiée en 1999 et révisée en 2009, est la plus connue en France. Elle précise des procédures techniques et organisationnelles permettant de garantir l'intégrité des documents lors de leur enregistrement, de leur stockage et de leur restitution. La première version de la norme NF Z 42-013 a été publiée à une époque où l'écrit électronique ne pouvait constituer qu'un commencement de preuve et non une preuve de même niveau que l'écrit sous forme papier (art. 1316-1 du Code civil). Sa révision de 2009 a élargi son périmètre.
- NF Z 43-400, homologuée le 20 août 2005, est relative à la micrographie informatique.
- ISO 14721 :2002 Systèmes de transfert des informations et données spatiales -- Système ouvert d'archivage d'information - Modèle de référence (modèle OAIS)
- ISO 19005-1, publiée en 2009, relative au format PDF destiné à l'archivage d'informations pour le long terme (format PDF/A).

Information géographique

- ISO 19101 :2002 : Information géographique -- Modèle de référence
- ISO/WD 19101 : Information géographique -- Modèle de référence
- ISO/TS 19101-2 :2008 : Information géographique -- Modèle de référence -- Partie 2 : Imagerie
- ISO/TS 19103 :2005 : Information géographique -- Langage de schéma conceptuel
- ISO/TS 19104 :2008 : Information géographique -- Terminologie
- ISO 19105 :2000 : Information géographique -- Conformité et essais
- ISO 19106 :2004 : Information géographique -- Profils

- ISO 19107 :2003 : Information géographique -- Schéma spatial
- ISO 19108 :2002 : Information géographique -- Schéma temporel
- ISO 19109 :2005 : Information géographique -- Règles de schéma d'application
- ISO 19110 :2005 : Information géographique -- Méthodologie de catalogage des entités
- ISO 19111 :2007 : Information géographique -- Système de références spatiales par coordonnées
- ISO 19111-2 :2009 : IG-- Système de références spatiales par coordonnées -- Partie 2 : Supplément pour valeurs paramétriques
- ISO 19112 :2003 : Information géographique -- Système de références spatiales par identificateurs géographiques
- ISO 19113 :2002 : Information géographique -- Principes qualité
- ISO 19114 :2003 : Information géographique -- Procédures d'évaluation de la qualité
- ISO 19115 :2003 : Information géographique -- Métadonnées
- ISO/NP 19115-1 : Information géographique -- Métadonnées -- Partie 1 : Titre manque
- ISO 19115-2 :2009 : Information géographique -- Métadonnées -- Partie 2 : Extensions pour les images et les matrices
- ISO 19116 :2004 : Information géographique -- Services de positionnement
- ISO/DIS 19117 : Information géographique -- Présentation
- ISO 19117 :2005 : Information géographique -- Présentation
- ISO/FDIS 19118 : Information géographique -- Codage
- ISO 19118 :2005 : Information géographique -- Codage
- ISO 19119 :2005 : Information géographique -- Services
- ISO 19119 :2005/Amd 1 :2008 : Extensions du modèle de métadonnées du service
- ISO/TR 19120 :2001 : Information géographique -- Normes fonctionnelles
- ISO/TR 19121 :2000 : Information géographique -- Imagerie et données quadrillées
- ISO/TR 19122 :2004 : Information géographique -- Qualification et accréditation du personnel
- ISO 19123 :2005 : Information géographique -- Schéma de la géométrie et des fonctions de couverture
- ISO 19125-1 :2004 : Information géographique -- Accès aux entités simples -- Partie 1 : Architecture commune
- ISO 19125-2 :2004 : Information géographique -- Accès aux entités simples -- Partie 2 : Option SQL
- ISO 19126 :2009 : Information géographique -- Dictionnaires de concepts de caractéristiques et registres
- ISO/TS 19127 :2005 : Information géographique -- Codes et paramètres géodésiques
- ISO 19128 :2005 : Information géographique -- Interface de carte du serveur Web
- ISO/TS 19129 :2009 : Information Géographique -- Structure de données pour les images, les matrices et les mosaïques
- ISO/TS 19130 :2010 : Information géographique -- Modèles de capteurs d'images de géopositionnement
- ISO/NP 19130-2 ; Titre manque
- ISO 19131 :2007 : Information géographique -- Spécifications de contenu informationnel
- ISO 19132 :2007 : Information géographique -- Services basés sur la localisation -- Modèle de référence

- ISO 19133 :2005 : Information géographique -- Services basés sur la localisation -- Suivi et navigation
- ISO 19134 :2007 : Information géographique -- Services basés sur la localisation -- Routage et navigation multimodaux
- ISO 19135 :2005 : Information géographique -- Procédures pour l'enregistrement d'éléments
- ISO/WD 19135-2 : Information géographique - Procédures pour l'enregistrement d'éléments -- Partie2 : Implémentation de schémas XML
- ISO 19136 :2007 : Information géographique -- Langage de balisage en géographie (GML)
- ISO 19137 :2007 : Information géographique -- Profil minimal du schéma spatial
- ISO/TS 19138 :2006 : Information géographique -- Mesures de la qualité des données
- ISO/TS 19139 :2007 : Information géographique -- Métadonnées -- Implémentation de schémas XML
- ISO 19141 :2008 : Information géographique -- Schéma des entités mobiles
- ISO 19142 :2010 : Information géographique -- Service d'accès aux entités géographiques par le web
- ISO 19143 :2010 : Information géographique -- Codage de filtres
- ISO 19144-1 :2009 : IG-- Systèmes de classification -- Partie 1 : Structure de système de classification
- ISO/DIS 19144-2 : Information géographique -- Systèmes de classification -- Partie 2 : Métalangage de couverture végétale (LCML)
- ISO/DIS 19145 : Information géographique -- Registre des représentations de localisation de point géographique
- ISO 19146 :2010 : Information géographique -- Vocabulaires interdomaines
- ISO/DIS 19148 : Information géographique -- Référencement linéaire
- ISO/DIS 19149 : IG-- Langue sur l'expression des droits pour l'utilisation de l'information géographique -- GeoREL
- ISO/CD 19151 : Titre manque
- ISO/CD 19152 : Information géographique -- Land Administration Domain Model (LADM)
- ISO/CD 19153 : Titre manque
- ISO/CD 19155 : Titre manque
- ISO/DIS 19156 : Information géographique -- Observations et mesures
- ISO/CD 19157 : Information géographique -- Qualité de données
- ISO/DTS 19158 : Titre manque
- ISO/NP TS 19159 : Titre manquant

6.2.7 Recherche d'information

- ISO 23950 Information and documentation - Information retrieval (Z39.50) - Application service definition and protocol specification.

6.2.8 Responsabilité sociétale des entreprises

- ISO 26000 : responsabilité sociale des organisations
- ISO 9001 : Qualité
- ISO 14001 : Environnement
- OHSAS 18001 : Santé Travail (pas une norme ISO)
- SA 8000 : Fournisseurs (standard)
- SD 21000 : Responsabilité sociétale et développement durable, norme française, n'a été mis en œuvre que pour des petites entreprises.

Le processus de rédaction de l'ISO 26000 a commencé à Salvador de Bahia du 7 au 11 mars 2005. les réunions suivantes ont eu lieu à Bangkok du 30 septembre au 7 octobre 2005 et à Lisbonne 13 au 19 mai 2006. Site de l'ISO sur la RS

Environnement, cycle de vie, échange de données

Article détaillé : Série des normes ISO 14000.

- ISO 14001 : Systèmes de management environnemental - Exigences et lignes directrices pour son utilisation (NF EN ISO 14001)

Sur l'analyse du cycle de vie :

- ISO 14040 : Principes et cadre de la série ISO 14040
- ISO 14041 : Définition de l'objectif, du champ d'étude et analyse de l'inventaire
- ISO 14042 : Évaluation de l'impact du cycle de vie
- ISO 14043 : Interprétation du cycle de vie
- ISO 14048 : Formats d'échange de données informatisées
- ISO 14049 : Rapports techniques sur des exemples d'analyse des inventaires selon ISO 14041

6.2.9 Sécurité

Sécurité de l'information

- ISO 27000 : Série de normes dédiées à la sécurité de l'information
 - ISO/CEI 27001 : Système de Management de la Sécurité de l'Information (SMSI) — Exigences
 - ISO/CEI 27002 : Code de bonnes pratiques pour la gestion de la sécurité de l'information (anciennement ISO/CEI 17799)
 - ISO/CEI 27003 : Système de Management de la Sécurité de l'Information (SMSI) — Guide d'implémentation
 - ISO/CEI 27004 : Mesure de la gestion de la sécurité de l'information
 - ISO/CEI 27005 : Gestion du risque en sécurité de l'information
 - ISO/CEI 27006 : Exigences pour les organismes réalisant l'audit et la certification de Systèmes de Management de la Sécurité de l'Information (SMSI)
 - ISO/CEI 27007 : Guide pour l'audit de Systèmes de Management de la Sécurité de l'Information (SMSI).
 - ISO/CEI 27799 : Informatique de santé - Gestion de la sécurité de l'information relative à la santé en utilisant l'ISO/CEI 27002

Nota : le standard Information Technology Security Evaluation Criteria (ITSEC) a été établi dans ce domaine en 1991 au niveau de l'Union européenne, mais il n'a pas obtenu le statut de norme.

Sécurité des denrées alimentaires

Voir aussi le paragraphe cycle de vie.

- ISO 22000 Systèmes de management de la sécurité des denrées alimentaires -- Exigences pour tout organisme appartenant à la chaîne alimentaire
- ISO 22002 Traçabilité de la chaîne alimentaire -- Principes généraux et exigences fondamentales s'appliquant à la conception du système et à sa mise en œuvre (projet en cours)
- ISO 22004 Systèmes de management de la sécurité des denrées alimentaires -- Recommandations pour l'application de l'ISO 22000 :2005
- ISO 22005 Traçabilité de la chaîne alimentaire -- Principes généraux et exigences fondamentales s'appliquant à la conception du système et à sa mise en œuvre

6.2.10 Notes et références

[1] http://www.iso.org/iso/fr/iso_catalogue.htm

6.2.11 Voir aussi

Article connexe

- Norme

Liens externes

- Site de l'ISO
- Liste des normes ISO
-  Portail du droit

6.3 ISO 8601

La norme internationale **ISO 8601** spécifie la représentation numérique de la **date** et de l'**heure** — respectivement basées sur le **calendrier grégorien** et le **système horaire de 24 heures**. Cette notation, créée en 1988, est particulièrement destinée à éviter tout risque de confusion dans les communications internationales dû au grand nombre de notations nationales différentes. Elle a en outre de nombreux avantages pour une utilisation **informatique** par rapport aux autres notations.

Il y a six niveaux de **granularité** dans ce format, selon les applications. Pour cette raison, il est possible d'omettre certains éléments.

Exemples :

- **1977-04-22T01:00:00-05:00** correspond au 22 avril 1977, à 1h du matin heure normale de l'est de l'Amérique du Nord (soit 5 heures de décalage).
- **1977-04-22T06:00:00Z** correspond au même instant.

6.3.1 Historique

La première édition de la norme ISO 8601 a été publiée en 1988. Elle a remplacé et unifié plusieurs normes ISO traitant les différents aspects de notation des dates et des heures (normes ISO 2014, 2015, 2711, 3307 et 4031)^[1]. Cette norme internationale a été mise à jour en 2000 puis 2004^[2].

La version de 1989 de la norme ISO était reprise par le système de normalisation français sous la référence **NF EN 28601 :1992**. Cette norme française a été annulée le 07 octobre 2011 ^[réf. souhaitée].

La notation numérique du plus étendu au plus précis (aaa-mm-jj) a été introduite par la norme ISO 2014 :1976 ; la numérotation des semaines par l'ISO 2015 :1976. Toutes deux sont intégrées dans la norme ISO 8601.

Le comité technique TC 154 s'occupe de la maintenance la norme ISO 8601^[3].

6.3.2 Principe

Date et heure

Les éléments les moins précis commencent pour aller vers les plus détaillés. L'année vient en tête (4 chiffres), suivie du mois (2 chiffres, de '01' à '12') et enfin du quantième (de '01' à '31'). Chaque élément est séparé de celui qui le précède par un tiret '-'. La date se fonde sur le calendrier grégorien. Ainsi, s'il était nécessaire de représenter des années avant la mise en application de ce calendrier (avant 1582-10-15 en France), une méthode de comptage entre les parties s'échangeant les informations devrait être convenue par avance.

Puis est indiquée l'heure de la journée qui débute par la lettre **T** (pour *Time*, soit temps en anglais), suivie des éléments suivants, dans cet ordre : heures (de '00' à '24'), minutes (de '00' à '59') puis secondes ('00' à '60') en utilisant deux-points ':' pour séparateur.

Et éventuellement suivi encore d'une virgule ',' puis de décimales de seconde. On rencontre parfois un point (toléré) au lieu d'une virgule, en particulier chez les anglophones. C'est toutefois la virgule qui est préférée par la norme.

Remarques importantes :

- Des décimales peuvent être ajoutées en quantité nécessaire à la précision recherchée. Exemple : **T15:23:56,9854**
- Si la notation de l'heure utilise le numéro '24' alors les minutes et les secondes doivent être à '00'. C'est une manière pratique en informatique de distinguer début et fin de journée quand la date n'est pas précisée. **T24 :00** équivaut strictement à **T00 :00 du jour suivant**.
- L'écriture '60' pour les secondes représentera par contre bel et bien une 61^e seconde. En effet, il est parfois nécessaire d'ajouter ou d'enlever une seconde intercalaire avant certains mois. (lire l'article sur le temps universel UTC pour les détails). **T23:59:60** représenterait donc, le cas échéant, une seconde ajoutée entre **T23:59:59** et **T24:00:00**.

Pour l'indication des heures habituelles, ces possibilités sont inutiles.

Par exemple, les horloges synchronisées n'indiquent généralement pas cette seconde intercalaire et préfèrent se resynchroniser progressivement. Utilisée ou non, toujours est-il que la norme a bel et bien prévu l'indication de ces secondes intercalaires.

Fuseau horaire

Pour tous les formats acceptés dans la norme ISO 8601, il est possible de spécifier le fuseau horaire. Le signe de départ est positif (+) lorsque l'heure locale est en avance sur le temps universel (UTC) (fuseaux horaires à l'Est du méridien horaire de Greenwich : fuseaux horaires d'Europe centrale, Afrique centrale et de l'Est, Asie, Pacifique Ouest et central). Négatif (-) sinon (fuseaux horaires de quelques pays d'Europe occidentale ou d'Afrique occidentale, Amériques, Pacifique Est). Il y a trois manières d'indiquer ce fuseau horaire :

1. en ajoutant **+hhmi** ou **-hhmi** dans les formats de base, *hh* représentant le nombre d'heures et *mi* le nombre de minutes de décalage par rapport au temps universel (le signe moins est habituellement le caractère ASCII *tiret-*

moins, mais certaines représentations permettent le signe moins mathématique, dont la largeur est identique à celle du signe plus) ;

2. en ajoutant *+hh :mi* ou *-hh :mi* dans les formats étendus, le séparateur horaire : étant écrit littéralement ;
3. en inscrivant littéralement un **Z** final lorsqu'il s'agit de l'heure UTC. (**Z** pour méridien **Z**éro, ou (*en*) **Z**ero *meridian*. Aussi connu sous le nom **Zulu** –prononcer *zoulou*– dans l'alphabet radio international).

Pour éviter d'éventuelles confusions, il s'agit bien d'un décalage exprimé *depuis* l'heure UTC (et non pas *vers*). Ce qui est logique puisqu'il s'agit d'une indication de fuseau horaire. Ainsi T09:00:00+01:00 correspond à T08:00:00Z (9 heures dans le fuseau ayant une heure de plus que l'UTC donc 8 heures en UTC). 09 :00+01 :00 ne représente pas l'addition de 09 :00 et de 01 :00, ce qui aurait été encore plus logique, mais plutôt cette addition-ci : **09 :00 = UTC +01 :00**.

Formats habituels

Ci-dessous, les notations suivantes sont utilisées pour symboliser les formats présentés :

- *aaaa* représente l'année grégorienne sur 4 chiffres fixes (1583 à 9999, les années 0001 à 1582 nécessitent un accord mutuel, de même que l'extension aux années à plus de 4 chiffres, il n'y pas de représentation standard des années avant Jésus-Christ, mais certains le permettent avec un accord mutuel en utilisant le préfixe littéral **B** suivi de 4 chiffres à partir de 0001 pour la désignation classique des années sans 0000, ou le préfixe **U** suivi d'un signe et de 4 chiffres à partir de 0000 pour l'extension grégorienne proleptique du calendrier UTC ; pour les années au-delà de 9999 = U+9999 ou avant B9999 ou avant U-9999, un accord mutuel est aussi nécessaire pour utiliser 5 chiffres fixes ou plus car cela entraîne un conflit avec les formats standards de base des dates incomplètes).
- – est écrit littéralement (signe moins -, ASCII ou mathématique –) dans les formats étendus (séparateur de date standard ISO), absent des formats de base.

En cas de notation de fractions (sur les heures, les minutes ou les secondes), le séparateur décimal est une virgule ou un point, la virgule étant le signe préféré.

- *mm* représente le numéro de mois dans l'année sur deux chiffres (01 à 12).
- *qq* représente le quantième dans le mois sur deux chiffres (01 à 31).
- *qqq* représente le quantième dans l'année sur 3 chiffres (001 à 366).
- *AAAA* représente l'année ISO sur 4 chiffres.
- **W** est toujours écrit littéralement.
- *ww* représente la semaine ISO sur deux chiffres (01 à 53).
- *j* représente le jour ISO de la semaine sur un chiffre (1 à 7).
- **T** est toujours écrit littéralement (peut être omis lorsqu'il n'y a pas de risque de confusion^[4]).
- *hh* représente l'heure du jour sur deux chiffres (00 à 23).
- : est écrit littéralement dans les formats étendus (séparateur horaire standard ISO).
- *mi* représente la minute sur deux chiffres (00 à 59).
- *ss* représente la seconde sur deux chiffres (00 à 60, en tenant compte des secondes ajoutées à la dernière minute de certains jours, selon la norme UTC)
- , est écrit littéralement dans les formats étendus (séparateur décimal ISO préféré^[5] ; le point étant acceptable).
- *n* représente la fraction de seconde sur un ou plusieurs chiffres.
- *zzzzz* représente le fuseau horaire dans le format de base (**Z** ou *+hhmi* ou *-hhmi*).

- zzzzzz représente le fuseau horaire dans le format étendu (**Z** ou **+hh :mi** ou **-hh :mi**).

Les formats suivants peuvent alors être utilisés pour indiquer :

- l'année seule :

format de base : "aaaa" (ex : 1997)

pas de format étendu

- l'année et le mois :

format de base : "aaaamm" (ex : 199707)

format étendu : "aaaa-mm" (ex : 1997-07)

- la date complète calendaire (année, mois, quantième) :

format de base : "aaaammqq" (ex : 19970716)

format étendu : "aaaa-mm-qq" (ex : 1997-07-16)

- la date complète ordinale (année, jour de l'année) :

format de base : "aaaaqqq" (ex : 1997206)

format étendu : "aaaa-qqq" (ex : 1997-206)

- la date complète hebdomadaire (année ISO, semaine ISO, jour de la semaine ISO) selon le numéro de semaine :

format de base : "AAAAWwwj" (ex : 2004W453)

format étendu : "AAAA-Www-j" (ex : 2004-W45-3)

- la date complète calendaire avec heures et minutes :

format de base : "aaaammqqThhmizzzzz" (ex : 19970716T1920+0100)

format étendu : "aaaa-mm-qqThh :mizzzzzz" (ex : 1997-07-16T19 :20+01 :00)

- la date complète calendaire plus l'heure, les minutes et les secondes :

format de base : "aaaammqqThhmisszzzz" (ex : 19970716T192030+0100)

format étendu : "aaaa-mm-qqThh :mi :sszzzzz" (ex : 1997-07-16T19 :20 :30+01 :00)

- la date complète calendaire et l'heure avec des fractions de seconde :

format de base : "aaaammqqThhmissnzzzz" (ex : 19970716T1920304+0100)

format étendu : "aaaa-mm-qqThh :mi :ss,nzzzzz" (ex : 1997-07-16T19 :20 :30,4+01 :00)

- l'heure seulement :

format de base : "Thhzzzzz" (ex : T19+0100)

format étendu : "Thhzzzzz" (ex : T19+01 :00)

- l'heure et les minutes :

format de base : "Thhmizzzzz" (ex : T1920+0100)

format étendu : "Thh :mizzzzz" (ex : T19:20+01:00)

- l'heure, les minutes et les secondes :

format de base : "Thhmisszzzz" (ex : T192030+0100)

format étendu : "Thh :mi :sszzzzzz" (ex : T19:20:30+01:00)

- l'heure complète avec des fractions de seconde :

format de base : "Thhmissnzzzzz" (ex : T1920304+0100)

format étendu : "Thh :mi :ss,nzzzzzz" (ex : T19:20:30,4+01:00)

Jour de l'année

Il est possible de représenter les dates en donnant l'année et le numéro du jour dans l'année. Donc un nombre sur 3 chiffres de '001' à '365' (voire '366' quand nécessaire). Exemple : **1970-033** représente le 33^e jour de l'année 1970. Autrement dit le 2 février 1970. Cette notation est distinctive grâce aux 3 chiffres après l'année. (Le mois seul ne prendrait que deux chiffres. Le mois et un **quantième** nécessiteraient quatre chiffres en notation abrégée.)

Numéro de semaine

Article détaillé : Numérotation ISO des semaines.

Système de numérotation L'industrie a pris l'habitude de numéroter les semaines de l'année (semaine 1, 2,... 52 voire 53) pour s'organiser et planifier sa production. En pratique, tout le monde ne respecte pas nécessairement la même convention de numérotation (par exemple, le Canada et les États-Unis considèrent le dimanche plutôt que le lundi comme le premier jour de la semaine). Cette norme ISO a défini de façon stricte une règle de numérotation, reconnue internationalement et couramment utilisée :

1. la semaine commence le lundi ;
2. les jours de chaque semaine sont numérotés de 1 (*un*) pour le lundi, à 7 (*sept*) pour le dimanche ;
3. la semaine 1 est celle qui contient le premier jeudi de l'année.

Remarques et conséquences :

- la semaine 1 est la première à contenir la majorité de ses jours (au moins 4 jours) dans l'année. Elle contient systématiquement le 4 janvier. Elle contient systématiquement le premier jour ouvré de l'année en considérant que le 1^{er} janvier, les samedis et les dimanches sont chômés. C'est aussi celle dont le lundi est le plus proche du jour de l'an. Elle commence au plus tôt le 29 décembre ou au plus tard le 4 janvier ;
- la dernière semaine de l'année (52 ou 53) est celle qui contient le dernier jeudi de l'année. Elle est la dernière à avoir la majorité de ses jours (au moins 4) dans l'année et contient systématiquement le 28 décembre. C'est aussi celle dont le dimanche est le plus proche du 31 décembre. Elle termine au plus tôt le 28 décembre ou au plus tard le 3 janvier ;
- il n'y a pas de semaine zéro (la semaine 1 d'une année succède immédiatement à la dernière semaine de l'année précédente) ;
- les premiers jours de la semaine 1 peuvent éventuellement être situés fin décembre de l'année précédente. De même, la dernière semaine (52 ou 53) d'une année peut avoir ses derniers jours au tout début de l'année suivante. Une année comporte toujours soit 52, soit 53 semaines au total ;

- les années ont 52 semaines en général ($365 \text{ jours} = 52 \times 7 + 1$). Donc, si une année a son 1^{er} janvier un lundi, son 31 décembre tombe aussi un lundi (un mardi si bissextile). Seules les années dont le jour de l'an est un jeudi auront 53 semaines (un mercredi ou un jeudi pour les années bissextiles).

Notation normalisée La norme prévoit également d'écrire une date d'après sa semaine. Il faut commencer par la lettre **W** (pour *week*, semaine en anglais) suivi du numéro de semaine sur deux chiffres (et éventuellement du numéro du jour de 1 à 7). Par exemple **2005-W52-1** correspond au premier jour (lundi) de la semaine 52 de l'année 2005. Soit autrement dit le lundi 26 décembre 2005 (ou **2005-12-26** ; les deux écritures sont équivalentes).

De la même façon, **2005-W52-7** correspond au dimanche de la dernière semaine de 2005 : donc au 1^{er} janvier 2006 ! (**2006-W01-1** correspond au lundi 2 janvier 2006, soit **2006-01-02**)

2004W48 désigne la semaine n°48 de l'année 2004.

Algorithme informatique Le calcul automatisé du numéro de semaine en informatique reste simple mais néanmoins non trivial, à cause des semaines charnières entre chaque année. Cette réflexion mérite donc d'être partagée.

Dans de nombreux langages de programmation informatiques, les outils comptent les jours de la semaine en partant du dimanche (valeur *zéro*), puis 1 pour le lundi, jusqu'à 6 pour le samedi. Dans certains pays, la semaine commence le dimanche et, mathématiquement (modulo 7), 0 et 7 reviennent au même dans les tests. Pour l'affichage, cela devient un cas particulier à traiter. Il existe parfois une fonction donnant directement une numérotation normalisée.

Pour calculer le numéro ISO de la semaine contenant une date 'D' donnée, il y a différents algorithmes. Cela peut être par exemple :

- calculer le jeudi 'J' de la même semaine que la date 'D' cherchée. (Ce jeudi est nécessairement situé dans la bonne année de numérotation, cela évite de gérer des cas particuliers). Ce jeudi peut donc tomber l'année précédente ou suivante ; le numéro ISO de la semaine de 'D' est identique à celui de 'J'
- calculer 'n' le numéro du jour de 'J' dans l'année ; le numéro de semaine est égal à 1 plus le résultat de la division euclidienne de $(n-1)$ par 7 (le numéro ISO de la semaine de 'J', et donc de 'D', est le nombre de jeudis écoulés entre le début de l'année où se situe 'J' et 'J').

Certains algorithmes utilisent des exceptions peu justifiables, voire retournent des valeurs erronées pour certaines dates. L'avantage de l'algorithme précédent est d'être fiable, tout en profitant des fonctions généralement disponibles (le compte des années bissextiles ou du nombre de jours est déjà implémenté).

Si on sait J le jour correspondant à une date (de J=1 pour lundi, à J=7 pour dimanche), en adaptant les fonctions existantes si nécessaire, alors le jeudi de la même semaine est situé (J-4) jours plus tôt (ou plus tard si négatif). Le lundi associé est (J-1) jours plus tôt.

La notation abrégée

Il est possible de n'indiquer que les éléments utiles. On peut par exemple indiquer le mois de février 2003 par : **2003-02**, sans devoir indiquer un jour précis ou un horaire. De manière générale on peut retirer toute information inutile (par exemple la date si seule l'heure est manipulée). Ce qui donne par exemple : **T23 :45**.

Dans un contexte donné, les notations doivent toutes être abrégées de la même manière. Donc toutes sur un même nombre de caractères. Toutes avec les mêmes informations. Sinon, on perd l'avantage de cette normalisation pour la cohérence et la facilité du tri.

Autre exemple, si le fuseau horaire est indiqué, il n'est pas toujours nécessaire d'indiquer les minutes (souvent inutiles). Ainsi **T23:05+05:00** pourra s'abrégé **T23:05+05** voire **T2305+05** d'après ce qui suit.

Enfin, si une taille compacte est requise au détriment de la lisibilité, il est possible de retirer certains séparateurs (tiret ou deux-points). Il convient alors de les retirer tous (tous les tirets et/ou tous les deux-points) pour éviter les confusions. Une écriture prise hors de son contexte doit rester compréhensible et on doit en particulier identifier s'il s'agit d'une heure ou d'une date. (2005 est une année, T2005 ou 20 :05 sont des heures)

Pour exemple, ces notations sont équivalentes :

20040204 = 2004-02-04 = 2004-W06-3 = 2004W063

Un dernier raccourci consiste à remplacer les éléments tronqués par un tiret '-'. Il est rarement employé. Toutefois il peut devenir nécessaire pour éviter des confusions. En omettant l'année, on noterait le 2 décembre ainsi : "--1202". On pourrait le réduire à "-1202" sans confusion avec "1202" seul qui représenterait l'année 1202.

2003--12 ou **2003-12** représenteraient le 12 de n'importe quel mois de l'an 2003, sans confusion avec **2003-12** qui représenterait décembre 2003.

Les durées et intervalles

Il est également possible d'indiquer une durée (mesure d'un temps écoulé) ou un intervalle de temps (entre deux dates précises). Pour ce faire on utilisera la barre oblique '/' pour séparateur.

Intervalle entre deux dates La barre oblique sépare la date de début et la date de fin. Si seul le dernier élément change il est possible de ne pas ré-écrire le début. Quelques exemples :

- **2004-12-02T23 :00/2005-01-01T12 :30** : une période entre deux dates et horaires différents ;
- **2004-12-02T22 :00/23 :30** : période comprise entre 22h et 23h30 le 2 décembre 2004.
2004-12-02T22 :00/2004-12-02T23 :30 : notation équivalente non abrégée ;
- **2004-12-02T22 :00/15** : période comprise entre 22h et 22h15 le 2 décembre 2004.
2004-12-02T22 :00/2004-12-02T22 :15 : notation équivalente non abrégée ;
- **2004-12-02T22/23** : période comprise entre 22h et 23h le 2 décembre 2004.
2004-12-02T22/2004-12-02T23 : notation équivalente non abrégée ;
- **2004-12-02/05** : la période de temps entre les 2 et 5 décembre 2004 (comprendre que les deux dates sont données comme des bornes à zéro heure, donc la journée du 5 décembre est exclue).
2004-12-02/2004-12-05 : notation équivalente non abrégée ;
- **T22:00/15** : intervalle entre 22h et 22h15 (la date est supposée connue dans le contexte).
T22:00/T22:15 : notation équivalente non abrégée ;
- **T22/23** : intervalle entre 22h et 23h (la date est supposée connue dans le contexte).
T22/T23 : notation équivalente non abrégée.

Durée Une durée représente une quantité de temps dans l'absolu, sans commencer à un instant précis. Il faut commencer l'écriture par la lettre 'P' (pour *Period* en anglais). Exemple :

P18Y9M4DT11H9M8S qui représente une (longue) durée de 18 ans, 9 mois, 4 jours, 11 heures, 9 minutes et 8 secondes. On remarque sur cet exemple l'emploi d'initiales anglaises pour séparer les éléments (*Years*, *Months*, *Days*, *Hours*, *Minutes* et *Seconds*). On remarque toujours la présence du **T** séparant les jours des heures. Il est vrai qu'ici la présence des H, M et S rend le T redondant. On pourrait aussi compter en semaines en réutilisant la lettre 'W' : **P3W** représente par exemple une durée de 3 semaines.

Si le contexte nécessitait de distinguer des durées positives ou négatives, il suffit d'ajouter le signe '+' ou '-' en tête d'écriture.

Une borne et la durée On peut enfin combiner les deux notations et indiquer une période en fonction de sa date de début et sa durée. Exemple :

2004-12-02/P3D étant l'intervalle de temps d'une durée de 3 jours débutant le 2 décembre 2004.

Ou encore, à l'inverse, on peut indiquer d'abord la durée puis la date finale ('PT1H/2006-01-01' représente l'heure qui précède le passage au jour de l'an de 2006).

Récurrence d'une durée Si la durée est répétitive, on précède la notation par la lettre R. Exemple : **R/PT01** signifie « toutes les heures ».

La barre oblique '/' permet un ajout après le R pour indiquer (éventuellement) un nombre de répétitions. **R5/PT01 :30** signifie donc « cinq occurrences, espacées de une heure trente entre elles ».

La durée entre chaque occurrence peut être indiquée par toutes les méthodes citées précédemment. (En indiquant deux dates ou en utilisant le symbole P pour des intervalles ou durées ; avec ou sans indication d'une date de départ ou de fin.)

R12/1985-04-12T23 :20 :50/P1Y2M signifierait « 12 occurrences séparées entre elles de un an et deux mois, en commençant la première le 12 avril 1985 à 23h20 et cinquante secondes »

6.3.3 Avantages

Les avantages de cette normalisation sont les suivants :

- Facile à lire et à écrire par un logiciel (pas de nom à traduire) ;
- Facile à comparer et à classer (en gardant un format fixe dans un contexte donné) ;
 - Le classement par ordre lexicographique correspond à l'ordre chronologique
- Indépendant de la langue ;
- Pas (ou peu) de risque de confusion avec d'autres notations ;
- La notation est petite et de taille constante ;
- Compréhension intuitive des éléments de date et d'heure de la notation.

6.3.4 Notes et références

- [1] « ISO 8601 :1988, revisions » (consulté le 31 mai 2014)
- [2] (en) « Data elements and interchange formats -- Information interchange -- Representation of dates and times » (consulté le 25 juin 2014)
- [3] « ISO/TC 154 Processus, éléments d'informations et documents dans le commerce, l'industrie et l'administration » (consulté le 31 juin 2014)
- [4] norme ISO 8601 :2004(E), ISO — ISO 4.3.2 NOTE : By mutual agreement of the partners in information interchange, the character [T] may be omitted in applications where there is no risk of confusing a date and time of day representation with others defined in this International Standard.
- [5] norme ISO 8601 :2004(E), ISO — ISO 4.2.2.4 ... the decimal fraction shall be divided from the integer part by the decimal sign specified in ISO 31-0, i.e. the comma [,] or full stop [.] . Of these, the comma is the preferred sign.

6.3.5 Liens externes

- « Représentation de la date et de l'heure - ISO 8601 », sur *iso.org* (consulté le 23 octobre 2013)

-  Portail du temps
-  Portail de l'informatique

6.4 Code lieu-fonction

Le **code lieu-fonction** ou *Global Location Number* (GLN) d'une entreprise est un identifiant unique attribué par l'antenne locale de GS1 (il y a des antennes locales dans plus de cent pays). Il est à la forme numérique d'un EAN 13. Ses premiers chiffres sont le code pays. Les chiffres suivants sont structurés à l'initiative de l'organisme national de codification. Le treizième et dernier chiffre est, comme pour tout EAN13, la clé de contrôle. En France, GS1 France a décidé de structurer la partie centrale de la manière suivante : le code pays (3 pour la France) est suivi de 01 (fournisseur) ou 02 (distributeur), puis du CNUF de l'entreprise, suivi d'un code interne choisi par l'entreprise pour différencier ses différents lieux-fonctions.

-  Portail des entreprises

6.5 Codage des caractères

Pour l'action de cacher le sens de l'information, voir chiffrement.



Carte perforée à 80 colonnes, sur laquelle est codée le texte de programmation « CALL RCLASS (AAA,21,NNC,PX3,PX4) »

En télécommunications et en informatique, un **jeu de caractères codés** est un code qui associe un jeu de caractères abstraits d'un ou plusieurs systèmes d'écriture (comme des alphabets ou des syllabaires) utilisés pour transcrire des langues naturelles avec une représentation numérique pour chaque caractère de ce jeu, ce nombre pouvant lui-même avoir des représentations numériques différentes. Par exemple, le code Morse (qui associe l'alphabet latin à une série de pressions longues et de pressions courtes sur le manipulateur morse du télégraphe) et le code ASCII (qui code les lettres, les chiffres et d'autres symboles comme des entiers codés sur 7 bits) sont des jeux de caractères codés.

Il est indispensable, pour l'échange d'information sur l'Internet, par exemple, de préciser le codage utilisé. Ne pas le faire peut rendre un document difficilement lisible (remplacement des lettres accentuées par d'autres suites de caractères connu sous le nom de *mojibake*). Toutefois, la convergence vers un standard commun devrait enfin répondre à ce problème.

Dans certains contextes (en particulier dans les communications et dans l'utilisation de données informatiques), il est important de distinguer un *répertoire de caractères*, qui est un jeu complet de caractères abstraits qu'un système supporte, et un *jeu de caractères codés* ou *codage de caractères* qui spécifie comment représenter un caractère en utilisant un entier.

6.5.1 Principe

Le *codage des caractères* est une convention qui permet, à travers un codage connu de tous, de transmettre de l'information textuelle, là où aucun support ne permet l'écriture scripturale.

Cela consiste à représenter chaque caractère, par un motif visuel, un motif sonore, ou une séquence abstraite. Les techniques des ordinateurs reposent sur l'association d'un caractère à un nombre, et/ou à un ou plusieurs codets.

D'autres techniques permettent, en utilisant un jeu restreint de caractères, d'en coder un plus grand nombre, ou de coder des octets qui peuvent en coder un plus grand nombre, notamment les bien connus Quoted printable, Entité de caractère et Percent-encoding (en) (défini par le RFC 1738).

Différence entre jeu de caractères codés et forme de codage

Le terme *jeu de caractères codés* est parfois confondu avec la façon dont les caractères sont représentés par une certaine séquence de bits, ce qui implique une *forme de codage* où le code entier est converti en un ou plusieurs *codets* (c'est-à-dire des valeurs codées) entiers qui facilitent le stockage dans un système qui gère les données par groupe de bits de taille fixe ou variable.

Par exemple le codage morse utilise un codage ternaire, dont deux des trois codets (impulsion courte ou longue) sont utilisés en groupes pour coder chaque valeur codée, le troisième (une pause plus longue) étant utilisé pour séparer les caractères.

De même, les entiers plus grands que 65535 ne peuvent pas être représentés sur 16 bits, c'est pourquoi la forme de codage UTF-16 représente ces grands entiers comme des couples d'entiers inférieurs à 65536 mais qui ne sont pas associés isolément à des caractères (par exemple, U+10000 - en hexadécimal - devient la paire 0xD800 0xDC00). Ce plan de codage convertit alors les valeurs de ces codes en une suite de bits et ce en prenant garde à un certain nombre de contraintes comme la dépendance vis-à-vis de la plateforme sur l'ordre final des octets (par exemple, D800 DC00 devient 00 D8 00 DC sur une architecture Intel x86). L'*Unicode Technical Report #17* explique cette terminologie en profondeur et fournit davantage d'exemples.

Une *page de code* abrège ce procédé en associant directement aux caractères abstraits des séquences de bits spécifiques de taille fixe (généralement 7 ou 8 bits par caractère).

Internationalisation

Les limites du standard américain ASCII ont conduit, sur trois périodes différentes, à trois approches de l'internationalisation :

- L'utilisation de standards régionaux à caractères-monooctets, techniquement les plus faciles à mettre en place ;
- L'utilisation de standards extensibles, où un même octet peut représenter un caractère différent suivant le contexte (famille ISO 2022) ainsi que des extensions où un caractère est codé sur plusieurs octets ;
- L'utilisation du Standard Unicode (famille UTF), qui est celui qui comprend le plus grand nombre de caractères.

Les standards régionaux ont l'inconvénient de ne permettre la représentation que d'un ensemble réduit de caractères, comme les caractères d'Europe occidentale. Avec cette approche, il est nécessaire d'indiquer l'encodage à l'extérieur du flot.

Les standards ISO 2022 ont l'inconvénient d'être contextuels. Il se peut que des logiciels utilisant certains algorithmes de recherche manquent d'interopérabilité à cet égard.

Les formes de codage définies par Le Standard Unicode ont l'inconvénient de la présence éventuelle de la fonctionnalité d'indicateur d'encodage en début de flot, qui le cas échéant est introduit par le caractère *Byte Order Mark*. Certains logiciels anciens ne sont pas compatibles avec la présence de ces trois octets, et ne pourront pas l'être en raison de la complexité conceptuelle que représente le fait de traiter ces trois octets. En particulier, certaines opérations deviennent plus complexes comme la concaténation de chaînes.

6.5.2 Historique

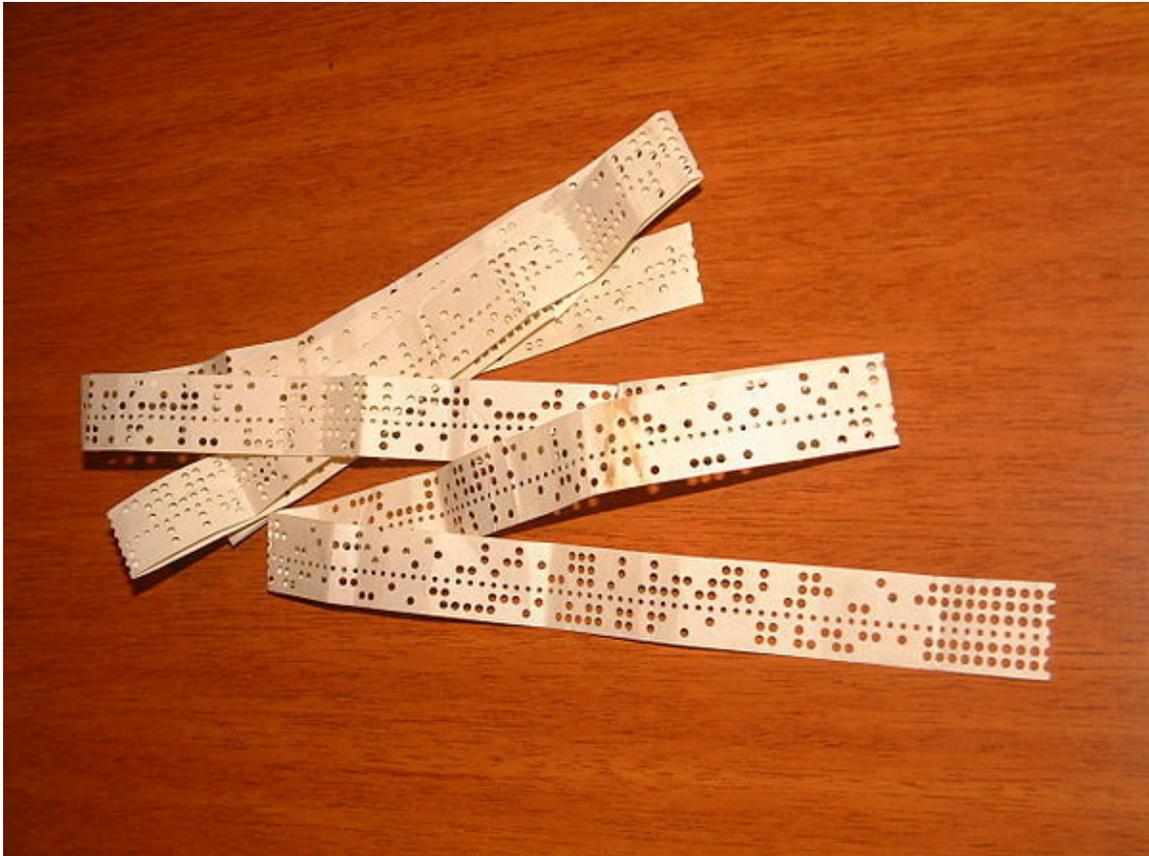
Premiers codages de caractères

Les premiers codages de caractères permettaient de s'affranchir de la distance. Ils utilisaient des techniques visuelles variées, mais ne codaient pas l'information sous forme binaire.

- *Code international des signaux maritimes*. Il s'agit de transmission visuelle, basée sur une séquence de drapeaux.
- *Tour Chappe* (1794). Il s'agit de transmission visuelles par positionnement des bras du télégraphe dans différentes postures.
- *Alphabet Morse* (1838). Il a été inventé en 1832 pour la télégraphie. Il est basé sur des séquences courtes et des séquences longues. Il permet de transmettre des textes non accentués, mais n'est pas limité en longueur, rendant des extensions possibles.

- Le télégraphe chinois convertissait les texte chinois avec des pages de codes en séquences de quatre chiffre décimaux, chiffres transmis en Morse.

Codage par séquence de bits



bande de papier avec des trous représentant les "Code Baudot"

Les premiers codages binaires de caractères furent introduits en France par l'ingénieur Mimault et par le Code Baudot pour le Téléx et aux États-Unis par des sociétés répondant aux appels d'offres de l'État fédéral. Pour le recensement puis des besoins particuliers.

Code Baudot (1874) : 32 codes (5 bits). Il permettait de transmettre des caractères non accentués et ne servait que pour le Telex.

Il permettait de transmettre les vingt-six lettres non accentuées, les dix chiffres, ainsi qu'une douzaine de symboles supplémentaires.

Vers 1901, le code Baudot original a été modifié par Donald Murray qui réorganisa les caractères, ajouta de nouveaux symboles, et introduisit les jeux de caractères.

Codage industriel, mécanique et télécommunications

Article détaillé : Tabulatrice.

Article détaillé : Codage des caractères sur carte perforée.

En 1890, on répartit les perforations arbitrairement sur la carte (recensement Hollerith 1890)^[1].

En raison de l'existence de brevets le codage Hollerith de IBM n'est pas utilisé par Bull. Ce manque d'interopérabilité fragmenta le marché entre "clients IBM" et "clients Bull".

Vers le changement de siècle développement d'un codage performant sur cartes perforées pour le recensement étatique. Le produit fut ensuite commercialisé par le Tabulating Machine Co.^[2].

Carte du recensement de Hollerith (1890)

- **Western Union** en utilisa une version modifiée jusque dans les années 1950. Avaient été rajoutées l'espace et la sonnerie.
- Dans les années 1930, le CCITT a introduit le International Telegraph Alphabet No. 2 (ITA2) code comme standard international, basé sur le code Western Union modifié. Les États-Unis ont standardisé une version de l'ITA2 comme American Teletypewriter code (USTTY) qui était une base de codes 5-bit teletypewriter jusqu'aux débuts de l'ASCII 7 bits en 1963.

Informatisation

En 1948, **Manchester Mark I**, le premier ordinateur électronique, utilisait les principes du **Code Baudot**.

- **RADIX-50** (1959) : 40 codes, stocke 3 caractères dans un mot de 16 bits en **big endian** (PDP-11). Il permet de stocker les dix chiffres, l'espace, vingt-six lettres, et les trois symboles « . », « \$ », « % ».
- **SIXBIT** (1968 et 1954) : 64 codes (6 bits), stocke six bits caractères par mot mémoire (PDP-10).

Six-bit BCD était utilisé par **IBM** sur les premiers ordinateurs comme le **IBM 704** en 1954^[3]. Par la suite, il donne naissance à l'**EBCDIC**.

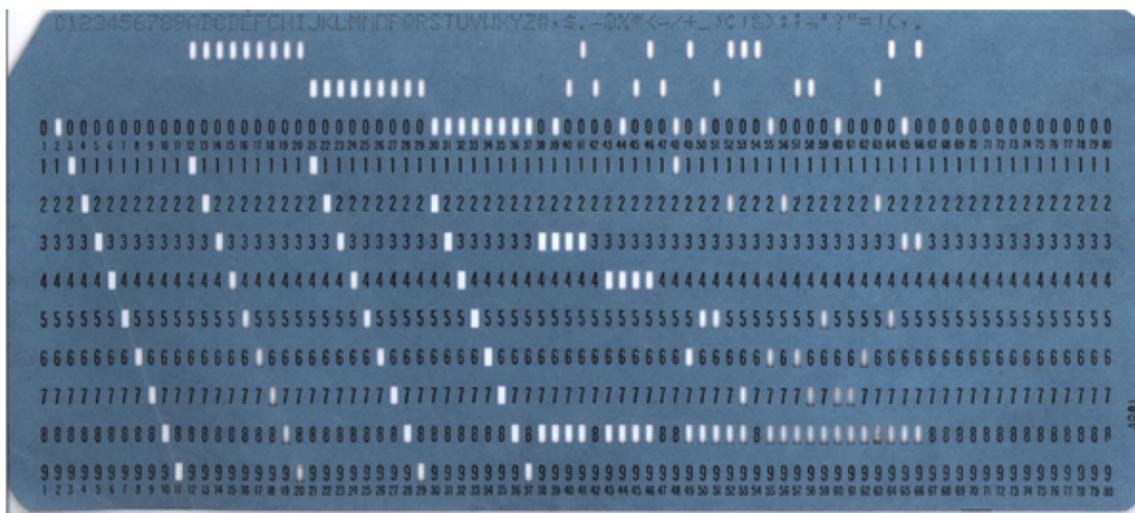
Voir aussi **BCD** (6-bit) et **Fieldata**.

Les années 1960

En 1960, **IBM**, **Univac**, **Burrough**, **Honeywell** et d'autres, se sont regroupés en consortium pour définir un standard commun^[4].

Dans les années 1960, dans un contexte de guerre froide, apparaissent le **GHOST** et l'**ASCII**.

- 1963 : naissance de l'**ASCII** dans le bloc des États-Unis. Prise en charge des caractères latins et anglais, sur sept bits. Dépassement de la limite précédente de six bits.
- 1963 : naissance de l'**Extended Binary Coded Decimal Interchange Code** (**EBCDIC**) sur 8-bit.
- 1964 : naissance du **GOST** (pour Государственный стандарт, norme d'État) dans le bloc soviétique **GOST 10859**. Prise en charge des caractères cyrilliques et anglais.



Une carte perforée 80 colonnes d'IBM d'un des types les plus utilisés au XX^e siècle. La photographie fait apparaître le jeu de caractère EBCDIC de 1964 : un chiffre est représenté par un trou, une lettre par deux et un symbole par trois. Le symbole affichable apparaît en noir bleuté sur fond bleu obscur sur le haut de la carte.

- 1968 : naissance de **MARC-8** (en) permettant de coder 15 000 caractères.

Naissance de l'ASCII En juillet 1971, le RFC 183, « The EBCDIC Codes and Their Mapping to ASCII » pose les bases de la conversion à l'ASCII, pour limiter les caractères utilisés.

Le standard **ASCII** (*American Standard Code for Information Interchange*) comporte 128 codes (7 bits). Quelques valeurs courantes :

- ASCII 10 : Saut de ligne abrégé « LF » pour Line Feed ;
- ASCII 32 : Espace ;
- ASCII 65 : Lettre A majuscule.

ASCII est standardisé en 1968 sous le nom « ANSI X3.4-1968 ». Puis les mises à jour suivent : ANSI X3.4-1977 et finalement ANSI X3.4-1986. On peut enfin le rencontrer sous le nom « US-ASCII ». Il existe en fait des dizaines de variantes de l'ASCII, mais c'est essentiellement la signification des codes de contrôles (caractères non imprimables) qui change.

Évolutions de l'ASCII vers les jeux de caractères codés sur 8 bits et la norme ISO 8859

Les variantes d'origine Le premier codage largement répandu fut l'ASCII. Pour des raisons historiques (les grandes sociétés associées pour mettre au point l'ASCII étaient américaines) et techniques (7 bits disponibles seulement pour coder un caractère), ce codage ne prenait en compte que 2^7 soit 128 caractères. De ce fait, l'ASCII ne comporte pas les caractères accentués, les cédilles, etc. utilisés par des langues comme le français. Ceci devint vite inadapté et un certain nombre de méthodes ad-hoc furent utilisées d'abord pour le modifier. Ainsi sont apparues des variantes multiples du codage ASCII, adaptées à des langues ou groupes de langues limités.

En 1972, pour mettre fin à leur prolifération, l'ISO a d'abord créé la norme **ISO/CEI 646** formalisant un code normalisé sur 7 bits, où des positions ont été désignées comme invariantes, et quelques-unes étaient autorisées à désigner d'autres caractères. Ce système a vite eu ses limites, car il n'était adapté qu'à un petit nombre de langues à écriture latine, et il y manquait de trop nombreux caractères, et ne permettait pas de représenter les écritures non latines.

Modification dynamique par séquence d'échappement Différentes solutions ont été envisagées : d'abord définir un système permettant de modifier dynamiquement le jeu de caractère codé utilisé ou le groupe de caractères utilisé dans une écriture complexe. Ce système a pu fonctionner pour conserver la compatibilité avec les systèmes

de communication restreints à 7 bits, et a donné lieu aux standards comme le **Videotex** ou les premiers systèmes de **Télétexte**, et reprenait en fait l'idée déjà en place pour les **télex**.

De leur côté, les pays asiatiques orientaux ont développé un système hybride utilisant des pages de code multiples, avec différentes représentations binaires selon que le système permettait de stocker les octets sur 8 bits ou sur 7 bits. L'ISO a aussi mis en place une norme destinée à leur interopérabilité, la norme **ISO 2022**, et pour l'Internet, l'IANA a mis en place un registre permettant de référencer tous ces codages.

Imposition du 8 bits Mais avec le développement des protocoles de communication sur Internet basés sur 8 bits, mais aussi pour les nombreux logiciels écrits en langues occidentales, il est apparu plus simple (au départ) de n'avoir à coder les caractères que sur un seul octet. Divers systèmes sont apparus d'abord sur des systèmes incompatibles entre eux, par exemple entre les premiers PC d'IBM et d'Apple (utilisant une puis de nombreuses pages de code), mais aussi les ordinateurs familiaux de divers marques, les imprimantes, le langage Postscript, ou encore les classiques *mainframes* d'IBM qui utilisaient encore l'EBCDIC sur une partie seulement des valeurs possibles sur 8 bits (avec là aussi de nombreuses variantes nationales mais dont les positions inutilisées depuis l'abandon du codage décimal des anciennes cartes perforées ont été utilisées pour étendre le jeu de caractères), ou encore les premiers systèmes Unix dont un produit par *Digital Equipment Corporation* (DEC) fera date en étant le vrai premier jeu de caractères « multinational » (le *Multinational Character Set* ou MCS) pour les langues à écriture latine, une idée qui sera aussi reprise aussi dans le développement de pages de codes nationales pour PC dans les premières versions de MS-DOS (produites par Microsoft et non IBM comme auparavant).

Ce développement a également eu lieu pour plusieurs autres écritures, mais avec de nombreuses variantes même au sein de chaque écriture : Apple popularisera le jeu de caractères codés **MacRoman** sur ses premiers Macintosh ainsi qu'une dizaine d'autres jeux similaires pour divers groupes de langues ou systèmes d'écriture, tandis que sur les PC subsisteront des pages de codes « multinationales » plus complètes mais encore insuffisantes pour couvrir convenablement les langues de plusieurs pays proches, et même parfois au sein de la même langue (par exemple entre les langues latines d'Europe occidentale, celles d'Europe centrale, et d'autres encore pour répondre aux besoins spécifiques de divers pays).

L'Union soviétique quant à elle, n'a pas accès officiellement aux systèmes d'exploitation développés essentiellement aux États-Unis en raison de l'embargo technologique à la fin de la **Guerre froide** (et donc pas aux discussions et décisions liées à leur développement) et développe sa propre norme **KOI8-R** pour l'écriture cyrillique en russe, et parvient à l'imposer en **Bulgarie** et en ex-**Yougoslavie** pour l'écriture cyrillique du serbo-croate, ainsi que dans d'autres pays utilisant cette écriture. Pour l'Ukraine elle crée une variante **KOI8-U** destinée à la langue ukrainienne, au départ utilisés sur les systèmes gouvernementaux bien avant que les PC s'imposent chez les particuliers. Ces jeux de caractères codés auront finalement (et ont encore) plus de succès que les pages de codes pour DOS développées par IBM et Microsoft, même dans les pays non soumis à l'embargo américain et qui commencent à utiliser aussi les PC d'IBM et MS-DOS, ou les jeux de caractères cyrilliques utilisés sur les premiers Macintosh d'Apple (souvent encore trop chers pour ces marchés).

En raison de l'utilisation de certains caractères locaux pour des programmations informatiques, une grande incohérence entre pays francophones est apparue. En 1983, cette incohérence a conduit la France à abandonner sa norme Z62010 au profit de l'Ascii.

Face à ce désordre qui nuisait à l'interopérabilité, l'ISO a développé une norme mieux étudiée, la norme **ISO/CEI 8859**, selon les principes des pages de code DOS mais cette fois compatible avec **ISO 646** et **ISO 2022**, mais là encore avec de multiples variantes, adaptées à des groupes de langue plus étendus qu'avec les pages de code pour PC. Une de ces variantes la plus connue est la page de code **ISO/CEI 8859-1**, développée sur la base du jeu de caractères MCS de DEC qui lui-même était basé sur la variante américaine (ASCII) du jeu ISO 646 normalisé mais seulement pour les langues européennes occidentales (le jeu dit « latin-1 »), les autres variantes étant créées de façon à réutiliser si possible les mêmes positions pour des caractères identiques ou proches lexicographiquement (cela n'a pas toujours été vrai pour toutes les variantes latines **ISO/CEI 8859-2**, etc. jusqu'à **ISO/CEI 8859-13**), y compris pour d'autres écritures à alphabet simples (grec, cyrillique, arabe, thaï).

Microsoft Windows intégrera cependant les caractères de la norme **ISO 8859** en abandonnant les caractères de contrôle supplémentaires codés au-delà des 128 premiers caractères ASCII (Windows abandonne toute idée d'interopérabilité avec EBCDIC), pour y coder quelques autres signes de ponctuation ou lettres manquantes dans certains des volets de la norme **ISO 8859** (puis en ajouter quelques autres, notamment le symbole de l'euro en 1998), réduisant ainsi le nombre de variantes nécessaires (notamment pour l'écriture latine), mais introduisant des problèmes d'interopérabilité notamment au niveau du symbole euro.

Émergence des systèmes multilingues

TRON Character Code et TRON Multilingual Environment Le Japon, qui a des besoins spécifiques a proposé un jeu de caractères multi-culturel dans la première moitié des années 1980s. Il s'agissait de projet **TRON** lancé en 1984.

Aujourd'hui, le TRON inclut les codages suivants :

- BTRON3 Character Sets (Partial) ;
- ISO 8859-1 ;
- GB 2312-80 ;
- JIS X 0208-1990 ;
- JIS X 0212-1990 ;
- KS C 5601-1992 ;
- Six-dot braille ;
- Eight-dot braille.

MULE et Emacs Dans la communauté Unix, le support des textes écrits dans des langues différentes était opéré à travers le logiciel **MULE** (en) d'Emacs à partir de 1993^[2]. Le système MULE prend en charge le ISO-2022-JP décrit dans le RFC 1554^[5].

Évolutions vers un jeu de caractères codés universel

De son côté, l'Inde a créé sa propre norme **ISCII**. La norme ISCII a été créée pour assurer la bonne interopérabilité d'une dizaine de jeux de caractères nécessaires pour autant de systèmes d'écriture différents. Elle aurait pu rentrer dans la norme ISO 8859 si l'ISCII avait utilisé la norme ISO 2022 pour basculer d'un jeu à l'autre ; de toute façon une telle intégration mise à jour de l'ISCII pour assurer cette compatibilité aurait plus nuit à l'interopérabilité des écritures indiennes, l'Inde ayant depuis aussi décidé que toute nouvelle extension se ferait via le jeu de caractères universel en développement et activement supporté par l'Inde. Il devenait urgent de définir un jeu de caractères codés dit « universel ».

L'ISO met en place un comité technique, mais tarde à publier sa norme (**ISO/CEI 10646**), et de plus un autre organisme de normalisation privé (le **Consortium Unicode**) se met en place sous l'impulsion de plusieurs éditeurs de logiciels : son but est de rendre utilisable le jeu de caractères universel en ajoutant des propriétés, mais dans un premier temps les deux organismes ne se mettent pas d'accord sur le codage de l'hangul (l'alphabet utilisé pour transcrire le coréen). De plus Unicode ne croit pas au début que plus de 65536 caractères seront nécessaires dans un seul plan, le consortium n'ayant au départ comme objectif que de couvrir que les systèmes d'écritures des langues modernes (les plus économiquement rentables), afin de les coder le plus rapidement possible.

La norme **Unicode 1.0** voit le jour, mais est en partie incompatible avec la première version de l'ISO 10646, ce qui freine finalement le développement (d'autant que le Consortium Unicode convient rapidement que plus d'un plan sera finalement nécessaire, en observant les travaux de codification des sinogrammes faits par un comité technique spécifique **UniHan**). Différentes solutions sont envisagées et **UTF-16** est expérimenté (mais pas encore standardisé par Unicode, ni approuvé dans la norme ISO/CEI 10646), tandis qu'apparaissent d'autres difficultés techniques d'interopérabilité avec **UTF-32** (dont le développement initial a lui aussi connu plusieurs variantes) et **UTF-8** (l'ISO/CEI 10646 en normalise une variante légèrement différente de celle définie par le standard Unicode !).

De leur côté, les systèmes d'exploitation et logiciels pour PC ou Mac ne sont pas encore prêts à supporter des jeux de caractères codés sur plusieurs octets, et d'autres variantes de l'ISO 8859 voient le jour, par exemple l'ISO/CEI 8859-14 (nécessaire pour les langues celtiques dont l'irlandais, une des langues officielles de l'Union européenne), l'ISO/CEI 8859-15 (introduisant en 1998 le symbole € de l'euro qui allait devenir la monnaie unique d'une douzaine de pays de l'Union européenne), et l'ISO/CEI 8859-16 (palliant l'absence de caractères nécessaires à la transcription correcte du roumain, à savoir des caractères utilisant une virgule souscrite et non une cédille).

Ce sera la dernière partie ajoutée à l'ISO 8859, l'ISO estimant que les besoins pour les autres langues étaient déjà couverts par les autres normes nationales (**ISCII** en Inde, **VISCII** au Viet Nam, **GB2312** en Chine, **EUC-JP** au Japon,

etc.) et préférant se consacrer au développement du jeu de caractères universel plutôt que d'en définir de nouveaux, codés sur 7 ou 8 bits et complètement incompatibles avec ceux déjà largement déployés dans ces pays, les autres pays ayant déjà décidé d'adopter pour leurs langues directement le jeu universel de caractères codés. Cet abandon sera d'autant plus facilité que les normes ISO 10646 et Unicode ont rapidement décidé de fusionner leurs répertoires et y sont parvenues rapidement en produisant une mise à jour majeure pour **Unicode 1.1** (rendant obsolète la version 1.0 du standard Unicode incompatible, mais en intégrant la forme de codage UTF-16 en tant que partie intégrante de son standard) et ISO 10646-1 (compatible avec ISO 10646, mais en abandonnant l'idée de supporter à l'avenir plus de 17 plans, et en acceptant d'intégrer et normaliser UTF-16), et en créant des procédures permettant aux deux comités techniques de collaborer et assurer leur coordination. De plus les systèmes d'exploitation, logiciels, polices de caractères et protocoles pour l'Internet ont également évolué pour accepter nativement le jeu universel.

La norme ISO 10646 qui définit le jeu de caractères universel (conjointement avec le standard Unicode lui ajoutant des propriétés facilitant le traitement) a largement réussi son ambition de coder la majorité des systèmes d'écriture utilisés dans le monde (il reste encore à coder quelques écritures complexes ou mal connues), et conserve son ambition de couvrir toutes les écritures utilisées dans l'histoire de l'humanité (donc y compris les écritures anciennes, ou des caractères rares dans des écritures déjà codées, ou de nouveaux caractères qui apparaîtraient dans l'usage courant dans certains pays du monde ou qui seraient rendus nécessaires pour certains traitements afin de lever des ambiguïtés).

Le répertoire universel (d'ISO 10646-1 et Unicode) contenait en fin 2009 environ 100 000 caractères (dont plus de la moitié pour les seuls sinogrammes), chacun possédant un unique code entier compris entre 0 et 10FFFF en hexadécimal (soit un peu plus de 1,1 million - il existe donc des entiers auxquels aucun caractère n'est associé). Deux autres répertoires fréquemment utilisés, les jeux de caractères codés ASCII (ISO 646, dans sa dernière variante américaine normalisée en 1984) et ISO 8859-1, sont respectivement identiques aux 128 et 256 premiers caractères codés dans ce répertoire universel.

Retour aux plages réduites

La popularisation des messages courts de 160 octets, dits SMS, de par leur taille réduite et leur coût, a conduit à la réapparition des techniques de codage de texte sur des *bytes* courts de sept bits. (**GSM 03.38 (en)**).

6.5.3 Jeux de caractères codés populaires, par pays

Sur Internet, l'UTF-8 et l'ASCII sont les deux encodages les plus populaires depuis 2010^[6]. En juillet 2012, leur utilisation est estimée à 80 %, (65 % + 15 %) contre 10 % environ pour les encodages occidentaux (latin1). L'utilisation des autres encodages est inférieure à 10 % sur internet.

Amérique

États-Unis

- ISO 646 ;
- ASCII.

Canada

- ASCII, standard de compatibilité, sans accent ;
- ISO 8859-1 ;
- Unicode, voir aussi UTF-8.

Europe

Union européenne

- UTF-8, notamment utilisé par <http://www.europa.eu/>.

France

- ISO 646-variante de référence internationale (*IRV, alias ASCII*), standard de compatibilité, sans accent, pour le développement de logiciel en anglais ;
- « Jeu de caractères graphiques français pour la langue française », 94 caractères codés sur 7 bits, normalisé ISO/CEI 646–69 le 1^{er} juin 1983, et normalisé AFNOR NF Z 62 010 en 1982. Aujourd'hui désuet, mais activable par séquence d'échappement ;
 - standard de compatibilité ;
 - dans l'une de ses deux variantes françaises (AFNOR) ;
 - standard du minitel (voir [viewdata](#)) ;
- ISO-2022 peu utilisé en France mais reprend tous les caractères français ;
- CP850 à l'époque du DOS ;
- Windows-1252, originaire de Microsoft Windows, encore appelé CP1252 ou Ansinew ;
- ISO 8859-1, avant l'euro ;
- ISO 8859-15, après l'euro ;
- Unicode, voir aussi UTF-8 (reprend tous les caractères des jeux de code précédent).

Historique des codages de caractères français ↑ → ↓

Grèce

- UTF-8 ;
- ISO 8859-7 (sans l'euro).

Roumanie

- ISO 8859-2 ;
- ISO 8859-16 ;
- UTF-8.

Lituanie

- Windows-1257, pour Windows, utilisé par <http://www.lrv.lt/> ;
- UTF-8, utilisé pour la langue française par <http://www.president.lt/fr/>.

Russie

- KOI8-R ;
- UTF-8.

Ukraine, Bulgarie

- KOI8-U ;
- UTF-8.

Afrique

Maghreb

- UTF-8, utilisé sur <http://www.maroc.ma/portailinst/Ar> ;
- Windows-1256, utilisé par exemple par <http://www.almaghribia.ma/>.

Asie

- ISCII (alphasyllabaires utilisés en Inde, au Sri Lanka et au Bangladesh) ;
- VISCII (alphabet latin moderne du vietnamien) ;
- TIS-620 (écriture thaïe) ;
- KSC 5601 (alphabet hangûl du coréen) ;
- GB18030 (écriture sinographique simplifiée des langues chinoises) ;
- Big5 (écriture sinographique traditionnelle des langues chinoises) ;
- Shift-JIS (syllabaires et écriture sinographique traditionnelle des langues japonaises).

6.5.4 Autres codages de caractères populaires

- Codage SMS ;
- ISO 646 : ASCII ;
- EBCDIC ;
- ISO 8859 : ISO 8859-1, ISO 8859-2, ISO 8859-3, ISO 8859-4, ISO 8859-5, ISO 8859-6, ISO 8859-7, ISO 8859-8, ISO 8859-9, ISO 8859-10, ISO 8859-11, ISO 8859-13, ISO 8859-14, ISO 8859-15, ISO 8859-16 ;
- Jeux de caractères codés DOS : CP437, CP737, CP850, CP852, CP855, CP857, CP858, CP860, CP861, CP863, CP865, CP866, CP869 ;
- Jeux de caractères codés Windows : Windows-1250, Windows-1251, Windows-1252, Windows-1253, Windows-1254, Windows-1255, Windows-1256, Windows-1257, Windows-1258 ;
- VISCII ;
- KOI8-R, KOI8-U ;
- TIS-620 ;
- ISCII ;
- ISO 2022, EUC ;
 - Big5 : HKSCS ;
 - Guobiao : GB2312, GBK, GB18030 ;
 - Shift-JIS ;
- ISO/CEI 10646 : Unicode, voir aussi UTF-8.

6.5.5 Notes et références

- [1] <http://www.kerleo.net/computers/mecanographie.htm>
- [2] Steven J. Searle, « A Brief History of Character Codes in North America, Europe, and East Asia », Sakamura Laboratory, University Museum, University of Tokyo, 6 août 2004
- [3] (en) IBM Corporation, 704 electronic data-processing machine : manual of operation, 1954 (lire en ligne), p. 35
- [4] Jacques André, « Caractères, codage et normalisation : de Chappé à Unicode », *Document numérique*, Lavoisier, vol. 6, n° 3-4 « Unicode, écriture du monde ? », 2002, p. 13-49 (ISBN 2-7462-0594-7, lire en ligne)
- [5] <http://www.ietf.org/rfc/rfc1554.txt>
- [6] <http://spectrum.ieee.org/telecom/standards/will-unicode-soon-be-the-universal-code>

6.5.6 Voir aussi

Bibliographie

- (en) Eric Fischer, The Evolution of Character Codes, 1874-1968, juin 2000, 33 p. (présentation en ligne, lire en ligne)
- Comité consultatif international télégraphique et téléphonique, Alphabet international de référence (ancien alphabet international n° 5 ou AI5) : Technologie de l'information – Jeux de caractères codés à 7 bits pour l'échange d'informations, Union internationale des télécommunications, coll. « Recommandations / Série T » (n° T.50 (09/92)), avril 1993, 22 p. (présentation en ligne, lire en ligne)
- (en) Victor Stinner, Programming with Unicode, 2011 (présentation en ligne, lire en ligne), chap. 5 (« Charsets and encodings »)

Liens externes

- (en) Character Sets, sur le site de l'Internet Assigned Numbers Authority
- (en) Unicode Character Encoding Model, sur le site du Consortium Unicode
- (en) The Punched Card, sur la page personnelle de John Savard
-  Portail de l'informatique
-  Portail de l'écriture

6.6 ISO 8859

ISO 8859, également appelée plus formellement **ISO/CEI 8859**, est une norme commune de l'ISO et de la CEI de codage de caractères sur 8 bits pour le traitement informatique du texte. Le standard est divisé en parties numérotées publiées séparément, telles que ISO/CEI 8859-1, ISO/CEI 8859-2, etc., chacune pouvant être référencée de façon informelle comme norme en tant que telle. La norme comprend actuellement 16 parties.

6.6.1 Introduction

Alors que les 96 caractères ASCII imprimables suffisent à l'échange d'informations en anglais courant, la plupart des autres langues qui utilisent l'alphabet latin ont besoin de symboles additionnels non couverts par l'ASCII, tels que *ß* (allemand), *å* (suédois et d'autres langues nordiques). ISO 8859 a cherché à remédier à ce problème en utilisant le huitième bit de l'octet, pour donner de la place à 128 caractères supplémentaires. (Ce bit était jadis utilisé pour

le contrôle de l'intégrité des données (bit de parité), ou était inutilisé.) Cependant, il fallait plus de caractères qu'on n'en pouvait mettre dans un jeu de caractères 8 bits, aussi plusieurs tables de correspondances ont été développées, en incluant au moins 10 tables pour couvrir uniquement l'écriture latine.

La norme ISO 8859-*n* n'est pas totalement identique aux codages de caractères bien connus ISO-8859-*n* approuvés par l'IANA pour l'utilisation sur l'Internet. Au-delà du trait d'union ajouté présent dans le nom approuvé par l'IANA, les codages diffèrent de sorte que chaque partie de la norme ISO assigne, au maximum, 191 caractères dans les étendues d'octets 32 à 126 et 160 à 255, alors que le codage de caractère correspondant approuvé par l'IANA fusionne ces tables de correspondances avec le jeu de contrôle C0 (caractères de contrôles positionnés aux octets de 0 à 31) de l'ISO 646 et le jeu de contrôle C1 (caractères de contrôles positionnés aux octets de 127 à 159) de l'ISO 6429, ce qui conduit à des tables de caractères 8 bits complètes, avec la plupart, sinon la totalité, des valeurs d'octets assignées. On utilise généralement ISO-8859-*n* comme nom MIME pour ces jeux de caractères. Beaucoup de gens emploient les termes ISO 8859-*n* et ISO-8859-*n* l'un pour l'autre.

6.6.2 Caractères

La norme ISO 8859 est conçue pour l'échange fiable d'informations, et non pour la typographie ; elle omet des symboles nécessaires pour la typographie de haute qualité, tels que les ligatures optionnelles, les guillemets incurvés, les tirets, etc. De ce fait, les systèmes de composition avancés utilisent souvent des extensions propriétaires ou idiosyncratiques au-delà des standards ASCII et ISO 8859, ou utilisent plutôt Unicode.

En général, si un caractère ou le symbole ne faisait pas déjà partie d'un jeu de caractères informatique largement utilisé et n'était pas non plus disponible sur la plupart des claviers de machines à écrire pour une langue nationale, il n'était pas inclus. Par conséquent, les guillemets doubles directionnels (« , ») utilisés pour quelques langues européennes ont été inclus, mais pas les guillemets doubles directionnels (“ , ”) utilisés pour l'anglais et quelques autres langues.

Le français n'a pas obtenu ses ligatures <Œ, œ> parce que les francophones n'avaient pas eu suffisamment besoin d'elles pour les exiger sur leurs claviers ; ni le <ÿ, ÿ>, bien que ce caractère soit utilisé dans des noms propres de famille et de lieux. Cependant, ces caractères ont été inclus plus tard dans l'ISO 8859-15, qui a introduit aussi le nouveau caractère de l'Euro €. De même le néerlandais n'a pas obtenu les lettres <IJ, ij> parce que les néerlandophones s'étaient habitués à la dactylographie de ces lettres plutôt comme deux caractères.

Au début, le roumain n'a pas obtenu ses lettres <Ș, ș> et <Ț, ț>, parce que ces lettres ont d'abord été unifiées avec <Ş, ş> et <Ț, ț> par le Consortium Unicode qui considérait les formes avec virgule souscrite comme des variantes de glyphe des formes avec cédille.

Cependant, les lettres avec virgule souscrite ont été ajoutées ultérieurement à la norme Unicode et sont aussi dans ISO 8859-16.

La plupart des codes ISO 8859 fournissent des signes diacritiques requis pour diverses langues européennes. D'autres fournissent des alphabets non-romains : grec, cyrillique, hébreu, arabe et thaï. La plupart des codages ne contiennent que des caractères avec chasse, bien que les codages hébreu et arabe contiennent aussi des caractères combinatoires. Cependant, la norme ne fournit rien pour les caractères de langues de l'est asiatique (CJC), car leur système d'écriture idéographique exige de nombreux milliers de points de code.

Bien qu'il utilise à la base des caractères latins, le vietnamien, lui aussi, n'est pas susceptible d'être représenté en 96 positions (sans utiliser des diacritiques combinatoires). Les syllabaires japonais Kana, quant à eux, le pourraient, mais, comme plusieurs autres alphabets du monde, ne sont pas codés dans le système ISO 8859.

6.6.3 Les parties du standard ISO 8859

ISO 8859 est constitué à ce jour des parties suivantes :

- ISO 8859-1 (*latin-1* ou *européen occidental*) — probablement la partie la plus largement utilisée de ISO 8859, couvrant la plupart des langues européennes occidentales : l'allemand, l'anglais, le basque, le catalan, le danois, l'écosseis, l'espagnol, le feringien, le finnois (partiellement²), le français (partiellement, il manque notamment l'*e dans l'o* œ et le ÿ²), l'islandais, l'irlandais, l'italien, le néerlandais (partiellement¹), le norvégien, le portugais, le rhéto-roman et le suédois, certaines langues européennes sud-orientales (l'albanais), ainsi que des langues africaines (l'afrikaans et le swahili). Le symbole de l'euro et la capitale Ÿ, qui manquaient, sont dans la version révisée ISO 8859-15 (latin-9). Le jeu de caractères correspondant ISO-8859-1, approuvé par l'IANA, est le

codage par défaut des anciens documents **HTML** ou des documents transmis par messages **MIME**, tels que les réponses **HTTP** quand le type de média du document est « text » (par exemple les documents « text/html »).

- ISO 8859-2 (*latin-2* ou *européen central*) — prend en charge celles des langues d'Europe centrale ou de l'Est basées sur un alphabet romain. Ceci inclut le bosnien, le croate, le polonais, le tchèque, le slovaque, le slovène et le hongrois. Le symbole de l'euro manquant est présent dans la version ISO 8859-16.
- ISO 8859-3 (*latin-3* ou *européen du Sud*) — le turc, le maltais, et l'espéranto ; largement supplanté par ISO 8859-9 pour le turc, et par Unicode pour l'espéranto.
- ISO 8859-4 (*latin-4* ou *européen du Nord*) — l'estonien, le letton, le lituanien, le groenlandais, et le sami.
- ISO 8859-5 (*cyrillique*) — Couvre la plupart des langues slaves utilisant un alphabet cyrillique, y compris le biélorusse, le bulgare, le macédonien, le russe, le serbe et l'ukrainien (partiellement³).
- ISO 8859-6 (*arabe*) — Couvre les caractères les plus courants de l'arabe. Ne prend pas en charge d'autres langues à alphabet arabe. Nécessite un moteur de rendu qui prend en charge l'affichage bi-directionnel et l'analyse contextuelle.
- ISO 8859-7 (*grec*) — Couvre la langue grecque moderne (orthographe monotonique). Peut être utilisé aussi pour le grec ancien écrit sans accents ou dans l'orthographe monotonique, mais il manque les signes diacritiques pour l'orthographe polytonique. Ceux-ci ont été introduits avec Unicode.
- ISO 8859-8 (*hébreu*) — Couvre l'alphabet hébraïque moderne tel qu'il est utilisé en Israël. En pratique, deux codes différents existent : ordre logique (nécessite un moteur de rendu bi-directionnel pour l'affichage) et ordre visuel (gauche à droite).
- ISO 8859-9 (*latin-5* ou *turc*) — Grosso modo le même que l'ISO 8859-1, où les lettres islandaises peu utilisées sont remplacées par des lettres turques. Il est aussi utilisé pour le kurde.
- ISO 8859-10 (*latin-6* ou *nordique*) — Un réarrangement du latin-4. Considéré plus utile pour les langues nordiques. Les langues baltes utilisent plus souvent le *latin-4*.
- ISO 8859-11 (*thai*) — Contient la plupart des glyphes requis pour la langue *thai*.
- ISO 8859-12 — Était supposé couvrir l'alphabet devanāgarī, mais ce projet a été abandonné en 1997. ISCII et Unicode/ISO/CEI 10646 couvrent le devanāgarī.
- ISO 8859-13 (*latin-7* ou *balte*) — Ajoute quelques caractères supplémentaires pour les langues baltes qui manquaient en latin-4 et latin-6.
- ISO 8859-14 (*latin-8* ou *celtique*) — Couvre des langues celtiques telles que l'irlandais (orthographe traditionnelle), le gaélique écossais, le mannois (langue disparue) et le breton (certaines anciennes orthographes).
- ISO 8859-15 (*latin-9* ou parfois de façon impropre *latin-0*) — une révision de 8859-1 qui abandonne quelques symboles peu utilisés, les remplaçant avec le symbole de l'euro € (remplaçant le signe de monnaie ₤) et les lettres Š, š, Ž, ž, Œ, œ, et Ÿ, ce qui complète la couverture du français, du finnois et de l'estonien.
- ISO 8859-16 (*latin-10* ou *européen du Sud-Est*) — Prévu pour l'albanais, le croate, le hongrois, l'italien, le polonais, le roumain et le slovène, mais aussi le finnois, le français, l'allemand et l'irlandais (en nouvelle orthographe). Cette police mise plus sur les lettres que les symboles.

Notes :

- ¹ : seul le *IJ*/*ij* (Dutch *Y*) manque, et peut être représenté comme *IJ*.
- ² : les caractères manquants sont dans l'ISO 8859-15.
- ³ : les caractères manquants “*Ŧ*/*ŧ*” ont été réintroduits en ukrainien en 1991, mais il y manque aussi le nouveau symbole (défini en 2004) de la devise monétaire ukrainienne, la hryvnia, ainsi que celui de l'euro ; la norme ISO 8859 étant désormais fermée, l'Ukraine a produit un autre jeu de caractères sur 8 bits dérivé de l'ancienne variante ukrainienne KOI-8U de l'ancienne norme soviétique (en voie d'obsolescence) mais utilise de plus en plus soit Unicode, soit le jeu de caractères Windows « ANSI cyrillique » (page de code 1253).

Chaque partie d'ISO 8859 est conçue pour représenter des langues qui empruntent souvent l'une de l'autre, ainsi les caractères nécessaires pour chaque langue sont d'habitude accommodés par une partie seule. Cependant, il y a quelques combinaisons de caractères et langue qui ne sont pas accommodées sans les transcriptions. Les efforts ont été faits pour rendre les conversions le plus facile possible. Par exemple, l'allemand a tous ses sept caractères spéciaux aux mêmes positions dans toutes les variantes latines (1-4, 9-10, 13-16), et en beaucoup de positions les caractères diffèrent seulement par les signes diacritiques qui varient entre les séries. En particulier, les variantes 1-4 ont été conçues conjointement et ont la propriété que chaque caractère codé apparaît à une seule position donnée ou pas du tout.

Table comparative des diverses parties d'ISO 8859

La position 0xA0 est toujours attribuée à l'espace insécable (*non breaking space*, *NBSP*) et la position 0xAD (sauf dans ISO 8859-11) au trait d'union conditionnel (*soft hyphen*, *SHY*), qui est seulement visible aux coupures de ligne. Les autres champs vides sont soit *non attribués* soit des caractères que le système utilisé n'est pas capable d'afficher.

6.6.4 Relations avec Unicode et UCS

Depuis 1991, le consortium Unicode a travaillé avec l'ISO pour développer Unicode et ISO/CEI 10646 (UCS : Universal Character Set) en tandem. Cette paire de normes a été créée pour unifier, entre autres, le répertoire de caractère ISO 8859, en associant, initialement, chaque caractère à une valeur de code sur 16 bits ; quelques valeurs de code restant non attribuées.

Progressivement, leurs modèles ont été adaptés pour permettre la mise en correspondance des caractères avec des points de code numériques abstraits plutôt qu'avec des valeurs de longueur fixe. Ce qui a permis à davantage de points de code et de méthodes de codage d'être supportées.

Unicode et l'ISO/CEI 10646 assignent actuellement de l'ordre de 100 000 caractères à un espace de code réservé qui consiste en plus d'un million de points de code (de 0x0 à 0x10FFFF en hexadécimal). Chaque caractère est noté *U+nnnn*, où *nnnn* représente une suite de 4 à 6 chiffres hexadécimaux. À titre d'exemple, le caractère A correspond au code *U+0041*. Unicode et l'ISO/CEI 10646 définissent également plusieurs normes de codage capables de représenter la totalité des points de code disponibles.

Les normes de codage Unicode et UCS utilisent des séquences de un à quatre codes de 8 bits (UTF-8), des séquences de un ou deux codes de 16 bits (UTF-16), ou une valeur de code de 32 bits (UTF-32 ou UCS-4). Il existe également un codage plus ancien qui utilise une valeur de 16 bits (UCS-2) mais qui n'est capable de représenter qu'un dix-septième des points de code disponibles : sur la totalité des points de code définis par Unicode, seuls ceux du BMP (*Basic Multilingual Plane*), c'est-à-dire ceux compris entre *U+0000* et *U+FFFF*, sont supportés par UCS-2. Parmi ces formes de codages, seul UTF-8 présente les suites d'octets dans un ordre fixe ; les autres sont sujets au problème lié à la manière dont les plates-formes codent les suites de 16 et 32 bits (voir sur ce point l'opposition *gros-boutiste* et *petit-boutiste*). Ce problème peut être résolu grâce à des codes spéciaux ou par des indications hors bande.

Les éditions d'ISO 8859 les plus récentes expriment les caractères en termes de leurs noms Unicode/UCS et de la notation *U+nnnn*. Ceci permet à chaque partie d'ISO 8859 d'être un schéma de codage de caractères Unicode/UCS qui permet de coder une petite partie de l'UCS par des octets. D'ailleurs, les 256 premiers caractères en Unicode et en UCS sont identiques à ceux de ISO 8859-1.

ISO 8859 a été favorisé dans les années 1990, car il avait les avantages d'être bien établi et plus facilement implémenté au niveau des logiciels : la correspondance d'un octet à un caractère est simple et adaptée à la plupart des applications ne supportant qu'une seule langue, et où il n'y a pas de caractères combinés ou de formes variantes.

Avec la baisse du coût relatif, en ressources informatiques, de l'utilisation de plus d'un octet par caractère, les langages de programmation et les systèmes d'exploitation proposent un support natif d'Unicode en plus de leur système de page de code. Les systèmes d'exploitation supportant Unicode devenant plus répandus, ISO 8859 et les autres codages anciens ont perdu de leur popularité. Alors qu'ISO 8859 et les modèles de caractères sur un seul octet restent ancrés dans de nombreux systèmes d'exploitation, langages de programmation, systèmes de stockage de données, applications réseau, matériels d'affichage et logiciels d'application pour utilisateur final, les applications informatiques les plus modernes utilisent Unicode en interne et s'appuient, lorsque c'est nécessaire, sur des tables pour convertir de ou vers un codage plus simple.

6.6.5 État de développement

La norme ISO/CEI 8859 était mise à jour par le groupe ISO/CEI Joint Technical Committee 1, Subcommittee 2, Working Group 3 (ISO/CEI JTC 1/SC 2/WG 3) mais en juin 2004, WG 3 s'est éteint et ses tâches ont été données au groupe SC 2. La norme n'est plus mise à jour car le seul groupe de travaux actif, WG 2, doit se concentrer au développement de ISO/CEI 10646.

6.6.6 Voir aussi

Références externes

- Les copies des versions publiées des normes sont à vendre au [ISO catalogue site](#) et au [ANSI eStandards Store](#).
- Les versions PDF des propositions finales de quelques sections d'ISO/ 8859 comme soumis pour revue & publication par ISO/CEI JTC 1/SC 2/WG 3 se trouvent au site [WG 3](#) :
 - [PDF] ISO/CEI 8859-1 :1998 - 8-bit single-byte coded graphic character sets, Part 1 : Latin alphabet No. 1 (*draft dated February 12, 1998, published April 15, 1998*)
 - [PDF] ISO/CEI 8859-4 :1998 - 8-bit single-byte coded graphic character sets, Part 4 : Latin alphabet No. 4 (*draft dated February 12, 1998, published July 1, 1998*)
 - [[PDF] <http://www.open-std.org/jtc1/sc2/open/02n3329.pdf> ISO/CEI 8859-7 :1999] - 8-bit single-byte coded graphic character sets, Part 7 : Latin/Greek alphabet (*draft dated 10 juin 1999 ; superseded by ISO/CEI 8859-7 :2003, published 10 octobre 2003*)
 - [PDF] ISO/CEI 8859-10 :1998 - 8-bit single-byte coded graphic character sets, Part 10 : Latin alphabet No. 6 (*draft dated February 12, 1998, published July 15, 1998*)
 - [PDF] ISO/CEI 8859-11 :1999 - 8-bit single-byte coded graphic character sets, Part 11 : Latin/Thai character set (*draft dated June 22, 1999 ; superseded by ISO/CEI 8859-11 :2001, published Dec 15, 2001*)
 - [PDF] ISO/CEI 8859-13 :1998 - 8-bit single-byte coded graphic character sets, Part 13 : Latin alphabet No. 7 (*draft dated April 15, 1998, published 15 octobre 1998*)
 - [PDF] ISO/CEI 8859-15 :1998 - 8-bit single-byte coded graphic character sets, Part 15 : Latin alphabet No. 9 (*draft dated August 1, 1997 ; superseded by ISO/CEI 8859-15 :1999, published 15 mars 1999*)
 - [PDF] ISO/CEI 8859-16 :2000 - 8-bit single-byte coded graphic character sets, Part 16 : Latin alphabet No. 10 (*draft dated 15 novembre 1999 ; superseded by ISO/CEI 8859-16 :2001, published 15 juillet 2001*)
- Les normes ECMA, écrites avec l'intention de correspondance exacte aux normes de polices de caractères ISO/CEI 8859, se trouvent au :
 - Standard ECMA-94 : 8-Bit Single Byte Coded Graphic Character Sets - Latin Alphabets No. 1 to No. 4 *2^e édition (juin 1986)*
 - Standard ECMA-113 : 8-Bit Single-Byte Coded Graphic Character Sets - Latin/Cyrillic Alphabet *3^e édition (décembre 1999)*
 - Standard ECMA-114 : 8-Bit Single-Byte Coded Graphic Character Sets - Latin/Arabic Alphabet *2nd édition (décembre 2000)*
 - Standard ECMA-118 : 8-Bit Single-Byte Coded Graphic Character Sets - Latin/Greek Alphabet (*décembre 1986*)
 - Standard ECMA-121 : 8-Bit Single-Byte Coded Graphic Character Sets - Latin/Hebrew Alphabet *2^e édition (décembre 2000)*
 - Standard ECMA-128 : 8-Bit Single-Byte Coded Graphic Character Sets - Latin Alphabet No. 5 *2^e édition (décembre 1999)*
 - Standard ECMA-144 : 8-Bit Single-Byte Coded Character Sets - Latin Alphabet No. 6 *3^e édition (décembre 2000)*
- Les tables de correspondance ISO/CEI 8859-1 à Unicode [tables de correspondance](#) en fichiers de texte ordinaire se trouvent au site FTP de Unicode.
- Les descriptions informelles et les tables de code pour la plupart des normes ISO 8859 se trouvent au ISO 8859 Alphabet Soup (Miroir)

Articles connexes

- Blocs de caractères Unicode contenant des caractères codés dans au moins une des 16 parties d'ISO 8859 :
 - pour ISO 8859-1 :
 - Latin basique
 - Supplément latin-1
 - pour les autres parties :
 - Latin étendu - A
 - Latin étendu - B
 - Diacritiques combinants
 - Grec et copte
 - Cyrillique
 - Hébreu
 - Arabe
 - Thaï
 - Symboles monétaires

-  Portail de l'informatique

-  Portail de l'écriture

6.7 American Standard Code for Information Interchange

 « ASCII » redirige ici. Pour les autres significations, voir [ASCII \(homonymie\)](#).

[[Fichier:ASCII full.svg|frame|Les 95 caractères ASCII affichables :

```
!"#$%&'()*+,-./ 0123456789 :;<=> ? @ABCDEFGHIJKLMNO PQRSTUVWXYZ[\]^_ `abcdefghijklmnopqrstuvwxyz{|}~
```

]]

Cette section doit être **recyclée**. Une réorganisation et une clarification du contenu sont nécessaires. Discutez des points à améliorer en page de discussion.

L'**American Standard Code for Information Interchange** (*Code américain normalisé pour l'échange d'information*), plus connu sous l'acronyme **ASCII** ([aski:]) est une norme de codage de caractères en informatique ancienne et connue pour son influence incontournable sur les codages de caractères qui lui ont succédé. Elle était la plus largement compatible pour ce qui est des caractères latins non accentués.

ASCII contient les caractères nécessaires pour écrire en anglais. Les caractères accentués sont fournis par d'autres normes, comme aujourd'hui, l'Unicode. Le jeu de caractères codés ASCII est le principal système qui a permis l'échange de textes en anglais à un niveau mondial, limitant ainsi l'usage des langues locales au travers d'extensions régionales.

L'ASCII est une des variantes de l'ISO/CEI 646. Il est inclus dans plusieurs dizaines de normes et standards augmentant les caractères disponibles ou extensions (ASCII étendu) Celles-ci peuvent être régionales (ISO/CEI 8859), nationales (GB 18030) ou internationales (ISO 2022, ISO/CEI 10646 ou Unicode). Avec l'avènement de la mondialisation et de l'internationalisation des systèmes d'information, les limitations de l'ASCII ne sont plus acceptées que dans des domaines très techniques qui requièrent la compatibilité avec des protocoles ou systèmes anciens. Malheureusement, l'ASCII n'a donné naissance qu'à des normes de codage incompatibles entre elles.

Dans la norme Unicode, le standard ASCII est défini sous le nom de « C0 Controls and Basic Latin ».

6.7.1 Histoire

Avant la standardisation, de nombreux codages de caractères incompatibles entre eux existaient^[1].

En 1960, l'ISO a créé le *technical committee on computers and information processing*^[2]. Il a été divisé en six groupes de travail.

- Working Group A : Glossary
- Working Group B : Character Sets and Coding
- Working Group C : Character Recognition
- Working Group D : Input and Output Media
- Working Group E : Programming Languages
- Working Group F : Digital Data Transmission

L'*American Standards Association* était chargé du standard des États-Unis.

L'ANSI a reconnu le consortium *Business Equipment Manufacturers Association* (BEMA, puis, CBEMA) comme le parrain du travail de standardisation du traitement des données. En 1960 BEMA a formé un groupe de traitement des données des partenaires, dont Minneapolis-Honeywell. Ce groupe a formé un *Plans and Policies Committee*, qui à son tour a formé l'*Engineering Committee*. L'*Engineering Committee* a formé le comité X3, qui a été reconnu par l'ASA comme *sectional committee*.

La liste des caractères à considérer et leur position ont été débattues^[3].

Elle a été inventée par l'américain Bob Bemer en 1961.^[réf. nécessaire]

En 1961, le DoD met au point un code standard de transmission de donnée sur 8 bits^[4]. Ce standard 8 bits est une variante des standards **FIELDATA** sixbits utilisés dans la décennie précédente par la défense. Il a eu une influence notable sur la première version de l'ASCII.

En 1963, la première version publiée de l'ASCII apparaît.

Auparavant, d'autres standards étaient utilisés notamment sur les cartes perforées^[5].

Sa dernière version stabilisée a été normalisée par l'ANSI en 1986 sous la désignation **ANSI X3.4 :1986** (après deux autres versions en 1967 et 1968, historiquement normalisées par l'ASI, devenu ANSI mais qui ne normalisait pas encore toutes les positions). C'est également la variante américaine des jeux de caractères codés selon la norme **ISO/CEI 646** avec laquelle on la confond souvent (d'où sa désignation également comme **US-ASCII** pour lever l'ambiguïté, désignation préférée dans le registre IANA des jeux de caractères codés).

À l'époque elle a été en concurrence avec des standards incompatibles. Par la suite, l'existence de nombreux codages reprenant les conventions de l'ASCII l'a rendu très populaire.

6.7.2 Principes

L'ASCII définit seulement 128 caractères numérotés de 0 à 127 et codés en binaire de 0000000 à 1111111. Sept bits suffisent donc pour représenter un caractère codé en ASCII. Toutefois, les ordinateurs travaillant presque tous sur un multiple de huit bits (multiple d'un octet) depuis les années 1970, chaque caractère d'un texte en ASCII est souvent stocké dans un octet dont le 8^e bit est 0. Aujourd'hui encore, certains systèmes de messagerie électronique et de SMS fonctionnent avec des bytes ou multiplét composés de seulement sept bits (contrairement à un octet qui est un byte ou multiplét standardisé à huit bits).

Les caractères de numéro 0 à 31 et le 127 ne sont pas affichables ; ils correspondent à des commandes de contrôle de terminal informatique. Le caractère numéro 127 est la commande pour effacer. Le caractère numéro 32 est l'espace. Les autres caractères sont les chiffres arabes, les lettres latines majuscules et minuscules sans accent et quelques symboles de ponctuation.

L'absence des caractères des langues étrangères à l'anglais rend ce standard insuffisant à lui seul pour des textes étrangers (par exemple en langue française), ce qui rend nécessaire l'utilisation d'autres encodages.

Lorsqu'il est employé seul pour la langue anglaise, il interdit l'usage des accents dans la langue anglaise (cf wikt:en:Appendix:English words with diacritics).

Limitations

Quelques-uns des caractères graphiques ASCII ont provoqué une polysémie. Ceci est en tout ou partie lié au nombre limité de codets dans un jeu à sept bits. Ceci se retrouve notamment dans les symboles de ponctuation et l'utilisation des guillemets. L'ASCII a été conservé parce qu'il est omniprésent dans de nombreux logiciels. Cet héritage se retrouve dans **Unicode** où ces signes sont dans un bloc disjoint des autres symboles similaires, se trouvant pour la plupart codés à partir de U+2000^[6].

6.7.3 Internationalisation

Les limites du standard américain ASCII ont conduit, sur trois périodes différentes, à trois approches de l'internationalisation :

- L'utilisation de standards régionaux à caractères-monooctets, techniquement les plus faciles à mettre en place ;
- L'utilisation de standards extensibles, où un même octet peut représenter un caractère différent suivant le contexte (famille **ISO 2022**) ainsi que des extensions où un caractère est codé sur plusieurs octets ;
- L'utilisation du Standard Unicode (famille **UTF**), qui est celui qui comprend le plus grand nombre de caractères.

Les standards régionaux ont l'inconvénient de ne permettre la représentation que d'un ensemble réduit de caractères, comme les caractères d'Europe occidentale. Avec cette approche, il est nécessaire d'indiquer l'encodage à l'extérieur du flot.

Les standards **ISO 2022** ont l'inconvénient d'être contextuels. Il se peut que des logiciels utilisant certains **algorithmes de recherche** manquent d'interopérabilité à cet égard.

Les formes de codage définies par le standard Unicode ont l'inconvénient de la présence éventuelle de la fonctionnalité d'indicateur d'encodage en début de flot, qui le cas échéant est introduit par le caractère **Byte Order Mark**. Certains logiciels anciens ne sont pas compatibles avec la présence de ces trois octets, et ne pourront pas l'être en raison de la complexité conceptuelle que représente le fait de traiter ces trois octets. En particulier, certaines opérations deviennent plus complexes comme la concaténation de chaînes.

6.7.4 Standardisation

Le jeu de codage ASCII est défini quasiment identiquement par plusieurs standards différents, a de nombreuses variantes et a donné naissance à une foison (des dizaines ou des centaines) d'extensions plus ou moins incompatibles entre elles.

Les principales extensions sont justifiées par le fait que l'ASCII ne répond pas aux divers besoins régionaux. Elles sont proposées par des organismes de normalisation, ou par des fournisseurs de produits et de services.

Les standards ASCII

Note : ne pas confondre USASI X3.4-1968 ou ANSI X3.4-1968 et ANSI X3.4 :1986

Normes définissant ou ayant défini ce système de codage des caractères :

- Les standards ASCII des États-Unis (les standards hérités, et le standard en vigueur)
 - *ASA X3.4-1963, (incomplet avec 28 positions libres, et un code de commande non assigné)*
 - *USASI X3.4-1967, (renommé rétroactivement ANSI X3.4-1967) ; ne normalisait pas encore toutes les positions.*
 - *USASI X3.4-1968, (renommé rétroactivement ANSI X3.4-1968) ; ne normalisait pas encore toutes les positions.*
 - *ANSI X3.4-1977,*
 - *ANSI X3.4 :1986 (en 1986, et en vigueur aujourd'hui)*

Les standards internationaux suivants sont généralement considérés compatibles (quasi-identiques) avec le standard ASCII en vigueur de 1986 à 2011, tout en constituant une normalisation internationale officielle :

- Norme ISO/CEI 646
 - ISO/CEI 646-US Variante des États-Unis
 - Variante IRV internationale
- Code page IBM 367.
- Alphabet International de Référence
 - Alphabet International de Référence n° 5 (de 1988)^[7].
 - Alphabet International de référence n° 5 (dans le jeu G0 de l'IRV)^[8].

La désignation US-ASCII ou ASCII é-u, ASCII des États-Unis est un mix des désignations précédentes. Le registre IANA lui attribue la dénomination US-ASCII, sans en définir le codage.

6.7.5 Approximation, variantes et extensions

- Norme ISO/CEI 646
 - Variante INV invariable (incomplète par rapport aux deux précédentes)

Trois types de codages de caractères se rapprochent de l'ASCII :

- ceux qui ne changent que par la dénomination ; ils sont essentiellement identiques à l'ASCII
- ceux qui sont des variantes, l'ASCII étant à l'origine la variante locale aux États-Unis de l'ISO 646
- ceux qui l'augmentent, dits extensions

Alias

En juin 1992, le RFC^[9] et la chambre d'enregistrement de jeux de caractères Internet Assigned Numbers Authority^[10] ont reconnu les alias insensibles à la casse suivants convenables pour l'utilisation dans des protocoles internet :

- ANSI_X3.4-1968 (canonical name)
- iso-ir-6
- ANSI_X3.4-1986
- ISO_646.irv:1991
- ASCII (with ASCII-7 and ASCII-8 variants)
- ISO646-US
- US-ASCII (preferred MIME name)^[10]
- us
- IBM367
- cp367
- csASCII

L'IANA promeut plus particulièrement la dénomination « US-ASCII » pour Internet.

Variantes

Article détaillé : [ISO/CEI 646](#).

ASCII a donné naissance à certaines variantes de l'ASCII qui conservent la plupart des caractères mais en remplacent certains. Dès lors, il ne s'agit plus d'ASCII à strictement parler.

Huitième bit et augmentations

Article détaillé : [ASCII étendu](#).

De nombreuses normes de codage de caractères ont repris les codes ASCII et ajouté d'autres caractères pour les codes supérieurs à 127.

Extensions mono-octets En particulier, beaucoup de pages de codes étendent l'ASCII en utilisant le 8^e bit pour définir des caractères numérotés de 128 à 255. La norme [ISO/CEI 8859](#) fournit des extensions pour diverses langues. Par exemple, l'[ISO 8859-1](#), aussi appelée *Latin-1*, étend l'ASCII avec les caractères accentués utiles aux langues originaires d'Europe occidentale comme le français ou l'allemand.

Par abus de langage, on appelle souvent « ASCII » des normes qui étendent ASCII, mais qui ne sont pas compatibles entre elles (et parfois même ne sont pas compatibles sur les 128 premiers caractères codés). En particulier, les standards [Windows-1252](#) (couramment utilisé sur Microsoft Windows dans les pays occidentaux), [ISO 8859-1](#) (couramment utilisé sur Internet et UNIX) et les pages de code pour PC numéro 437 et 850 (couramment utilisées sur DOS) ne sont pas la norme ASCII. Cet abus de langage ne va pas sans causer des confusions causant des incompatibilités, souvent rendues visibles par le fait que les caractères non ASCII comme les « lettres accentuées » (éÈç) s'affichent mal. On écrit parfois **ASCII de base** pour bien identifier ASCII, et pas un standard plus étendu.

Extensions asiatiques, à base de séquences d'échappement Afin d'unifier les différents codages de caractères complétant l'ASCII et y intégrer les codages complètement différents (le [JIS](#) pour le japonais par exemple, qui bien que développé aussi sur la base de l'US-ASCII, en diffère dans l'assignation d'un des 128 premiers codets), la norme [ISO/CEI 10646](#) a été inventée (et aussi développée au départ séparément par le Consortium Unicode dans une version de sa norme Unicode 1.0 initialement incompatible avec [ISO/CEI 10646](#)).

Voir notamment [ISO 2022](#).

Extensions Unicode La version 1.0 a été abandonnée depuis la version 1.1 afin d'unifier et fusionner les deux répertoires dans un jeu universel de caractères codés). [ISO/CEI 10646](#) codifie des dizaines de milliers de caractères, mais les 128 premiers restent compatibles avec ASCII (dans sa dernière version [X3.4-1986](#)) ; la norme Unicode y ajoute des sémantiques supplémentaires.

Toutefois, certains pays d'Asie orientale (la République populaire de Chine, les anciens dominions britannique et portugais en Chine, de Hong Kong et Macao, qui sont devenus depuis des régions administratives spéciales de Chine, la République de Chine à Taïwan, et le Japon) ont choisi de continuer à développer leur propre norme pour coder le jeu de caractères universel, tout en choisissant de les maintenir entièrement convertibles avec l'[ISO/CEI 10646](#) ; parmi ces normes asiatiques, seule la norme nationale japonaise continue à maintenir une différence dans ses 128 premières positions avec le jeu ASCII, en codant le symbole monétaire du yen à la place de la barre oblique inversée (comme c'est aussi le cas dans la variante japonaise de la norme [ISO/CEI 646](#)).

Parmi les nombreuses extensions 8 bits de l'ASCII, le [Multinational Character Set](#) créé par Digital Equipment Corporation pour le terminal informatique [VT220](#) est considéré comme à la fois l'ancêtre de l'[ISO 8859-1](#) et de l'[Unicode](#)^[11].

6.7.6 Influence

L'ASCII a eu une influence importante dans le monde informatique. En particulier il a pendant longtemps limité les caractères disponibles aux caractères latins non accentués, notamment dans le monde de l'internet, que ce soit pour les noms de domaine, les adresses de courrier électronique, les caractères disponibles dans le Bios, ou les caractères dans lesquels peuvent être écrits des programmes informatiques.

6.7.7 Description

Table des 128 caractères ASCII

ASCII TABLE

Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char	Decimal	Hex	Char
0	0	[NULL]	32	20	[SPACE]	64	40	@	96	60	`
1	1	[START OF HEADING]	33	21	!	65	41	A	97	61	a
2	2	[START OF TEXT]	34	22	"	66	42	B	98	62	b
3	3	[END OF TEXT]	35	23	#	67	43	C	99	63	c
4	4	[END OF TRANSMISSION]	36	24	\$	68	44	D	100	64	d
5	5	[ENQUIRY]	37	25	%	69	45	E	101	65	e
6	6	[ACKNOWLEDGE]	38	26	&	70	46	F	102	66	f
7	7	[BELL]	39	27	'	71	47	G	103	67	g
8	8	[BACKSPACE]	40	28	(72	48	H	104	68	h
9	9	[HORIZONTAL TAB]	41	29)	73	49	I	105	69	i
10	A	[LINE FEED]	42	2A	*	74	4A	J	106	6A	j
11	B	[VERTICAL TAB]	43	2B	+	75	4B	K	107	6B	k
12	C	[FORM FEED]	44	2C	,	76	4C	L	108	6C	l
13	D	[CARRIAGE RETURN]	45	2D	-	77	4D	M	109	6D	m
14	E	[SHIFT OUT]	46	2E	.	78	4E	N	110	6E	n
15	F	[SHIFT IN]	47	2F	/	79	4F	O	111	6F	o
16	10	[DATA LINK ESCAPE]	48	30	0	80	50	P	112	70	p
17	11	[DEVICE CONTROL 1]	49	31	1	81	51	Q	113	71	q
18	12	[DEVICE CONTROL 2]	50	32	2	82	52	R	114	72	r
19	13	[DEVICE CONTROL 3]	51	33	3	83	53	S	115	73	s
20	14	[DEVICE CONTROL 4]	52	34	4	84	54	T	116	74	t
21	15	[NEGATIVE ACKNOWLEDGE]	53	35	5	85	55	U	117	75	u
22	16	[SYNCHRONOUS IDLE]	54	36	6	86	56	V	118	76	v
23	17	[ENG OF TRANS. BLOCK]	55	37	7	87	57	W	119	77	w
24	18	[CANCEL]	56	38	8	88	58	X	120	78	x
25	19	[END OF MEDIUM]	57	39	9	89	59	Y	121	79	y
26	1A	[SUBSTITUTE]	58	3A	:	90	5A	Z	122	7A	z
27	1B	[ESCAPE]	59	3B	;	91	5B	[123	7B	{
28	1C	[FILE SEPARATOR]	60	3C	<	92	5C	\	124	7C	
29	1D	[GROUP SEPARATOR]	61	3D	=	93	5D]	125	7D	}
30	1E	[RECORD SEPARATOR]	62	3E	>	94	5E	^	126	7E	~
31	1F	[UNIT SEPARATOR]	63	3F	?	95	5F	_	127	7F	[DEL]

Table de correspondance ASCII

Dans la table suivante, les 33 caractères de contrôle (codes 0 à 31 et 127) sont présentés avec leur nom en anglais suivi d'une traduction entre parenthèses.

On peut aussi présenter la table des caractères ASCII sous cette forme plus condensée qui met en évidence une organisation fondée sur la base 16.

Caractères de contrôle

NUL Originellement une **NOP**, un caractère à ignorer. Lui donner le code 0 permettait de prévoir des réserves sur les bandes perforées en laissant des zones sans perforation pour insérer de nouveaux caractères *a posteriori*. Avec le développement du langage C il a pris une importance particulière quand il a été utilisé comme indicateur de fin de chaîne de caractères.

SOH *Start of heading* : début d'en-tête. Il est aujourd'hui souvent utilisé dans les communications séries pour permettre la synchronisation après erreur^[14].

DEL *Delete* : effacement. Lui donner le code 127 (1111111 en binaire) permettait de supprimer *a posteriori* un caractère sur les bandes perforées qui codaient les informations sur 7 bits. N'importe quel caractère pouvait être transformé en DEL en complétant la perforation des 7 bits qui le composaient.

LF, CR, fin de ligne Dans un fichier texte, la fin d'une ligne est représentée par un ou deux caractères de contrôle. Plusieurs conventions existent :

- sur les systèmes Multics, Unix, Type Unix (Linux, AIX, Xenix, Mac OS X, etc.), BeOS, AmigaOS, RISC OS entre autres, la fin de ligne est indiquée par un saut de ligne (LF) ;

- sur les machines **Apple II** et **Mac OS** jusqu'à la version 9, la fin de ligne est indiquée par un retour chariot (CR) ;
- sur les systèmes **DEC**, **RT-11** et généralement tous les premiers systèmes non-Unix et non-IBM, **CP/M**, **MP/M**, **MS-DOS**, **OS/2** ou **Microsoft Windows**, la fin de ligne est indiquée par un retour chariot suivi d'un saut de ligne (CR suivi de LF).

Ainsi, lorsqu'on transfère un fichier ASCII entre des systèmes ayant des conventions de fin de ligne différentes, il faut convertir les fins de ligne pour pouvoir le manipuler confortablement sur le système cible. Autrement, il faut utiliser un éditeur de texte capable de gérer les diverses conventions de fin de ligne, ce qui n'est par exemple pas le cas du classique **Bloc-notes** de **Microsoft Windows**. Les programmes utilisant les fichiers ASCII ne sont en général pas perturbés par un changement de type de fin de ligne.

SUB Il est souvent associé à la combinaison de touche Contrôle + z, et est utilisé dans les communications séries pour permettre l'envoi des données en lieu et place de la touche entrée.

6.7.8 Voir aussi

Notes et références

- [1] <http://www.trailing-edge.com/~{ }bobbemer/SURVEY.HTM>
- [2] <http://discover.lib.umn.edu/cgi/f/findaid/findaid-idx?c=umfa;cc=umfa;rgn=main;view=text;didno=cbi00067>
- [3] <http://www.cs.tut.fi/~{ }jkorpe/latin1/ascii-hist.html>
- [4] Appendix §A.6.8 de la norme de A963
- [5] Pascal Becker, Histoire des supports informatiques (lire en ligne)
- [6] (fr) *Unicode 5.0 en pratique*, chapitre 7 « Ponctuation », Patrick Andries.
- [7] Alphabet International de référence n° 5 : RECOMMANDATION T50, Union international des télécommunications, 1988 (lire en ligne)
- [8] Alphabet International de référence : ANCIEN ALPHABET INTERNATIONAL n° 5 ou AI5, RECOMMANDATION T50, Union international des télécommunications, 1992 (lire en ligne)
- [9] (en) K. Simonsen, Rationel Almen Planlaegning, Request for Comments : 1345 : Character Mnemonics & Character Sets, Network Working Group, juin 1992 (RFC 1345).
- [10] Internet Assigned Numbers Authority (May 14, 2007). "Character Sets". Accessed 2008-04-14.
- [11] Roman Czyborra, « ISO 8859-1 and MCS, from ISO 8859 Alphabet Soup », sur *czyborra.com*
- [12] La norme ANSI X3.4 définit le caractère 39 par *apostrophe (closing single quotation mark, acute accent)* et les anciennes tables de caractères le représentaient souvent incliné. Les encodages plus récents restreignent ce code à la représentation de l'apostrophe verticale (ni penchée à droite, ni à gauche, mais neutre). Voir (en) Latin-1's apostrophe, grave accent, acute accent.
- [13] Le code 96 est également employé comme *guillemet ouvrant simple* en ASCII. En Unicode, il existe un code plus approprié.
- [14] (en) *ASCII character set*

Bibliographie

- (en) Business Equipment Manufacturer Associations, American Standard Code for Information Interchange : ASA standard X3.4-1963, American Standards Association Incorporated, 17 juin 1963 (lire en ligne)
- (en) Eric Fischer, The Evolution of Character Codes, 1874-1968, .transbay.net (lire en ligne)

Articles connexes

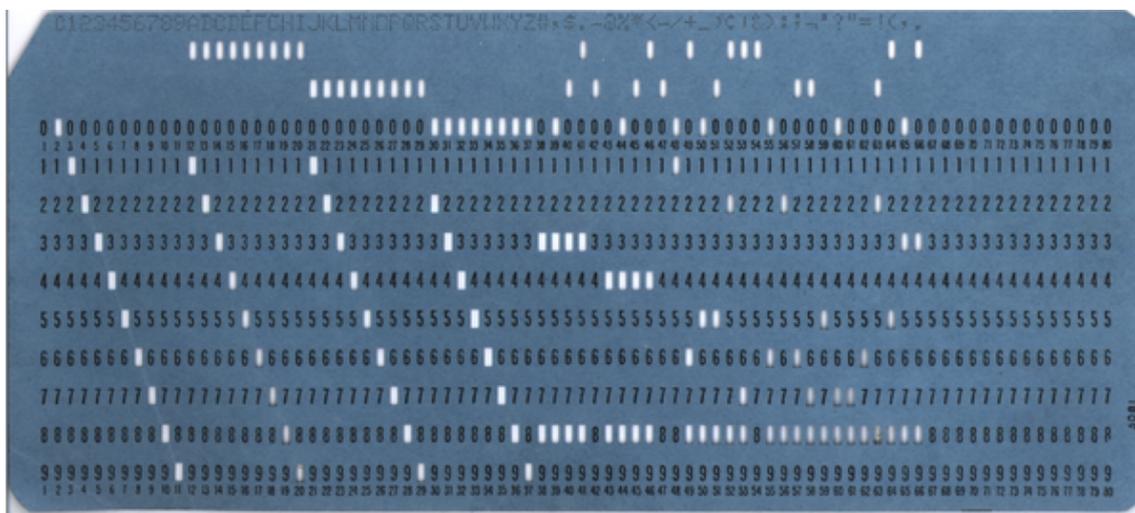
- Art ASCII
- ASCII porn
- Fichier texte
- Vidéotex
- libcaya : bibliothèque permettant des rendus ASCII à partir de vidéos
- Unicode
-  Portail de l'informatique
-  Portail de l'écriture

6.8 Extended Binary Coded Decimal Interchange Code

L'Extended Binary Coded Decimal Interchange Code (EBCDIC) est un mode de codage des caractères sur 8 bits créé par IBM à l'époque des *cartes perforées*. Il existe au moins 6 versions différentes bien documentées (et de nombreuses variantes parfois créées par des concurrents d'IBM), incompatibles entre elles. Ce mode de codage a été critiqué pour cette raison, mais aussi parce que certains caractères de ponctuation ne sont pas disponibles dans certaines versions. Ces disparités ont parfois été interprétées comme un moyen pour IBM de conserver ses clients captifs.

EBCDIC est encore utilisé dans les systèmes AS/400 d'IBM ainsi que sur les mainframes sous MVS, VM ou DOS/VSE.

6.8.1 Historique



Une carte perforée 80 colonnes d'IBM d'un des types les plus utilisés au XX^e siècle. La photographie fait apparaître le jeu de caractère EBCDIC de 1964, qui ajoutait de nouveaux caractères aux caractères précédents.

La disposition apparemment « étrange » du codage EBCDIC est historique et provient des anciens lecteurs de bandes et cartes perforées, qui devaient pouvoir être lues et percées manuellement par un humain en utilisant le système décimal plus mnémotechnique (et il était lui-même dérivé de l'ancien jeu de caractère télégraphique sur 5 bits).

Les bits de poids fort sont apparus après et ont permis de coder dans une colonne supplémentaire de perforation les distinctions entre chiffres et lettres, ou entre minuscules et majuscules. Le jeu était conçu pour qu'à chaque caractère corresponde aussi au moins un trou, afin de permettre la synchronisation des lecteurs de bandes perforées,

c'est pourquoi la colonne 0 était à l'origine inutilisée pour les lettres, les chiffres ayant un trou dédié supplémentaire correspondant à un bit de poids fort.

Le jeu EBCDIC à l'origine était bien un jeu sur 7 bits, le huitième étant alors uniquement utilisé sur les cartes perforées pour indiquer que la colonne de perforation était bien occupée par un caractère codé (c'est pourquoi les principaux codes EBCDIC occupent la seconde moitié de la table).

De même, les ponctuations étaient codées aussi de façon visuelle et mnémotechnique, et pouvaient facilement être distinguées des chiffres et lettres par le fait qu'elles n'avaient aucun trou dans les positions binaires correspondant aux marques de chiffres ou de lettres.

L'absence de toute perforation dans une colonne de perforations était interprétée comme du bourrage non signifiant, lié à la technologie employée (ce code en début de table correspond aujourd'hui au caractère de commande C0 "NUL"). De même, on pouvait corriger une erreur manuelle de perforation en perforant toutes les positions d'une colonne, et donc la dernière position est aussi un caractère de bourrage non signifiant (ce code en fin de table correspond aujourd'hui au caractère de commande C1 "APC", fonction plutôt dévolue au caractère de commande C0 "DEL" de l'ISO 646 et de l'ASCII mais dont l'usage est plus ambigu).

L'EBCDIC est donc le fruit d'une évolution historique bien plus ancienne (développée de longue date par IBM en continuité avec les anciens systèmes télégraphiques) que l'ASCII (plus pratique à manipuler dans les programmes) qui l'a remplacé ensuite pratiquement partout et a donné ensuite lieu à une normalisation dans l'ISO 646.

Les bandes et cartes perforées ont cependant continué à être utilisées avec l'EBCDIC jusque vers les années 1990, par exemple en **Turquie** pour les échanges de données militaires alors que les problèmes de compatibilité des systèmes d'enregistrement magnétiques ou des réseaux étaient encore loin d'être tous résolus.

Le succès des normes de communication entre systèmes hétérogènes (notamment **TCP/IP** et **l'Internet**, ainsi que les systèmes de **cryptographie** pour la transmission sécurisée et moins coûteuse via des réseaux publics) a mis fin à l'intérêt de l'EBCDIC pour tout nouveau développement, et les nombreuses versions nationales de l'EBCDIC ont également de moins en moins d'intérêt depuis l'apparition de l'**UTF-EBCDIC**.

6.8.2 Exemple de codage

Ce tableau représente le codage d'une variante d'EBCDIC compatible avec l'ISO 8859-1. Les caractères codés de 0x00 à 0x3F ainsi que 0xFF sont des caractères de contrôle, 0x40 est l'espace, 0x41 est l'espace insécable. Le caractère codé en 0x54 est le tiret de **césure** visible uniquement en fin de ligne.

Cette variante est compatible, octet par octet, avec l'**UTF-EBCDIC** qui utilise les positions variantes colorées en vert pour coder les caractères Unicode (hors de l'ASCII et des codes de contrôle) sous forme de séquences d'octets.

Chaque variante nationale ou internationale de l'EBCDIC est codée d'après cette table (cependant des permutations sont possibles entre deux CCSID pour les mêmes caractères). Toutes les variantes de l'EBCDIC ne codent pas les lettres minuscules latines de base, au contraire des jeux de caractères basés sur l'ISO 646 ou ISO 8859.

D'autre part certaines variantes nationales de l'ISO 646 (ou autres jeux de caractères compatibles avec cette norme) contiennent des positions variantes supplémentaires, invariantes dans les jeux de caractères basés sur l'EBCDIC.

Enfin, les positions 0x5A et 0x7F de l'EBCDIC sont variantes — la plupart des variantes EBCDIC y codent respectivement le point d'exclamation et le guillemet anglais (double quote) — au contraire des caractères correspondants des variantes nationales de l'ISO 646 aux positions respectives 0x21 et 0x22.

Sur les systèmes EBCDIC, le saut de ligne est normalement codé avec le caractère de commande C1 "NEL" (U+0085 en Unicode, ou 0x25 dans toutes les variantes standard EBCDIC) et non avec les caractères de commande C0 "CR" et/ou "LF" de l'ISO 646 et de l'ASCII (U+000D et/ou U+000A, c'est-à-dire 0x0D et/ou 0x15 en EBCDIC, où ces commandes ont une fonction bien définie et unique de gestion de position du curseur sur un terminal, ou bien permettent de distinguer les sauts de lignes forcés dans un même paragraphe, ou encore permettaient d'afficher une ligne en surimpression pour produire des caractères gras, soulignés, ou accentués supplémentaires).

6.8.3 Transcodage de l'ISO 8859-1 vers l'EBCDIC

La table suivante permet de transcoder l'ASCII (caractères Unicode U+0000 à U+007F) et le jeu de commandes C1 (caractères Unicode U+0080 à U+009F) en EBCDIC. Les caractères de l'extensions ISO 8859 (en verts) sont indiqués ici dans l'ordre compatible avec l'**UTF-EBCDIC**, mais cette assignation ne correspond pas à une variante

nationale particulière de l'EBCDIC.

Cette table est l'inverse de la table précédente et est compatible octet par octet avec la seconde phase (de permutation des valeurs d'octets) de l'UTF-EBCDIC.

Ensembles, ces deux tables permettent d'adapter facilement n'importe quel jeu de caractère compatible ISO 646 ou ISO 8859 pour des traitements sur des systèmes EBCDIC standards. Par contre, pour adapter à une version précise d'un jeu EBCDIC (et faire reconnaître exactement les caractères non ASCII), il faudra substituer les positions marquées en vert dans les deux tables, en fonction de leur assignation réelle dans les variantes nationales de jeux EBCDIC correspondants, voire substituer les lettres minuscules latines remplacées dans certaines versions obsolètes de certains jeux EBCDIC (jeu EBCDIC japonais ou cyrillique, sauf les jeux EBCDIC russes qui unifient avec le même code EBCDIC certaines lettres cyrilliques et latines comme le A, d'apparences identiques dans les deux écritures).

Le détail des assignations de ces positions (marquées ici en vert dans les deux tables) dans les variantes nationales de l'EBCDIC est référencé sur le site IBM mentionné dans les liens externes. Notamment, la variante encore la plus utilisée est l'EBCDIC CCSID 500 (version internationale du jeu Latin n°1 dont le jeu est très proche de celui de l'ISO 8859-1, mais avec quelques différences, et triée différemment des tables présentées ici).

6.8.4 Voir aussi

Article connexe

- EBCDIC-codepages avec Latin-1-charset
- EBCDIC-codepage 297 (France)
- UTF-EBCDIC

Liens externes

- Tables EBCDIC sur le site IBM
-  Portail de l'informatique

tsm

6.9 UTF

La norme **Unicode** définit plus de cent mille caractères. Chaque caractère abstrait est identifié par un nom unique (un en anglais et un en français) et associé à un nombre entier positif appelé son point de code (ou position de code).

UTF (pour **Unicode Transformation Format** en anglais) désigne les méthodes de conversion possibles pour le transport informatique de ces nombres :

- UTF-1 (en), une méthode de conversion annulée et remplacée par UTF-8, son objectif était de maximiser la compatibilité avec ISO 2022 ;
- UTF-7, une méthode de conversion considérée comme obsolète, n'a jamais fait partie de la norme **Unicode**, initialement décrite dans le RFC 1642 ;
- UTF-8, conversion par une suite de codes de 8 bits ;
- UTF-EBCDIC, conversion similaire à UTF-8 mais conçu pour être compatible avec l'EBCDIC ;
- UTF-16, conversion par une suite de codes de 16 bits ;
- UTF-32, conversion par une suite de codes de 32 bits.

6.10 UTF-8

UTF-8 (abréviation de l'anglais *Universal Character Set Transformation Format - 8 bits*) est un codage de caractères informatiques conçu pour coder l'ensemble des caractères du « répertoire universel de caractères codés », initialement développé par l'ISO dans la norme internationale ISO/CEI 10646, aujourd'hui totalement compatible avec le standard Unicode, en restant compatible avec la norme ASCII limitée à l'anglais de base (et quelques autres langues beaucoup moins fréquentes), mais très largement répandue depuis des décennies.

L'UTF-8 est utilisé par 82,2 % des sites web en décembre 2014^[1]. De par sa nature, UTF-8 est d'un usage de plus en plus courant sur Internet, et dans les systèmes devant échanger de l'information. Il s'agit également du codage le plus utilisé dans les systèmes GNU, Linux et compatibles pour gérer le plus simplement possible des textes et leurs traductions dans tous les systèmes d'écritures et tous les alphabets du monde.

6.10.1 Liens avec la norme internationale ISO/CEI 10646 et les standards Unicode et de l'internet

UTF-8 est un « format de transformation » issu à l'origine des travaux pour la norme ISO/CEI 10646, c'est-à-dire que UTF-8 définit un codage pour tout point de code scalaire (caractère abstrait ou « non-caractère ») du répertoire du jeu universel de caractères codés (*Universal Character Set*, ou *UCS*). Ce répertoire est aujourd'hui commun à la norme ISO/CEI 10646 (depuis sa révision 1) et au standard Unicode (depuis sa version 1.1).

UTF-8 est officiellement défini dans la norme ISO/CEI 10646 depuis son adoption dans un amendement publié en 1996. Il fut aussi décrit dans le standard Unicode et fait partie de ce standard depuis la version 3.0 publiée en 2000. En 1996 fut publiée la RFC 2044 (« *UTF-8, a transformation format of ISO 10646* ») dans le but de fournir une spécification accessible d'UTF-8 et d'entamer sa standardisation au sein de l'*Internet Engineering Task Force* (IETF). Cette RFC fut révisée en 1998 (RFC 2279) puis finalement en 2003 (RFC 3629), cette dernière version faisant d'UTF-8 un des standards de l'internet (STD 63).

6.10.2 Description technique

Techniquement, il s'agit de coder les caractères Unicode sous forme de séquences de un à quatre codets d'un octet chacun. La norme Unicode définit entre autres un ensemble (ou répertoire) de caractères. Chaque caractère est repéré dans cet ensemble par un index entier aussi appelé « point de code ». Par exemple le caractère « € » (euro) est le 8365^e caractère du répertoire Unicode, son index, ou point de code, est donc 8364 (on commence à compter à partir de 0).

Le répertoire Unicode peut contenir plus d'un million de caractères, ce qui est bien trop grand pour être codé par un seul octet (limité à des valeurs entre 0 et 255). La norme Unicode définit donc des méthodes standardisées pour coder et stocker cet index sous forme de séquence d'octets : UTF-8 est l'une d'entre elles, avec UTF-16, UTF-32 et leurs différentes variantes.

La principale caractéristique d'UTF-8 est qu'elle est rétro-compatible avec la norme ASCII, c'est-à-dire que tout caractère ASCII se code en UTF-8 sous forme d'un unique octet, identique au code ASCII. Par exemple « A » (A majuscule) a pour code ASCII 65 et se code en UTF-8 par l'octet 65. Chaque caractère dont le point de code est supérieur à 127 (caractère non ASCII) se code sur 2 à 4 octets. Le caractère « € » (euro) se code par exemple sur 3 octets : 226, 130, et 172.

6.10.3 Description

Le numéro (valeur scalaire) de chaque point de code dans le jeu universel de caractères (UCS) est donné par la norme ISO/CEI 10646 qui assigne un point de code à chaque caractère valide, puis permet leur codage en leur attribuant une valeur scalaire identique au point de code ; cette norme est reprise dans le standard Unicode (qui utilise depuis la version 1.1 le même répertoire).

Tous les « points de code » (*code points* en anglais) de U+0000 à U+D7FF et de U+E000 à U+10FFFF sont représentables en UTF-8 — même ceux qui sont attribués à des « non-caractères » (*non-character*) et tous ceux qui ne sont pas encore attribués — et uniquement ceux-là. Les seuls points de codes valides dans l'espace de l'UCS et qui ne doivent pas être représentés dans UTF-8 sont ceux qui sont attribués aux « demi-codets » (*surrogates* en anglais), car

ils ne sont pas représentables de façon bijective dans le codage UTF-16 et ne sont pas non plus par eux-mêmes des caractères : contrairement aux autres points de codes, les demi-codets n'ont donc *pas* de « valeur scalaire » (*scalar value* en anglais) définie.

Les points de code ayant une valeur scalaire de 0 à 127 (points de codes U+0000 à U+007F, attribués aux caractères du jeu codé sur 7 bits dans l'ASCII) sont codés sur un seul octet dont le bit de poids fort est nul.

Les autres points de code (attribués ou non à des caractères) ayant une valeur scalaire supérieure à 127 (sauf ceux auxquels sont attribués des « demi-codets » qui ne sont pas eux-mêmes des caractères) sont codés sur plusieurs octets ayant chacun leur bit de poids fort non nul : les bits de poids fort du premier octet de la séquence codée forment une suite de 1 de longueur égale au nombre total d'octets (au moins 2) utilisés pour la séquence entière suivie d'un 0 et les octets suivants nécessaires ont leurs deux bits de poids fort positionnés à 10.

Ce principe pourrait être étendu jusqu'à huit octets pour un seul point de code (pour représenter des points de code comprenant jusqu'à 42 bits), mais la version normalisée actuelle d'UTF-8 pose la limite à quatre^[2].

Le codage interdit la représentation des points de code réservés aux *demi-codets* (qui n'ont pas de valeur scalaire définie, afin de préserver la compatibilité avec UTF-16 qui ne permet pas non plus de les représenter). Il autorise cependant la représentation des points de code assignés à des *non-caractères* (alors même que leur présence est interdite dans un texte conforme).

Exemples

Caractéristiques

Dans toute chaîne de caractères codée en UTF-8, on remarque que :

- tout octet de bit de poids fort nul désigne un unique « point de code » assigné à un caractère du répertoire de l'US-ASCII et codé sur ce seul octet, d'une valeur scalaire identique à celle du codet utilisé dans le codage US-ASCII ;
- tout octet de bits de poids fort valant 11 est le premier octet d'une séquence unique représentant un « point de code » (assigné à un caractère ou un non-caractère) et codé sur plusieurs octets ;
- tout octet de bits de poids fort valant 10 est un des octets suivants d'une séquence unique représentant un « point de code » (assigné à un caractère ou un non-caractère) et codé sur plusieurs octets ;
- aucun octet ne peut prendre une valeur hexadécimale entre C0 et C1, ni entre F5 et FF (le plus haut point de code valide et assigné à un caractère représentable est U+10FFFF ; c'est un caractère à usage privé alloué dans le 17^e plan valide).

Le plus grand point de code valide assignable à un caractère valide *non privé* est U+EFFFD dans le 15^e plan (il n'est pas encore assigné mais peut le devenir dans l'avenir), mais le codage UTF-8 peut être utilisé aussi, de façon conforme aux normes, pour représenter n'importe quel caractère valide à usage privé (dans une des trois plages U+E000 à U+F8FF, U+F000 à U+FFFF, et U+10000 à U+10FFFF).

L'acceptation ou non des *non-caractères* ou des *caractères d'usage privé* est laissée aux applications ou protocoles de transport de texte. Cependant les *non-caractères* ne sont normalement pas acceptés dans des textes strictement conformes au standard Unicode où à la norme ISO/CEI 10646.

Certaines applications imposent des restrictions supplémentaires sur les points de code utilisables (par exemple, les standards HTML et XML interdisent, dans tout document conforme à ces spécifications, la présence de la plupart des caractères de contrôle entre U+0000 et U+001F et entre U+0080 et U+009F, en dehors du contrôle de la tabulation U+0009 considéré comme un caractère blanc, et interdisent aussi les *non-caractères*).

Tout point de code est toujours représenté par exactement la même séquence binaire, quelle que soit sa position relative dans le texte, et ces séquences sont autosynchronisées sur la position indivise des codets significatifs (ici les octets : on peut toujours savoir si un octet débute ou non une séquence binaire effective) ; ce codage autorise donc les algorithmes rapides de recherche de texte, tel que l'algorithme de Boyer-Moore.

Ce n'est pas toujours le cas des codages contextuels (qui utilisent généralement la compression de données, par exemple SCSU défini dans la note technique standard UTS#6 optionnelle complétant le standard Unicode) et qui peuvent nécessiter de lire le texte complètement depuis le début, ni des codages basés sur plus d'une seule variable

d'état (ou qui incorporent des codes supplémentaires de redondance) ; au mieux certains de ces codages peuvent demander d'utiliser des algorithmes complexes de resynchronisation, basés souvent sur des heuristiques qui peuvent échouer ou conduire à de fausses interprétations si on ne lit pas le texte depuis le début (par exemple BOCU-1).

Principe et unicité du codage Dans le tableau ci-dessus on voit que le caractère « € » se trouve au point de code 8364, soit en binaire : 00100000 10101100

$\lceil \log_2 8364 \rceil = 14$ donc 14 bits au moins sont nécessaires pour encoder le caractère « € ». Selon le tableau « Définition du nombre d'octets utilisés », il nous faut 3 octets pour placer les 14 bits dont nous avons besoin.

Mais bien sûr si nous avons 4 octets (21 bits), nous pourrions là aussi placer nos 14 bits. Il est dit ^[3] : « Pour des raisons de sécurité, un programme qui décode des caractères au format UTF-8 ne doit pas accepter les séquences UTF-8 qui sont plus longues que nécessaires pour coder ces caractères. ». (« Il risquerait d'abuser un test de sous-chaîne, qui ne regarderait que les codages les plus courts ».)

Ainsi « € » se codera : **11100010 10000010 10101100**, mais ne se codera pas : **11110000 10000010 10000010 10101100**.

Une telle forme, plus longue que nécessaire s'appelle en anglais *overlong*. De telles formes (initialement autorisées dans des spécifications anciennes avant qu'elles soient normalisées successivement par la RFC initiale publiée par le Consortium X/Open, puis parallèlement par la norme ISO 10646 et le standard Unicode) sont interdites et doivent être traitées comme invalides.

Types d'octets, séquences valides et décodage

Le codage est prédictif et permet toujours de retrouver la position du premier octet d'une séquence représentant un point de code, à partir de la valeur d'un octet quelconque et de la lecture d'un nombre limité d'octets voisins, dans les deux directions de lecture (ce sera toujours l'octet lui-même, ou le premier éligible dans un des 1 à 3 octets voisins).

- Tout octet de continuation dans une séquence UTF-8 valide ne peut prendre que les valeurs hexadécimales **80 à BF** ;
 - il ne peut exister qu'à la suite d'un octet de début de séquence (représentant un point de code), qui sera le dernier codé dans un des 1 à 3 octets précédents et qui n'est pas non plus un octet de continuation ;
 - le point de code suivant, s'il y en a un, ne peut commencer au maximum que dans les 1 à 3 octets suivants.
- Le premier octet d'une séquence UTF-8 valide ne peut prendre que les valeurs hexadécimales **00 à 7F ou C2 à F4** :
 - le premier octet hexadécimal **00 à 7F** d'une séquence n'est suivi d'**aucun** octet de continuation ;
 - le premier octet hexadécimal **C2 à DF** d'une séquence est toujours suivi d'**un seul** octet de continuation (chacun de valeur hexadécimale entre 80 et BF) ;
 - le premier octet hexadécimal **E0 à EF** d'une séquence est toujours suivi de **deux** octets de continuation (chacun de valeur hexadécimale entre 80 et BF) ;
 - cependant, si le premier octet d'une séquence prend la valeur hexadécimale **E0**, le premier octet de continuation est restreint à une valeur hexadécimale entre **A0** et BF ;
 - cependant, si le premier octet d'une séquence prend la valeur hexadécimale **ED**, le premier octet de continuation est restreint à une valeur hexadécimale entre 80 et **9F** ;
 - le premier octet hexadécimal **F0 à F4** d'une séquence est toujours suivi de **trois** octets de continuation (chacun de valeur hexadécimale entre 80 et BF) ;
 - cependant, si le premier octet d'une séquence prend la valeur hexadécimale **F0**, le premier octet de continuation est restreint à une valeur hexadécimale entre **90** et BF ;
 - cependant, si le premier octet d'une séquence prend la valeur hexadécimale **F4**, le premier octet de continuation est restreint à une valeur hexadécimale entre 80 et **8F**.

Séquences interdites

- Les points de code sont toujours représentés par la séquence d'octets la plus courte possible :
 - par conséquent, aucune séquence d'octets ne contient des octets initiaux de valeur hexadécimale **C0** ou **C1** dans un texte valide codé en UTF-8 ;
 - de même, aucune séquence commençant par l'octet initial **E0** ne peut avoir un premier octet de continuation de valeur hexadécimale **80 à 9F**.
- Les points de code allant de U+D800 à U+DFFF sont interdits (leur valeur scalaire est réservée pour la représentation UTF-16 des points de code supplémentaires avec des paires de demi-codets) :
 - par conséquent, le premier octet de continuation d'une séquence qui commence par l'octet hexadécimal **ED** ne peut prendre aucune des valeurs hexadécimales **A0 à BF** ;
 - ces séquences interdites en UTF-8 sont en revanche autorisées dans la transformation CESU-8 (non recommandée, qui ne doit en aucun cas être confondue avec UTF-8, car CESU-8 utilise ces séquences pour coder les caractères des plans supplémentaires en 2 séquences de 3 octets chacune, au lieu d'une seule séquence de 4 octets en UTF-8).
- De même que tout codage pouvant donner un point de code de valeur supérieure à U+10FFFF est interdit :
 - par conséquent, le premier octet de continuation d'une séquence qui commence par l'octet hexadécimal **F4** ne peut prendre aucune des valeurs hexadécimales **90 à BF** ;
 - et aucune séquence d'octets ne contient des octets initiaux de valeur hexadécimale **F5 à FF**.

De telles séquences sont dites mal formées (*ill-formed*). (Voir la référence ci-dessus, notamment la seconde table dans la clause de conformité D36 du standard ou l'article Unicode).

En revanche, les points de code réservés (pas encore alloués à des caractères) sont autorisés (même si l'interprétation des caractères peut rester ambiguë) : il appartient aux applications de décider si ces caractères sont acceptables ou non, sachant que les mêmes applications continueront probablement à être utilisées alors que ces positions auront été assignées dans les normes Unicode et ISO 10646 à de nouveaux caractères parfaitement valides.

De même les autres points de code assignés de façon permanente aux autres « *non-caractères* » sont interdits dans les textes conformes à la norme ISO/CEI 10646 ou au standard Unicode : par exemple U+xFFFF à U+xFFFF (où x indique un numéro de plan hexadécimal de 0 à 10). Mais ils restent encodables et décodables en tant que tels en UTF-8 (les *non-caractères* sont à disposition des applications qui peuvent en faire un usage au sein d'API internes, par exemple comme codes intermédiaires nécessaires à l'implémentation de certains traitements).

La restriction de l'espace de représentation aux seuls points de code inférieurs ou égaux à U+10FFFF (non compris les points de codes assignés aux *demi-codets*) n'a pas toujours été appliquée :

- Cela n'a pas toujours été le cas dans la norme ISO/CEI 10646, qui prévoyait à l'origine de pouvoir coder un très grand nombre de plans possibles (l'UCS-4 permettait un codage jusqu'à 31 bits), alors que Consortium Unicode (depuis la fusion du répertoire commun dans sa version 1.1) n'utilisait encore que le plan multilingue de base et n'avait pas encore envisagé de couvrir autant d'écritures qu'aujourd'hui.
- L'introduction par Unicode du codage UTF-16 dans une annexe standard (quand il a admis que plus de 65536 caractères seraient rapidement nécessaires) a demandé l'allocation préalable par l'ISO/CEI 10646 d'un bloc de points de codes pour des « demi-codets » qui étaient considérés au début par l'ISO/CEI 10646 comme des caractères spéciaux (une concession faite à Unicode alors que l'UCS-4 avait été créé comme un espace de codage linéaire où tous les points de code avaient une valeur scalaire), alors qu'Unicode n'utilisait encore que le sous-espace UCS-2 et pas l'espace UCS-4 complet.
 - Pour éviter des problèmes d'interopérabilité avec les autres applications (non Unicode) basées sur UCS-2, une première révision de l'UTF-8 a été publiée par l'ISO en 1998, mentionnant que ces demi-codets n'avaient donc pas de valeur scalaire définie et qu'aucun point de code assignés aux « demi-codets » dans les deux blocs successifs alloués ne devait pas être encodés en UTF-8.
- Mais selon l'accord final intervenu entre le comité technique du Consortium Unicode et celui chargé de la norme ISO/CEI 10646, toute utilisation de plus de 17 plans a été proscrite, afin d'assurer l'interopérabilité totale avec le codage UTF-16 défini par Unicode, un codage déjà massivement déployé dans les systèmes

d'exploitation (par exemple **Microsoft Windows**), ou sous-systèmes (par exemple **OpenType**), ou encore dans de nombreux langages de programmation qui en dépendent pour leur interopérabilité (dont certains issus de normes nationales ou internationales, tels que les langages C et C++ qui ont eu des difficultés à supporter le répertoire universel).

- En effet, après une vingtaine d'années d'efforts pour la définition de l'UCS pour toutes les écritures du monde, des règles plus strictes ont été établies pour limiter les caractères encodables selon un modèle assurant une compatibilité ascendante mais aussi de meilleures pratiques de codage. Pratiquement toutes les écritures modernes du monde ont été codées, et on dispose d'estimations fiables de l'ordre de grandeur sur la quantité de caractères nécessaires pour le support des autres écritures, ainsi que sur les besoins de codage pour de nouveaux caractères.
- La croissance initiale très forte des allocations dans l'UCS (ou des caractères restant encore à coder) s'est fortement ralentie, et seulement 6 des 17 plans sont utilisés fin 2011 (mais deux seulement ont un taux de remplissage significatif : le plan multilingue de base, pratiquement plein, et le plan idéographique supplémentaire ; c'est sur le plan multilingue complémentaire que se concentrent la majorité des autres écritures anciennes restant à encoder, ou des nouveaux ensembles de symboles et caractères de notation technique).
- Le rythme de croissance des allocations dans l'UCS pour la norme ISO/CEI 10646 ne permet pas d'envisager sa saturation avant un terme dépassant de très loin le cycle de vie des normes internationales (et encore plus celui des standards industriels comme Unicode). À ce terme trop lointain, il est tout à fait possible que UTF-16 soit devenu obsolète depuis fort longtemps, ou qu'une nouvelle norme de codification ait vu le jour et ait été massivement déployée (et que les outils de conversion automatique aient aussi été normalisés et déployés). Rien ne justifie encore de maintenir une extension possible non nécessaire au besoin immédiat d'interopérabilité des normes et standards actuels ou des futures normes envisagées.
- Un ou deux autres plans seulement sont envisagés pour les écritures sinographiques, anciennes écritures cunéiformes ou hiéroglyphiques, et éventuellement un autre plan pour des collections de symboles et pictogrammes nécessaires à l'interopérabilité de certaines applications modernes (par exemple les *emojis* des messageries et services interactifs est-asiatiques, ou des symboles nécessaires à des normes internationales de signalisation ou de sécurité).
- Les « groupes » supplémentaires d'usage privé à la fin de l'UCS-4, ainsi que les « plans » supplémentaires d'usage privé dans l'UCS-4 à la fin du groupe 0, qui avaient été envisagés par l'ISO depuis le début de ses travaux de normalisation, ont été abandonnés pour ne garder, parmi les 17 premiers plans du premier groupe, que les deux derniers plans à cet usage privé (en plus du bloc d'usage privé U+E000 à U+F8FF déjà alloué dans le plan multilingue de base), ce qui s'avère suffisant pour toutes les applications.
- Cela a fait l'objet de la révision en 2003 de la RFC publiée par le comité technique de l'ISO définissant le codage UTF-8 pour la norme ISO/CEI 10646, et simultanément d'une mise à jour de l'annexe standard au standard Unicode (une annexe standard qui a, depuis, été intégrée au standard lui-même).
- Depuis ces mises à jour de 2003, le codage UCS-4 défini par la norme ISO/CEI 10646 est devenu en pratique équivalent à UTF-32 (défini dans la norme Unicode qui adjoint des propriétés supplémentaires mais sans différence de codage). Et la dernière RFC publiée par l'ISO et approuvée par l'IETF en 2003 fait d'ailleurs maintenant une référence normative à la définition de l'UTF-8 publiée conjointement avec (ou dans) le standard Unicode.

6.10.4 Avantages

Universalité Ce codage permet de représenter les milliers de caractères du répertoire universel, commun à la norme ISO/CEI 10646 et au standard Unicode (du moins depuis sa version 1.1).

Compatibilité avec US-ASCII

Un texte en US-ASCII est codé identiquement en UTF-8 (lorsque le BOM n'est pas utilisé)

Interopérabilité

Du fait qu'un caractère est découpé en une suite d'octets (et non en mots de plusieurs octets), il n'y a pas de problème d'*endianness* (ou « boutisme »).

- Ce problème apparaît avec les codages UTF-16 et UTF-32 par exemple, si on ne les utilise pas avec un marqueur d'ordonnement (appelé *BOM* pour *Byte Order Mark*) codé en début de fichier à l'aide du caractère U+FEFF, qui était auparavant destiné à un autre usage (*ZWNBS* pour *zero-width non-breaking space*, une fonction d'agglutination de mots à afficher sans espace séparatrice ni césure que remplit aujourd'hui le caractère WJ pour *word-joiner*). En revanche, les codages dérivés UTF-16BE, UTF-16LE, UTF-32BE et UTF-32LE sont conçus avec un ordonnancement précis ne nécessitant l'emploi d'aucun BOM.
- Pour différentes raisons de compatibilité (notamment via des processus de transcodage), il est cependant resté admis qu'un *BOM* (U+FEFF), non absolument nécessaire, puisse encore être codé en tête d'un fichier UTF-8 (leur interprétation reste celle du caractère *ZWNBS*, même si de nombreux protocoles ont choisi d'ignorer et filtrer silencieusement ce caractère puisqu'il ne sert plus qu'à cet usage et que son ancienne fonction, quand elle reste nécessaire à l'interprétation du texte lui-même, est désormais transférée sur un autre caractère codé exprès).

Efficacité Pour la plupart des langues à écriture latine, les fichiers de données numériques ou les codes sources de programmes, ou de nombreux protocoles textuels de communication (comme FTP, HTTP ou MIME), qui utilisent abondamment (voire parfois exclusivement dans certaines parties) les caractères US-ASCII, UTF-8 nécessite moins d'octets que l'UTF-16 ou l'UTF-32.

Réutilisabilité De nombreuses techniques de programmation informatique valables avec les caractères uniformément codés sur un octet le restent avec UTF-8, notamment :

- la manière de repérer la fin d'une chaîne de caractères C, car tout octet binaire 00000000 trouvé dans une chaîne de caractères codés en UTF-8 est toujours le caractère nul (en revanche il est alors impossible de représenter le caractère NUL lui-même comme membre de la chaîne de caractères, à moins que l'information de longueur effective du texte codé soit stockée ou transportée en dehors de celui-ci, auquel cas cet octet sera interprété comme tel au sein même des chaînes codées en UTF-8).
- la manière de trouver une sous-chaîne est identique.

Fiabilité Il s'agit d'un codage auto-synchronisant (en lisant un seul octet on sait si c'est le premier d'un caractère ou non).

- Il est possible, depuis n'importe quelle position dans un texte codé, de remonter au premier octet de la séquence en lisant une toute petite quantité d'octets antérieurs, soit au maximum 3 octets, ou de trouver facilement le début de la séquence suivante, là encore en ne sautant qu'au maximum 3 octets) ;
- Une séquence décrivant un caractère n'apparaît jamais dans une séquence plus longue décrivant un autre caractère (comme c'est le cas de Shift-JIS).
- Il n'existe pas de code « d'échappement » changeant l'interprétation (comme caractères) de la suite d'une séquence d'octets.

6.10.5 Inconvénients

Taille variable Les points de code sont représentés en UTF-8 par des séquences d'octets de taille variable (de même qu'en UTF-16), ce qui rend certaines opérations sur les chaînes de points de code plus compliquées : le calcul du nombre de points de code ; le positionnement à une distance donnée (exprimée en nombre de points de code) dans un fichier texte et en règle générale toute opération nécessitant l'accès au point de code de position *N* dans une chaîne.

Une taille variable des caractères d'une chaîne empêche l'exploitation d'algorithmes efficaces en matière de comparaisons de chaînes, telles que l'algorithme de Knuth-Morris-Pratt et pénalise donc fortement les traitements de données en masse comme par exemple dans l'exploitation des bases de données. Ce problème est toutefois davantage lié aux aspects de normalisation que d'encodage.

Efficacité Pour les langues utilisant beaucoup de caractères extérieurs à US-ASCII, UTF-8 occupe sensiblement plus d'espace.

Par exemple, les idéogrammes courants employés dans les textes de langues asiatiques comme le chinois, le coréen ou le japonais (kanji, par exemple) utilisent 3 octets en UTF-8 contre 2 octets en UTF-16.

De manière générale, les écritures employant beaucoup de points de code de valeur égale ou supérieure à U+0800 occupent plus de mémoire que s'ils étaient codés avec UTF-16 (UTF-32 sera plus efficace uniquement pour les textes utilisant majoritairement des écritures anciennes ou rares codées hors du plan multilingue de base, c'est-à-dire à partir de U+10000, mais il peut aussi s'avérer utile localement dans certains traitements pour simplifier les algorithmes, car les caractères y ont toujours une taille fixe, la conversion des données d'entrée ou de sortie depuis ou vers UTF-8 ou UTF-16 étant triviale).

Séquences invalides Par son système de codage, il était éventuellement possible de représenter un code de différentes manières en UTF-8, ce qui pouvait poser un problème de sécurité : un programme mal écrit peut accepter un certain nombre de représentations UTF-8, normalement invalides selon la RFC 3629 et dans les spécifications (maintenant équivalentes entre elles) publiées par l'ISO 10646 et Unicode ; mais ce n'était pas le cas selon la spécification originale, qui permettait de les convertir comme un seul et même caractère.

De fait, un logiciel détectant certaines chaînes de caractères (pour prévenir les injections SQL, par exemple) pouvait échouer dans sa tâche (ce n'est plus le cas si la conformité du codage avec la définition stricte et normalisée d'UTF-8 est vérifiée avant toute chose).

Prenons un exemple tiré d'un cas réel de virus attaquant des serveurs HTTP du Web en 2001 ((en)Crypto-Gram : July 15, 2000 Microsoft IIS and PWS Extended Unicode Directory Traversal Vulnerability Microsoft IIS 4.0/5.0 Web Directory Traversal Vulnerability). Une séquence à détecter pourrait être « *./.* » représentée en ASCII (*a fortiori* en UTF-8) par les octets « 2F 2E 2E 2F » en notation hexadécimale.

Cependant, une manière malformée de coder cette chaîne en UTF-8 serait « 2F C0 AE 2E 2F », appelée aussi en anglais *overlong form* (*forme superlongue*). Si le logiciel n'est pas soigneusement écrit pour rejeter cette chaîne, en la mettant par exemple sous forme canonique, une brèche potentielle de sécurité est ouverte. Cette attaque est appelée *directory traversal*.

Les logiciels acceptant du texte codé en UTF-8 ont été blindés pour rejeter systématiquement ces formes longues car non conformes à la norme : soit le texte entier est rejeté ; mais parfois les séquences invalides sont remplacées par un caractère de substitution (généralement U+FFFD si l'application accepte et traite ce caractère normalement ; parfois un point d'interrogation ou le caractère de contrôle de substitution SUB U+001A de l'ASCII, qui peuvent poser d'autres problèmes de compatibilité) ; moins souvent, ces séquences interdites sont éliminées silencieusement (ce qui est très peu recommandé).

Caractère nul UTF-8 ne peut représenter le caractère de contrôle nul (U+0000) qu'avec un seul octet nul, ce qui pose des problèmes de compatibilité avec le traitement de chaînes qui ne codifient pas séparément leur longueur effective car cet octet nul ne représente alors aucun caractère mais la fin de chaîne (cas très courant en langage C ou C++ et dans les API des systèmes d'exploitation). Si un caractère nul doit être stocké dans un texte sur de tels systèmes, il sera nécessaire de recourir à un système d'échappement, spécifique de ce langage ou système avant de coder en UTF-8 le texte ainsi transformé. En pratique, aucun texte valide ne devrait contenir ce caractère.

Une autre solution est d'utiliser une des séquences interdites dans le codage UTF-8 standard afin de coder le caractère par cette séquence ; mais le texte ainsi codé ne sera pas conforme au codage UTF-8 standard, même si le codage ainsi modifié reste un format de transformation universelle conforme (qui ne doit cependant pas être désigné comme « UTF-8 »). Voir la section ci-dessous relative aux variantes non standards basées sur UTF-8.

6.10.6 Histoire

UTF-8 a été inventé par Kenneth Thompson lors d'un dîner avec Rob Pike aux alentours de septembre 1992^[4]. Appelé alors *FSS-UTF*, il a été immédiatement utilisé dans le système d'exploitation Plan 9 sur lequel ils travaillaient. Une contrainte à résoudre était de coder les caractères nul et '/' comme en ASCII et qu'aucun octet codant un autre caractère n'ait le même code. Ainsi les systèmes d'exploitation UNIX pouvaient continuer à rechercher ces deux caractères dans une chaîne sans adaptation logicielle.

FSS-UTF a fait l'objet d'un standard préliminaire X/Open de 1993^[5] qui fut proposé à l'ISO. Cette dernière l'adopta dans le cadre de la norme ISO/CEI 10646 sous le nom d'abord d'UTF-2, puis finalement UTF-8.

Restrictions successives

Le codage original FSS-UTF était destiné à remplacer le codage multi-octets UTF-1 initialement proposé par l'ISO 10646. Ce codage initialement permissif, permettait plusieurs représentations binaires pour le même caractère (cela

a été interdit dans la version normalisée dans la RFC publiée par le Consortium X/Open, et approuvé par Kenneth Thompson).

De plus il pouvait (dans une version préliminaire non retenue) coder tous les caractères dont la valeur de point de code comprenait jusqu'à 32 bits en définissant un huitième type d'octet (dans des séquences comprenant jusqu'à 6 octets), au lieu des 7 types d'octets finalement retenus pour ne coder (dans des séquences comprenant aussi jusqu'à 6 octets) que les points de code jusqu'à 31 bits dans la version initiale d'UTF-8 (publiée par le Consortium X/Open sous le nom FSS-UTF, puis proposé par le comité technique d'ISO 10646 comme la proposition « UTF-2 » alors encore en concurrence avec la proposition « UTF-1 », jusqu'à ce que la proposition UTF-2 soit retenue et adopte le nom UTF-8 déjà retenu et utilisé dans X/Open et Plan 9).

Ce codage UTF-8 a été restreint encore lorsque Unicode et ISO 10646 sont convenus de n'allouer des caractères que dans les 17 premiers plans afin de maintenir indéfiniment la compatibilité avec UTF-16 (sans devoir le modifier), en restreignant les séquences jusqu'à 4 octets seulement et en n'utilisant que les 5 premiers des 7 types d'octets (ce qui a nécessité de définir comme invalides de nouvelles valeurs d'octet et certaines séquences d'octets pourtant valides individuellement).

Prise en charge

L'IETF exige maintenant qu'UTF-8 soit pris en charge par défaut (et non pas simplement supporté en tant qu'extension) par tous les nouveaux protocoles de communication d'Internet (publiés dans ses RFC numérotées) qui échangent du texte (les plus anciens protocoles n'ont toutefois pas été modifiés pour rendre ce support obligatoire, mais seulement étendus si possible, pour le supporter de façon optionnelle, si cela produit des incompatibilités ou introduit de nouveaux risques de sécurité : c'est le cas de protocoles Internet très utilisés comme DNS, HTTP, FTP, Telnet et de HTML dans ses versions initiales alors pas encore standardisés par le W3C et l'ISO).

Il est devenu incontournable, notamment dans les principaux logiciels de communication du web et aujourd'hui les systèmes d'exploitation :

- **Navigateurs web** : la prise en charge d'UTF-8 commença à être répandue à partir de 1998.
 - Les anciens navigateurs web ne supportant pas UTF-8 affichent tout de même correctement les 127 premiers caractères ASCII ;
 - Netscape Navigator supporte UTF-8 à partir de sa version 4 (juin 1997) ;
 - Microsoft Internet Explorer supporte UTF-8 à partir de sa version 4 (octobre 1997) pour Microsoft Windows et pour Mac OS (janvier 1998) ;
 - Les navigateurs basés sur le moteur de rendu Gecko (1998) supportent UTF-8 : Mozilla, Mozilla Firefox, SeaMonkey, etc.
 - Opera supporte UTF-8 à partir de sa version 6 (novembre 2001) ;
 - Konqueror supporte UTF-8 ;
 - Safari sur Macintosh et Windows supporte UTF-8 ;
 - OmniWeb sur Macintosh supporte UTF-8 ;
 - Chrome de Google supporte UTF-8.
- **Fichiers et noms de fichiers** : de plus en plus courant sur les systèmes de fichiers GNU/Linux et Unix, mais pas très bien supporté sous les anciennes versions de Windows (antérieures à Windows 2000 et Windows XP, lesquels peuvent maintenant les prendre en charge sans difficulté puisque le support d'UTF-8 est maintenant totalement intégré au système, en plus d'UTF-16 présent depuis les premières versions de Windows NT et de l'API système Win32).

Noter que le système de fichiers FAT historique de MS-DOS et Windows a été étendu depuis Windows NT 4 et Windows XP pour prendre en charge UTF-16 (pour une meilleure compatibilité avec NTFS) et non UTF-8 (mais cette transformation est équivalente et transparente aux applications) avec quelques restrictions sur les caractères autorisés ou considérés comme équivalents (de telles exceptions ou restrictions existent aussi sur tous les systèmes de fichiers pour GNU/Linux, Unix, Mac OS X et d'autres systèmes, y compris les systèmes de fichiers distribués sur Internet comme HTTP/WebDAV).

- Client de messagerie : tous les logiciels de messagerie utilisés aujourd'hui supportent UTF-8.
 - Thunderbird ;
 - Windows Mail ;
 - Microsoft Outlook ;
 - Lotus Notes ;
 - Novell GroupWise ;
 - etc.

Extensions non standards

Toutefois, des variantes d'UTF-8 (basées sur les possibilités de codage de la version initiale non restreinte) ont continué à être utilisées (notamment dans l'implémentation de la sérialisation des chaînes Java) pour permettre de coder sous forme d'un échappement multioctets certains caractères ASCII réservés normalement codés sur un seul octet (par exemple le caractère nul).

De plus, certains systèmes utilisent des chaînes de caractères non restreints : par exemple, Java (et d'autres langages y compris des bibliothèques de manipulation de chaînes en C, PHP, Perl, etc...) représentent les caractères avec des unités de codage sur 16 bits (ce qui permet de stocker les chaînes en utilisant le codage UTF-16, mais sans les contraintes de validité imposées par UTF-16 concernant les valeurs interdites et l'appariement dans l'ordre des « demi-codets » ou *surrogates*) ; dans ce cas, les unités de codage sont traitées comme des valeurs binaires et il est nécessaire de les sérialiser de façon individuelle (indépendamment de leur interprétation possible comme caractères ou comme demi-points de code). Dans ce cas, chaque unité de codage 16 bits qui représente un « caractère » (non-contraint) est sérialisé sous forme de séquences comprenant jusqu'à 3 octets chacune, et certains octets interdits par l'implémentation (par exemple les caractères nuls ou la barre de fraction « / » dans un système de fichiers ou d'autres caractères codés sur un octet dans d'autres protocoles) sont codés sous forme de séquences d'échappement à deux octets dont aucun n'est nul, en utilisant simplement le principe de codage de la première spécification de FSS-UTF (avant celle qui a été retenue par le Consortium X/Open dans sa RFC initiale où ces échappements étaient spécifiquement interdits et le sont restés).

Avant l'adoption de la proposition UTF-2 retenue pour UTF-8, il a également existé une variante UTF-1, où les codages multiples étaient impossibles, mais nécessitait un codage/décodage plus difficile devant prendre en compte la position de chaque octet et utilisant un certain nombre de valeurs « magiques ».

Ces variantes ne **doivent pas** être appelées « UTF-8 ».

Une de ces variantes non standards a fait cependant l'objet d'une standardisation ultérieure (en tant qu'alternative à UTF-16 et utilisant des paires de « demi-codets » codés chacun sur 3 octets ; soit 6 octets en tout au lieu de 4 avec UTF-8) : voir CESU-8.

Exemple de variante utilisée en Java Par exemple les API d'intégration des machines virtuelles Java (pour JNI, *Java Native Interface* ou pour la sérialisation des classes précompilées), qui permettent d'échanger les chaînes Java non contraintes sous forme de séquences d'octets (afin de les manipuler, utiliser ou produire par du code natif, ou pour le stockage sous forme de fichier natif codés en suites d'octets) sont suffixées par "UTFChars" ou "UTF", mais ce codage propre à Java n'est pas UTF-8 (La documentation de Sun la désigne comme *modified UTF*, mais certains documents plus anciens relatifs à JNI désignent encore ce codage incorrectement sous le nom *UTF-8*^[6], ce qui a produit des anomalies de comportement de certaines bibliothèques natives JNI, notamment avec les API systèmes d'anciennes plateformes natives qui ne supportent pas nativement les codages de caractères sur plus de 8 bits), car :

- le caractère nul, présent en tant que tel dans une chaîne Java, est codé sous forme de deux octets non nuls (et non un seul octet nul utilisé pour indiquer la fin de séquence) ;
- les demi-codets (en anglais *surrogates*, assignés entre U+D000 à U+D7FF) peuvent être codés librement, dans un ordre quelconque, de même que les points de code interdits normalement pour le codage de caractère (par exemple U+FFFF ou U+FFFE) : aucun test de validité n'est demandé ;
- les séquences d'octets plus longues (sur 4 octets pour représenter les caractères hors du plan multilingue de base) normalisées et valides dans UTF-8 ne sont pas reconnues par la machine virtuelle dans ses API basées sur *modified UTF* (ce qui déclenche des exceptions lors de la conversion demandée par le code natif de la

chaîne 8 bits vers une chaîne Java gérée par la machine virtuelle) : il faut alors réencoder les caractères hors du plan de base (et codés sur 4 octets en UTF-8) sous la forme de deux séquences de 3 octets en *modified UTF*, une pour chaque *surrogate* (comme dans le codage CESU-8) et la conformité des chaînes Java avec Unicode doit être vérifiée et les exceptions gérées ;

- les chaînes Java (de la classe système String) et le type numéral char sont utilisés aussi pour le stockage (sous forme compacte, non modifiable et partageable) de données binaires quelconques (pas seulement du texte), et peuvent aussi être manipulées dans d'autres codages que l'UTF-16 (la seule contrainte étant que les unités de codage individuelle ne doivent pas dépasser 16 bits et doivent être de valeur positive, le bit de poids fort n'étant pas évalué comme un bit de signe).

En conséquence,

- les applications écrites en pur Java (sans code natif) et qui nécessitent l'implémentation de contraintes de codage pour être conformes à Unicode pour le texte **doivent** le demander explicitement et utiliser un des filtres de codage fournis (pour UTF-8, comme aussi pour UTF-16), ou construire et utiliser des classes basées sur la classe String et le type numéral char.
- un texte UTF-8 valide (et manipulé en code natif dans des chaînes sans caractères nuls) nécessite un prétraitement avant de pouvoir être transmis à la machine virtuelle Java via JNI ; notamment, toute séquence codée sur 4 octets (pour un caractère hors du plan de base) doit être transcodée en deux séquences de 3 octets.
- les chaînes obtenues depuis une machine virtuelle Java via les interfaces UTF de JNI nécessitent un prétraitement de contrôle de validité ou de filtrage dans le code natif, avant de pouvoir être utilisées comme du texte UTF-8 valide (il faut détecter les occurrences du caractère null codé en deux octets et, si ce caractère est acceptable par le code natif, le transcoder en un seul octet ; il faut vérifier dans le code natif l'appariement correct des *surrogates*, codés chacun sur 3 octets, et les filtrer si ces séquences ne sont pas refusées comme invalides, puis transcoder toute paire valide de *surrogates* en une seule séquence de 4 octets seulement et non deux séquences de 3 octets).

Ces traitements peuvent être inefficaces pour l'interfaçage de grosses quantités de texte car ils demandent l'allocation de tampons mémoire supplémentaires pour s'interfacer ensuite dans le code natif avec des interfaces système ou réseau qui n'acceptent que l'UTF-8 standard.

Cependant JNI fournit aussi une API binaire plus efficace permettant d'utiliser UTF-16 directement, capable de s'interfacer directement avec les protocoles réseau et les interfaces système (par exemple les API Windows) qui supportent l'UTF-16, sans nécessiter aucune allocation mémoire supplémentaire pour le transcodage (seule la vérification de conformité peut être nécessaire, principalement pour vérifier dans le texte codé l'appariement correct des demi-codets ou *surrogate*, que Java (comme aussi d'autres langages de programmation) permet de manipuler sans restriction de validité dans ses propres chaînes de caractères non dédiées au stockage des seuls textes conformes avec l'UCS). Cette API binaire est supportée sur tous les systèmes où Java a été porté, même ceux dont le système d'exploitation n'offre pas d'API de texte Unicode (la prise en charge pouvant se faire dans l'application native hôte ou en utilisant les bibliothèques standard fournies avec la machine virtuelle Java ou d'autres bibliothèques natives indépendantes.

6.10.7 Notes et références

[1] « Usage of character encodings for websites », sur *W3Techs* (consulté le 18 décembre 2014).

[2] Cependant sur cette page retraçant l'histoire du codage UTF-8 avant 1996 il est dit : « *UTF-8 encoded characters may theoretically be up to six bytes long* », faisant par là référence à l'ensemble des valeurs possibles (plus de deux milliards, codées sur 31 bits au maximum) dans l'édition initiale (aujourd'hui obsolète) de la norme ISO/CEI 10646, cf. section Restrictions successives.

[3] UTF-8 and Unicode FAQ

[4] Histoire de la création d'UTF-8 - par Rob Pike

[5] File System Safe UCS Transformation Format (FSS-UTF)

[6] <http://java.sun.com/docs/books/jni/html/types.html#58973>

6.10.8 Voir aussi

Articles connexes

- UTF-16, UTF-32, CESU-8
- Unicode, ISO/CEI 10646
- ISO 646
- ISO 8859, ISO 8859-1
- Analyse syntaxique

Blocs de caractères Unicode spéciaux contenant des non-caractères

- Table des caractères Unicode - Demi-zone haute d'indirection
- Table des caractères Unicode - Demi-zone basse d'indirection
- Table des caractères Unicode - Formes A de présentation arabes (2^e partie)
- Table des caractères Unicode - Fins de plans : Caractères spéciaux, plan multilingue complémentaire, plan idéographique complémentaire, *plan 3, plan 4, plan 5, plan 6, plan 7, plan 8, plan 9, plan 10, plan 11, plan 12, plan 13*, plan complémentaire spécialisé, zone supplémentaire A à usage privé, zone supplémentaire B à usage privé

Liens externes

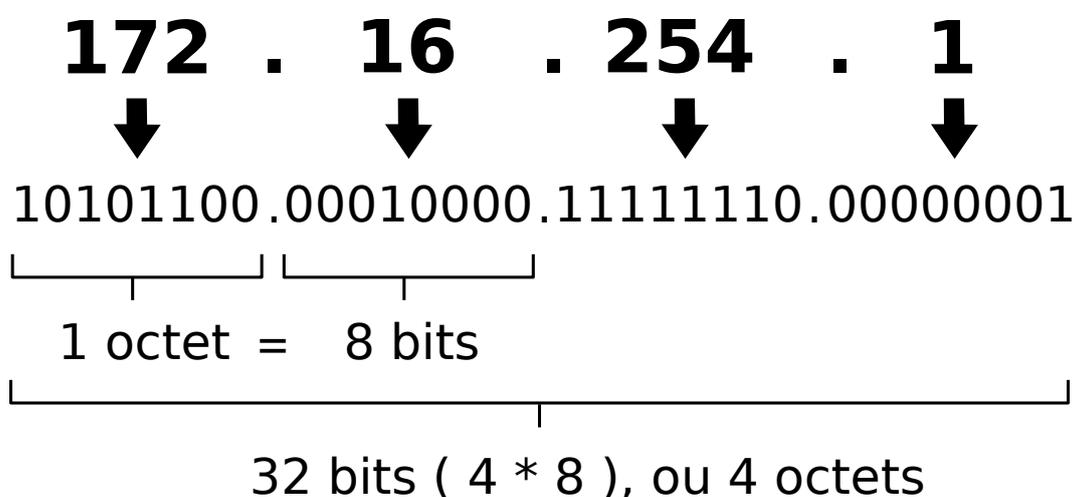
- (fr) Formulaire de conversion UTF-8, UTF-16, UTF-32
- (fr) *Conformité* - Traduction du Standard Unicode en français [PDF]
- (en) The Unicode Consortium, The Unicode Standard, Version 6.0.0, Mountain View (Californie, États-Unis), février 2011 (ISBN 978-1-936213-01-6, présentation en ligne, lire en ligne), chap. 3 (« Conformance »), pp. 88–100.
- (en) RFC 3629 - *UTF-8, a transformation format of ISO 10646*, novembre 2003 (standard, totalement compatible avec Unicode) :
 - (en) RFC 2279 - *UTF-8, a transformation format of ISO 10646*, janvier 1998 (ancienne révision, obsolète) ;
 - (en) RFC 2044 - *UTF-8, a transformation format of Unicode and ISO 10646*, octobre 1996 (version initiale approuvée par l'IETF, obsolète) ;
 - (en) Papier original sur UTF-8, de Rob Pike et Ken Thompson (informatif, obsolète) [PDF].
- (en) RFC 2277 - *IETF policy on character sets and languages*, janvier 1998.
- (en) Histoire de la création d'UTF-8, par Rob Pike.
- (en) UTF-8 Everywhere Manifesto, 2012-09-14 (lire en ligne)
-  Portail de l'informatique
-  Portail de l'écriture

6.11 Adresse IP

Une **adresse IP** (avec IP pour *Internet Protocol*) est un numéro d'identification qui est attribué de façon permanente ou provisoire à chaque appareil connecté à un réseau informatique utilisant l'Internet Protocol. L'adresse IP est à la base du système d'acheminement (le routage) des messages sur Internet.

Il existe des adresses IP de *version 4* (sur 32 bits, soit 4 octets) et de *version 6* (sur 128 bits, soit 16 octets). La version 4 est actuellement la plus utilisée : elle est généralement représentée en notation décimale avec quatre nombres compris entre 0 et 255, séparés par des **points**, ce qui donne par exemple : 212.85.150.134. Les plages d'adresses IPv4 étant proches de la saturation, les opérateurs incitent à la transition d'IPv4 vers IPv6^[1].

Une adresse IPv4 (notation décimale à point)



Adresse IP v4.

6.11.1 Utilisation des adresses IP

L'adresse IP est attribuée à chaque interface avec le réseau de tout matériel informatique (routeur, ordinateur, modem ADSL, imprimante réseau, etc.) connecté à un réseau informatique utilisant l'*Internet Protocol* comme protocole de communication entre ses nœuds. Cette adresse est assignée soit individuellement par l'administrateur du réseau local dans le sous-réseau correspondant, soit automatiquement via le protocole *DHCP*. Si l'ordinateur dispose de plusieurs interfaces, chacune dispose d'une adresse IP spécifique, une interface peut également disposer de plusieurs adresses IP^[2].

Chaque paquet transmis par le protocole IP contient l'adresse IP de l'émetteur ainsi que l'adresse IP du destinataire. Les routeurs IP acheminent les paquets vers la destination de proche en proche. Certaines adresses IP sont utilisées pour la diffusion (*multicast* ou *broadcast*) et ne sont pas utilisables pour adresser des ordinateurs individuels. La technique *anycast* permet de faire correspondre une adresse IP à plusieurs ordinateurs répartis dans Internet.

Les adresses IPv4 sont dites *publiques* si elles sont enregistrées et routables sur Internet, elles sont donc uniques *mondialement*. À l'inverse, les adresses *privées* ne sont utilisables que dans un réseau local, et ne doivent être uniques que dans ce réseau. La traduction d'adresse réseau permet de transformer des *adresses privées* en *adresses publiques* et d'accéder à Internet à partir d'un poste d'un réseau privé.

An IPv6 address (in hexadecimal)

2001:0DB8:AC10:FE01:0000:0000:0000:0000

↓ ↓ ↓ ↓
2001:0DB8:AC10:FE01:: Zeroes can be omitted

↘ ↘ ↘ ↘
 1000000000000001:0000110110111000:1010110000010000:1111111000000001:
 0000000000000000:0000000000000000:0000000000000000:0000000000000000

Adresse IP v6.

6.11.2 Adresse IP et nom de domaine

La plupart des adresses IP des *serveurs* peuvent être converties en un *nom de domaine* et inversement. Le nom de domaine est plus facilement lisible : fr.wikipedia.org est le nom de domaine correspondant à 91.198.174.225. Il s'agit du système de *résolution de noms* (DNS pour *Domain Name System* en anglais).

6.11.3 Classe d'adresse IP

Article détaillé : [Classe d'adresse IP](#).

Jusqu'aux années 1990, on a distingué des classes d'adresse IP qui étaient utilisées pour l'assignation des adresses et par les protocoles de routage. Cette notion est désormais obsolète pour l'attribution et le routage des adresses IP. Attention toutefois : dans la pratique, de nombreux matériels et logiciels se basent sur ce système de classe, y compris les algorithmes de routage des protocoles dit *classless* (cf. *Cisco CCNA Exploration - Protocoles et concepts de routage*).

6.11.4 Sous-réseau

Article détaillé : [Sous-réseau](#).

En 1984, devant la limitation du modèle de classes, la RFC 917 (*Internet subnets*) crée le concept de *sous-réseau*. Ceci permet par exemple d'utiliser une adresse de Classe B comme 256 sous-réseaux de 256 ordinateurs au lieu d'un seul réseau de 65536 ordinateurs, sans toutefois remettre en question la notion de classe d'adresse.

Le masque de sous-réseau permet de déterminer les deux parties d'une adresse IP correspondant respectivement au numéro de réseau et au numéro de l'hôte.

Un masque a la même longueur qu'une adresse IP. Il est constitué d'une suite de chiffres 1 (éventuellement) suivie par une suite de chiffres 0.

Pour calculer la partie sous-réseau d'une adresse IP, on effectue une opération ET logique bit à bit entre l'adresse et le masque. Pour calculer l'adresse de l'hôte, on effectue une opération ET bit à bit entre le complément à un du masque et l'adresse.

En IPv6, les sous-réseaux ont une taille fixe de /64, c'est-à-dire que 64 des 128 bits de l'adresse IPv6 sont réservés à la numérotation d'un hôte dans le sous-réseau.

6.11.5 Agrégation des adresses

En 1992, la RFC 1338 (*Supernetting : an Address Assignment and Aggregation Strategy*) propose d'abolir la notion de classe qui n'était plus adaptée à la taille d'Internet.

Le *Classless Inter-Domain Routing* (CIDR), est mis au point en 1993^[3] afin de diminuer la taille de la table de routage contenue dans les routeurs. Pour cela, on agrège plusieurs entrées de cette table en une seule.

La distinction entre les adresses de classe A, B ou C a été ainsi rendue obsolète, de sorte que la totalité de l'espace d'adressage unicast puisse être gérée comme une collection unique de sous-réseaux indépendamment de la notion de classe. Le masque de sous-réseau ne peut plus être déduit de l'adresse IP elle-même, les protocoles de routage compatibles avec CIDR, dits *classless*, doivent donc accompagner les adresses du masque correspondant. C'est le cas de Border Gateway Protocol dans sa version 4, utilisé sur Internet (RFC 1654 *A Border Gateway Protocol 4*, 1994), OSPF, EIGRP ou RIPv2. Les registres Internet régionaux (RIR) adaptent leur politique d'attribution des adresses en conséquence de ce changement.

L'utilisation de *masque de longueur variable* (*Variable-Length Subnet Mask*, VLSM) permet le découpage de l'espace d'adressage en blocs de taille variable, permettant une utilisation plus efficace de l'espace d'adressage.

Le calcul du nombre d'adresses d'un sous-réseau est le suivant, $2^{\text{taille de l'adresse} - \text{masque}}$.

Un fournisseur d'accès internet peut ainsi se voir allouer un bloc /19 (soit $2^{32-19} = 2^{13} = 8192$ adresses) et créer des sous-réseaux de taille variable en fonction des besoins à l'intérieur de celui-ci : de /30 pour des liens points-à-point à /24 pour un réseau local de 200 ordinateurs. Seul le bloc /19 sera visible pour les réseaux extérieurs, ce qui réalise l'agrégation et l'efficacité dans l'utilisation des adresses.

6.11.6 Base de données des adresses IP

L'IANA, qui est depuis 2005 une division de l'ICANN, définit l'usage des différentes plages d'adresses IP en segmentant l'espace en 256 blocs de taille /8, numérotés de 0/8 à 255/8.

Les adresses IP *unicast* sont distribuées par l'IANA aux registres Internet régionaux (RIR). Les RIR gèrent les ressources d'adressage IPv4 et IPv6 dans leur région. L'espace d'adressage unicast IPv4 est composé des blocs d'adresse /8 de 1/8 à 223/8. Chacun de ces blocs est soit réservé, assigné à un réseau final ou à un registre Internet régional (RIR) ou libre^{[4],[5]}. En février 2011, il ne reste plus aucun bloc /8 libre.

En IPv6, le bloc 2000 ::/3 est réservé pour les adresses unicast globales^[6]. Des blocs /23 sont assignés aux RIR depuis 1999.

Il est possible d'interroger les bases de données des RIR pour savoir à qui est assignée une adresse IP grâce à la commande *whois* ou via les sites web des RIR.

Les RIR se sont regroupés pour former la *Number Resource Organization* (NRO) afin de coordonner leurs activités ou projets communs et mieux défendre leurs intérêts auprès de l'ICANN (l'IANA), mais aussi auprès des organismes de normalisation (notamment l'IETF ou l'ISOC).

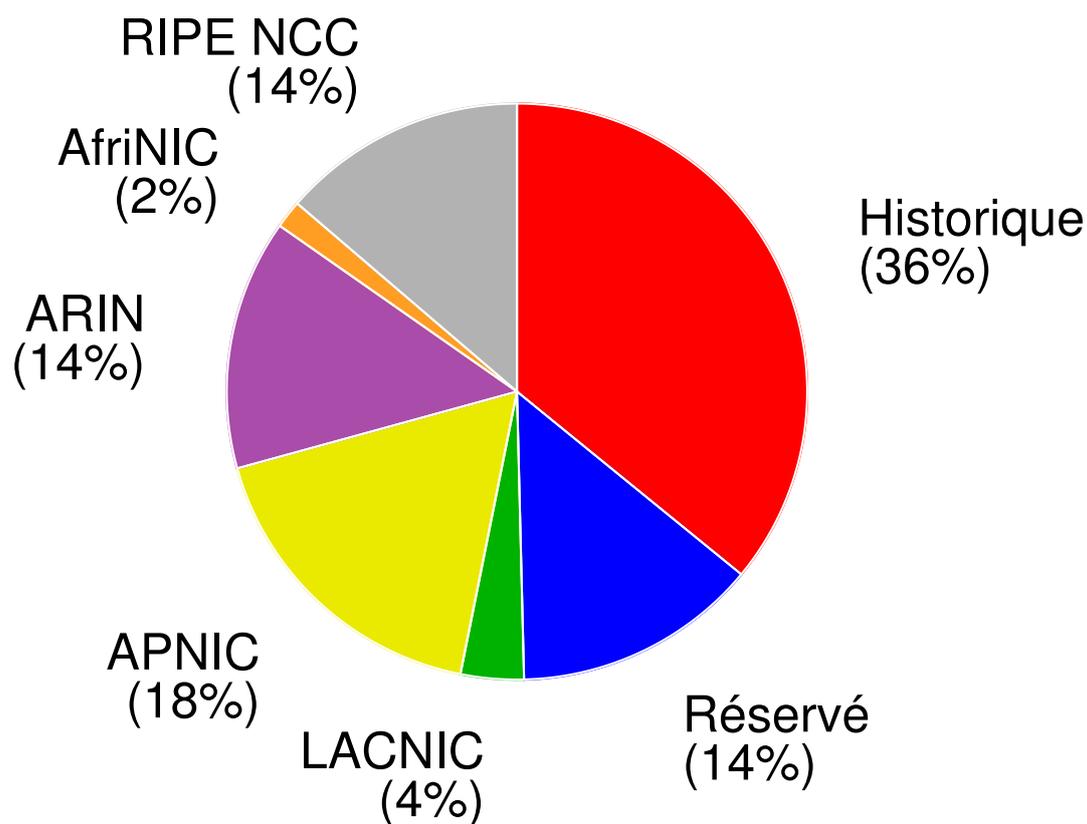
Plages d'adresses IP spéciales

IPv4 Certaines adresses sont réservées à un usage particulier (RFC 5735) :

Adresses privées

Ces adresses ne peuvent pas être routées sur Internet. Leur utilisation par un réseau privé est encouragée pour éviter de réutiliser les adresses publiques enregistrées. Il faut toutefois prévoir qu'il n'y ait pas de doublon lors de l'interconnexion de réseaux privés non prévue lors de leurs créations.

Adresses de diffusion



Distribution de l'espace d'adressage IPv4 en février 2011.

- L'adresse 255.255.255.255 est une adresse de diffusion.
- La première adresse d'un réseau spécifie le réseau lui-même, la dernière est une adresse de diffusion (*broadcast*).

Adresses multicast

En IPv4, tout détenteur d'un numéro d'AS 16 bit peut utiliser un bloc de 256 adresses IP multicast, en 233.x.y.z où x et y sont les 2 octets du numéro d'AS (RFC 3180).

IPv6 Article détaillé : Adresse IPv6.

Les plages d'adresses IPv6 suivantes sont réservées (RFC 5156) :

Adresses spéciales

- `::/128` indique une adresse non spécifiée. Celle-ci est illégale en tant qu'adresse de destination, mais elle peut être utilisée localement dans une application pour indiquer n'importe quelle interface réseau ou sur le réseau dans une phase d'acquisition de l'adresse.

Adresses locales

En IPv6, les adresses locales de site `fec0 ::/10` étaient réservées par la RFC 3513 pour le même usage privé, mais sont considérées comme obsolètes par la RFC 3879 pour privilégier l'adressage public et décourager le recours aux

NAT. Elles sont remplacées par les adresses locales uniques `fc00::/7` qui facilitent l'interconnexion de réseaux privés en utilisant un identifiant aléatoire de 40 bits.

En IPv6, les adresses `fe80::/64` ne sont uniques que sur un lien. Un hôte peut donc disposer de plusieurs adresses identiques dans ce réseau sur des interfaces différentes. Pour lever une ambiguïté avec ces adresses de scope lien local, on devra donc préciser l'interface sur laquelle l'adresse est configurée. Sous les systèmes de type **Unix**, on ajoute à l'adresse le signe pourcent suivi du nom de l'interface (par exemple `ff02::1%eth0`), tandis que sous **Windows** on utilise le numéro de l'interface (`ff02::1%11`).

Adresses expérimentales obsolètes

- `3ffe::/16` et `5f00::/8` étaient utilisés par le **6bone** entre 1996 et 2006.
- `::a.b.c.d/96` (où a.b.c.d est une adresse IPv4) étaient des adresses compatibles IPv4 définies dans la RFC 1884 mais rendues obsolètes par la RFC 4291 en 2006.

Épuisement des adresses IPv4

Article détaillé : [Épuisement des adresses IPv4](#).

La popularité d'Internet a abouti à l'épuisement en 2011 des blocs d'adresses IPv4 disponibles, ce qui menace le développement du réseau.

Pour remédier à ce problème ou repousser l'échéance, plusieurs techniques existent :

- **IPv6**, dont la capacité d'adressage est considérable^[7],
- **NAT**, qui permet à de nombreux ordinateurs d'un réseau privé de partager une adresse publique, mais qui complique le fonctionnement de certains protocoles,
- les **registres Internet régionaux** ont développé des politiques d'affectation d'adresses plus contraignantes, qui tiennent compte des besoins réels à court terme. L'affectation de blocs d'adresses plus petits diminue cependant l'efficacité de l'agrégation des adresses.
- la récupération des blocs attribués généreusement autrefois, certaines entreprises disposent ainsi d'un bloc /8, soit plus de 16 millions d'adresses publiques.

6.11.7 Identification par adresse IP

Article connexe : [Loi Création et Internet#L'identification par l'adresse IP](#).

Celle-ci pose problème^[8], pour plusieurs raisons :

- dans certains cas, une adresse IP publique peut être utilisée par plusieurs personnes simultanément et indépendamment (voir **NAT** et **Carrier Grade NAT**) ;
- il est possible d'usurper l'adresse IP d'autrui dans la source d'un paquet IP dans la mesure où les routeurs utilisent normalement l'adresse destination, l'établissement d'une connexion TCP implique cependant un routage bidirectionnel correct ;
- possibilité de dérouter les systèmes de traçage^[réf. nécessaire] ;
- le traçage de l'adresse IP est souvent utilisé à des fins de marketing^[9], et soupçonné d'influencer les politiques tarifaires^[10].

6.11.8 Notes et références

- [1] IPv6 : opérateurs et hébergeurs en plein chantier, sur le site zdnet.fr du 15 avril 2011.
- [2] [Configuring Multiple IP Addresses and Gateways](#)
- [3] RFC 1518, *An Architecture for IP Address Allocation with CIDR*
- [4] RFC 2373, *IP Version 6 Addressing Architecture*
- [5] IPv4 address space
- [6] IPv6 Unicast Address Assignments
- [7] Internet : nouvelle norme pour les IP, *Le Figaro*, 6 juin 2012.
- [8] Des problèmes juridiques notamment, comme le montre cet article en s'interrogeant sur l'adresse IP au regard de la loi informatique et libertés
- [9] Ariane Krol et Jacques Nantel, « Pêcher le client dans une baignoire : Nouveaux outils du marketing », *Le Monde diplomatique*, n° 711, juin 2013 (lire en ligne)
- [10] « IP tracking : l'eurodéputée Françoise Castex veut une enquête de la Commission », *ZDNet*, janvier 2013 (lire en ligne)

6.11.9 Annexe

Articles connexes

- Adresse IPv6
- Adresse IP virtuelle
- *ARP poisoning*
- *Broadcast*
- *Internet Protocol*
- Usurpation d'adresse IP

Références

Les définitions des adresses IP versions 4 et 6, la notion de classe et la notation CIDR sont documentées dans les *Request for comments* suivants (en anglais) :

Communes

- RFC 997 — *Internet numbers*, mars 1987
- RFC 791 — *Internet Protocol*, septembre 1981 (IP).
- RFC 1519 — *Classless Inter-Domain Routing (CIDR) : an Address Assignment and Aggregation Strategy*, septembre 1993
- RFC 1918 — *Address Allocation for Private Internets*, février 1996
- RFC 1531 — *Dynamic Host Configuration Protocol*, octobre 1993 (DHCP).

IPv4

- RFC 3330 — *Special-Use IPv4 Addresses*, septembre 2002
- RFC 903 — *A Reverse Address Resolution Protocol*, juin 1984 (RARP).

IPv6

- RFC 2460 — *Internet Protocol, Version 6 (IPv6) Specification*, décembre 1998
- RFC 2373 — *IP Version 6 Addressing Architecture*, juillet 1998
- RFC 2893 — *Transition Mechanisms for IPv6 Hosts and Routers*, août 2000

La liste des RIR ainsi que la table d'allocation des adresses se trouvent sur la page *Number Resources* de l'IANA.

Liens externes

- Réseaux TCP/IP, sur Wikibooks
- L'adresse IP est-elle une donnée personnelle (article juridique)

-  Portail de l'informatique

-  Portail d'Internet

-  Portail des télécommunications

6.12 Port (logiciel)

 Pour les articles homonymes, voir Port (homonymie).

Correspondant à la **couche de transport** du modèle OSI, la notion de **port** logiciel permet, sur un ordinateur donné, de distinguer différents interlocuteurs. Ces interlocuteurs sont des programmes informatiques qui, selon les cas, écoutent ou émettent des informations sur ces ports. Un port est distingué par son numéro.

6.12.1 Origine du mot

Port, en informatique, est une traduction erronée de l'anglais *port* (en) ; l'étymologie du mot au sens informatique est le latin *porta* (→ *porte*), et non *portus* (→ *port*)^[1].

6.12.2 Explication métaphorique

Pour simplifier, on peut considérer les ports comme des *portes donnant accès au système d'exploitation* : (Microsoft Windows, Mac OS, GNU/Linux, Solaris...). Pour fonctionner, un programme (par exemple un jeu à accélération 3D/2D, ou un logiciel de retouche photo) ouvre des portes pour entrer dans le système d'exploitation, mais lorsque l'on quitte le programme, la porte n'a plus besoin d'être ouverte.

6.12.3 Utilité

Grâce à cette abstraction, on peut exécuter plusieurs logiciels serveurs sur une même machine, et même simultanément des logiciels clients et des serveurs, ce qui est fréquent sur les systèmes d'exploitation multitâches et multiutilisateurs.

6.12.4 Attribution des ports

Article détaillé : [Liste des ports logiciels](#).

Un port est identifié par un numéro exprimé sur 16 bits. Il existe donc un maximum de 65 535 ports ($2^{16} - 1$) par ordinateur.

L'attribution des ports est faite par le système d'exploitation, sur demande d'une application. Cette dernière peut demander que le système d'exploitation lui attribue n'importe quel port disponible. L'application peut ensuite l'utiliser comme bon lui semble.

Lorsqu'un logiciel client veut dialoguer avec un logiciel serveur, aussi appelé service, il a besoin de connaître le port écouté par ce dernier. Les ports utilisés par les services devant être connus par les clients, les principaux types de services utilisent des ports qui sont dits réservés. Par convention, ce sont tous ceux compris entre 0 et 1 023^[2] inclus et leur utilisation par un logiciel serveur nécessite souvent que celui-ci s'exécute avec des droits d'accès particuliers. Les services utilisant ces ports sont appelés les *services bien connus* ("Well-Known Services").

Le fichier *services* indique la liste de ces services dits well-known. Sous UNIX, ce fichier est directement dans */etc* ; sous Windows, ce fichier est par défaut dans *C:\Windows\System32\drivers\etc*. Les services les plus utilisés sont :

- 20/21, pour l'échange de fichiers via FTP
- 22, pour l'accès à un shell sécurisé Secure SHell, également utilisé pour l'échange de fichiers sécurisés SFTP
- 23, pour le port telnet
- 25, pour l'envoi d'un courrier électronique via un serveur dédié SMTP
- 53, pour la résolution de noms de domaine en adresses IP : DNS
- 67/68, pour DHCP et bootpc
- 80, pour la consultation d'un serveur HTTP par le biais d'un navigateur web
- 110, pour la récupération de son courrier électronique via POP
- 123 pour la synchronisation de l'horloge : Network Time Protocol (NTP)
- 143, pour la récupération de son courrier électronique via IMAP
- 389, pour la connexion à un LDAP
- 443, pour les connexions HTTP utilisant une surcouche de sécurité de type SSL : HTTPS
- 465, pour l'envoi d'un courrier électronique via un serveur dédié utilisant une surcouche de sécurité de type SSL : SMTPS
- 500, port utilisé pour le canal d'échange de clés IPsec
- 636, pour l'utilisation d'une connexion à un LDAP sécurisé par une couche SSL/TLS
- 1352, pour le protocole Lotus Notes Domino
- 1433, serveur de base de données MS SQL
- 1521, serveur de base de données Oracle Database
- 1723, pour l'utilisation du protocole de VPN PPTP
- 3306, serveur de base de données MySQL
- 3389, pour la prise de contrôle à distance RDP
- 5432, serveur de base de données PostgreSQL
- 6667, pour la connexion aux serveurs IRC

Toutefois, les conventions ci-dessus peuvent ne pas être respectées pourvu que le client et le serveur soient cohérents entre eux et que le nouveau numéro choisi ne soit pas déjà utilisé par ailleurs (on peut ainsi activer différentes versions d'un même service sur des numéros de ports différents, pour des tests). La plupart des logiciels de communication permettent d'effectuer ce réglage.

6.12.5 Notes et références

[1] L'anglais *port* peut se traduire alors en français par *sabord*

[2] Liste des ports Tcp et Udp, sur le site frameip.com

6.12.6 Voir aussi

Articles connexes

- Liste des ports logiciels
- Port matériel

Lien externe

- (en) Service Name and Transport Protocol Port Number Registry, sur le site iana.org
- Une base de donnée pour trouver la description d'un port TCP et UDP, sur le site frameip.com



Portail d'Internet



Portail de l'informatique

Chapitre 7

Sources, contributeurs et licences du texte et de l'image

7.1 Texte

- **Échange de données informatisé** *Source* : <http://fr.wikipedia.org/wiki/%C3%89change%20de%20donn%C3%A9es%20informatis%C3%A9?oldid=111093028> *Contributeurs* : Looxix, Gérard, Boism, Emmanuel legrand, MedBot, Sam Hocevar, Francois Trazzi, Phe-bot, Kwisark, Ancalagon, Nis, CafeFroid, Xavier Combelle, Benjamin Pineau, Ellywa, Sherbrooke, BrightRaven, Laurent Jerry, Epommate, Chobot, RobotE, Romanc19s, Volkmar, Arnaud.Serander, Zwobot, RobotQuistnix, FlaBot, YurikBot, Eskimbot, MistWiz, Jerome66, MMBot, TED, Sfick, Ben81, Pautard, Geoffroy.aubry, Thidras, Xofc, Jmax, Malost, PieRRoBoT, PhilFree, Thijs !bot, Chaoborus, Laurent Nguyen, Le Pied-bot, Ce'dric, .anaconadobot, Ftiercel, Rei-bot, Salebot, Speculos, Jmnil, Isaac Sanolnacov, DorganBot, VolkovBot, Fluti, Hdeumpty, Cerebellum, Kyro, Hercule, SkoRo, PixelBot, Lejmr, Ertezoute, BotSottile, Factory, Elfis, Tanhabot, JackPotte, Luckas-bot, Fnaudin, ArthurBot, Rubinbot, Lomita, RedBot, ContactGALIA, EmausBot, Kilith, EditionPage, ChuispastonBot, Utilisateur disparu, Nevenlca, Mister Toufou, Fhal.daniel, Addbot et Anonyme : 86
- **Organization for the Advancement of Structured Information Standards** *Source* : <http://fr.wikipedia.org/wiki/Organization%20for%20the%20Advancement%20of%20Structured%20Information%20Standards?oldid=97613451> *Contributeurs* : Olrick, Nataraja, Traroth, Céréales Killer, (:Julien :), MedBot, Sam Hocevar, Phe-bot, NeMeSiS, PieRRoMaN, Ork, Sherbrooke, Fobos, ZeMeilleur, Romanc19s, RobotQuistnix, YurikBot, Odyssee, Bortzmeyer, Toutoune25, Loveless, Best friend fr, TonioVins, PieRRoBoT, NicoV, En passant, Reibot, DodekBot, VolkovBot, Chicobot, MonstroLinux, Ange Gabriel, JeanTwix, BOTarate, Alexbot, Escherichia coli, Δ, SebGR, Cantons-de-l'Est, Xqbot, Asturcon01, RedBot, ZéroBot, Addbot, Fenestride et Anonyme : 10
- **GS1** *Source* : <http://fr.wikipedia.org/wiki/GS1?oldid=98372236> *Contributeurs* : Ofol, Sherbrooke, BrightRaven, Loveless, Pautard, MetalGearLiquid, GaMip, PhilFree, Thijs !bot, Speculos, Walpole, Pinprenelle, VolkovBot, Louperibot, MystBot, Ken123BOT, LordAnubisBOT, Amstramgrampikepekeolegram, DumZiBoT, Purbo T, JackPotte, Merimac, Cantons-de-l'Est, Thierryrnet, Komeil, MastiBot, TobeBot, KamikazeBot, Ripchip Bot, EmausBot, Memole, InfomanTheOnlyOne, MerIiwBot, Addbot et Anonyme : 2
- **GS1 France** *Source* : <http://fr.wikipedia.org/wiki/GS1%20France?oldid=88506524> *Contributeurs* : Alno, MedBot, Vincnet, Leag, BrightRaven, David Berardan, Litlok, Chaps the idol, Le gorille, Pautard, Davric, Salebot, Walpole, Nicourse, Alecs.bot, DumZiBoT et Anonyme : 7
- **Comité français d'organisation et de normalisation bancaires** *Source* : <http://fr.wikipedia.org/wiki/Comit%C3%A9%20fran%C3%A7ais%20d'organisation%20et%20de%20normalisation%20bancaires?oldid=103200795> *Contributeurs* : Céréales Killer, Marc Lacoste, Gdgourou, Régis B., Xfigpower, BMR, MMBot, ChoumX, Orthomaniaque, Alecs.bot, Expertom, Olivier.descout, Gmetman, Nouisser et Anonyme : 4
- **Groupement pour l'amélioration des liaisons dans l'industrie automobile** *Source* : <http://fr.wikipedia.org/wiki/Groupement%20pour%20l'am%C3%A9lioration%20des%20liaisons%20dans%20l'industrie%20automobile?oldid=108331445> *Contributeurs* : Chris a liege, Pixeltoo, Bob08, Jrcourtois, Ascaron, OdIn, Fabriced28, Critias, Durifon, Ir4ubot, ZetudBot, Loreleil, ContactGALIA, Berdea, Mario93, Pierre.alix, Gratus et DickensBot
- **VDA** *Source* : <http://fr.wikipedia.org/wiki/VDA?oldid=108858138> *Contributeurs* : Archeos, Phe-bot, Alecs.y, Sfick, Pautard, Bouktin, MirgolthBot, Staffilo, AlleborgoBot, Catalysebot, HerculeBot, Luckas-bot, SassoBot, Alex-F, Fharchi, Addbot et Anonyme : 2
- **EDI-Optique** *Source* : <http://fr.wikipedia.org/wiki/EDI-Optique?oldid=97639954> *Contributeurs* : Pautard, Fm790, JLM, Kyro, Ir4ubot, Herr Satz, Sundgauvien38, Lomita et Anonyme : 2
- **Échange de données informatisées pour l'administration, le commerce et le transport** *Source* : <http://fr.wikipedia.org/wiki/%C3%89change%20de%20donn%C3%A9es%20informatis%C3%A9es%20pour%20l'administration%2C%20le%20commerce%20et%20le%20transport?oldid=110675088> *Contributeurs* : Med, Koyuki, Archeos, Meodudlye, Kwisark, Olivier Mengué, Romary, ClementSeveillac, Jef-Infojef, Bayo, FlaBot, MistWiz, BMR, Litlok, Pio, Traumrune, Silex6, Pautard, Geoffroy.aubry, GaMip, FritzG, Thijs !bot, Bombastus, Fm790, Analphabot, Nicourse, Mv0001, FritzGG, Alecs.bot, LordAnubisBOT, Hxhbot, Hercule, DumZiBoT, Expertom, SilvononBot, ZetudBot, LaaknorBot, Luckas-bot, Alla.bruno, EmausBot, Fcarcena01, MahdiBot, Addbot, Fabercas et Anonyme : 9
- **Electronic Business using XML** *Source* : <http://fr.wikipedia.org/wiki/Electronic%20Business%20using%20XML?oldid=109064053> *Contributeurs* : Jyp, Haypo, Fafnir, Francois Trazzi, Phe-bot, Kwisark, Olivier Mengué, O. Morand, Leag, Poulos, Elg, ZeMeilleur, Romanc19s, RobotQuistnix, FlaBot, Eskimbot, Alain r, Jerome66, Jcmorand, Litlok, Loveless, Traumrune, Neokilly, Pautard, GaMip, Genium, PieRRoBoT, NicoV, Thijs !bot, Escarbot, Deep silence, Nono64, CommonsDelinker, Analphabot, Salebot, Isaac Sanolnacov,

Jonathan1, Azerty72, VolkovBot, Nodulation, Edificas, Eric JARRY, Asaintguilhem, Virginie G, Pascal Boulerie, Psouhard, Alecs.bot, Hercule, HerculeBot, WikiCleanerBot, ZetudBot, JackPotte, Penjo, Cantons-de-l'Est, KamikazeBot, Esnico30, EdoBot, OrlodrimBot, Mattho69, Addbot et Anonyme : 10

- **Electronic Data Interchange for the Electric Industry** *Source* : <http://fr.wikipedia.org/wiki/Electronic%20Data%20Interchange%20for%20the%20Electric%20Industry?oldid=98409403> *Contributeurs* : Garwiel, Arnaud.Serander, Loveless, Pautard, Ahbon ?, Fm790, Jmnil, Dsant, Vlaam, Dhatier et Lebronj23
- **Échange télématique banque-clients** *Source* : <http://fr.wikipedia.org/wiki/%C3%89change%20%C3%A9%C3%A9matique%20banque-clients?oldid=106871610> *Contributeurs* : Phe-bot, Gdgourou, Stéphane FILLOD, Xfigpower, Vlad, Nostradamus, Jmax, GaMip, Thargos, VolkovBot, BlueGinkgo, Nicourse, Hrouvroy, Philippe Pujol, Zil, Rastacouette, Webmastercedricom, Archimèa, RatBOT, Drongou, DUTERQUE, GrrrrBot, MedoFree, David2paris15, Gmetman, Addbot et Anonyme : 8
- **Electronic Banking Internet Communication Standard** *Source* : <http://fr.wikipedia.org/wiki/Electronic%20Banking%20Internet%20Communication%20Standard?oldid=110846629> *Contributeurs* : Leag, Milvus, Vlad, Jmax, Sebleouf, Consulnico, Isaac Sanolnacov, TXiKiBoT, Olivier.descout, Mro, ZetudBot, Luckas-bot, Drongou, Nsonnier, EmausBot, Gmetman, Le pro du 94 :, Groupe cedricom, Dref74, Addbot et Anonyme : 6
- **Edificas** *Source* : <http://fr.wikipedia.org/wiki/Edificas?oldid=79739854> *Contributeurs* : Phe-bot, MaCRoEco, Nicolas Ray, Bob08, Matpib, MMBot, Pautard, GaMip, Lacrymocéphale, PieRRoBoT, NicoV, Le Pied-bot, Cqui, Edificas, Tmo85, EDIFICAS et Anonyme : 1
- **Protocole de communication** *Source* : <http://fr.wikipedia.org/wiki/Protocole%20de%20communication?oldid=108835538> *Contributeurs* : Athymik, Calo, Valéry Beaud, Med, Ryo, Youandme, Mikue, Vargenau, Oz, Ploum's, Howard Drake, HasharBot, Koyuki, MedBot, Phe-bot, MG, Weft, Kassus, François-Dominique2, Poleta33, Nicolas Ray, Len'Alex, Velisarius, Epommate, Apitux, Korg, Aris, Mike2, Medium69, Loveless, Silex6, Christopheg, SashatoBot, Malost, Asibasth, Linan, Thijs !bot, Escarbot, Piglop, Laurent Nguyen, Rhizome, Fm790, Zedh, Catskingloves, Salebot, Zorrobot, Guiillaume, Moonray, Guillipetiot, SieBot, SniperMaské, GLec, Alphos, Ir4ubot, DragonBot, HerculeBot, Linedwell, JackPotte, Jotterbot, ArthurBot, Nairod.brain, Dabour, GrrrrBot, Kilith, 0x010C, MerllwBot, Addbot et Anonyme : 37
- **Réseau à valeur ajoutée** *Source* : <http://fr.wikipedia.org/wiki/R%C3%A9seau%20%C3%A0%20valeur%20ajout%C3%A9?oldid=96947710> *Contributeurs* : Semnoz, Serged, Bob08, Mi Ga, Givet, DumZiBoT, Michco, WikiCleanerBot, JackBot, MastiBot, GrrrrBot et Addbot
- **European Network Exchange (ENX) Association** *Source* : [http://fr.wikipedia.org/wiki/European%20Network%20Exchange%20\(ENX\)%20Association?oldid=107912278](http://fr.wikipedia.org/wiki/European%20Network%20Exchange%20(ENX)%20Association?oldid=107912278) *Contributeurs* : Polmars, Pautard, Macassar, Sebleouf, Speculos, Orthomaniaque, Nemesis 12, Linedwell, Penjo, Coyote du 57, Orlodrim, StefG en, Galia-Assoc, RustyBSD, Autonovisti et Anonyme : 4
- **X.25** *Source* : <http://fr.wikipedia.org/wiki/X.25?oldid=112124442> *Contributeurs* : Vargenau, Capbat, Solarus, Phe, Bibi Saint-Pol, Lucasbfr, Bap, Jef-Infojef, Bob08, Elg, Stéphane33, Plyd, FlaBot, Ash Crow, YurikBot, Eskimbot, Soccarfr, Pautard, Eiki, Mwarf, Manu1400, MetalGearLiquid, Liquid-aim-bot, Daniel*D, Thijs !bot, Doom*, BeLz, JAnDbot, Zedh, Huzzlet the bot, Speculos, Isaac Sanolnacov, Bapti, VolkovBot, BZP, -Nmd, SieBot, Catalysebot, Ken123BOT, Alecs.bot, Smeet666, FitzSai, DumZiBoT, GLec, Alexbot, Davgrps, HerculeBot, Slycooper, Luckas-bot, ABACA, NegroTruc, Xqbot, Tango Panaché, JackBot, ViRally77, Rag.lien, Lostinthiswhirlpool, GrrrrBot, Helgismidh, ZéroBot, ChuispastonBot, YmFzZTY0, Jean-Luc Cavey, AvicBot, Addbot et Anonyme : 43
- **X.400** *Source* : <http://fr.wikipedia.org/wiki/X.400?oldid=111969567> *Contributeurs* : Vargenau, Bibi Saint-Pol, Neuceu, YurikBot, Bortz-meyer, Litlok, Traumrune, Mrambil, Thijs !bot, Escarbot, Kyle the bot, Le Pied-bot, TXiKiBoT, Louperibot, Wanderer999, Vlaam, LaaknorBot, Xqbot, JackBot, GrrrrBot, Addbot et Anonyme : 3
- **EDIINT** *Source* : <http://fr.wikipedia.org/wiki/EDIINT?oldid=97750155> *Contributeurs* : Kwisark, Vincnet, Benjamin Pineau, Litlok, ChloeD, Pautard, Laocian, Le Pied-bot, Speculos, Zil, Vlaam, JackBot et Anonyme : 2
- **Simple Mail Transfer Protocol** *Source* : <http://fr.wikipedia.org/wiki/Simple%20Mail%20Transfer%20Protocol?oldid=113412372> *Contributeurs* : Athymik, Calo, PhD, Alvaro, Orthogaffe, Oz, Ploum's, HasharBot, Wilinckx, Sanao, Marc Mongenet, MedBot, Francois Trazzi, Phe-bot, Turb, Bibi Saint-Pol, BernardM, Vincnet, Xavier Combelle, Mcannac, YolanC, Laurent Jerry, KiaN, Elfi, DocteurCosmos, Chobot, Vanheu, Romanc19s, MagnetiK, Record, RobotQuistnix, FlaBot, YurikBot, Tos42, Mike2, Medium69, Lt-wiki-bot, MMBot, Sitelec, Cyrildemont, MelancholieBot, Shawn, Grecha, Pautard, Maxxwayne, Manu1400, MetalGearLiquid, Mathgb, Daniel*D, Escarbot, Kyle the bot, Arkanosis, .anacondabot, Nono64, Mickey45, Piquart, Jotum, Salebot, Speculos, Zorrobot, Dracou, AlnoktaBOT, VolkovBot, Tmaquet, Melkor73, -Nmd, Ptbotgourou, Grumlin, Pascal Boulerie, SieBot, Louperibot, DaBot, OKBot, Alecs.bot, Bub's, Gnieark, Theodore.m, GLec, Alphos, Web33, Djmoa, Trazfr, Alexbot, Mro, HerculeBot, JackPotte, Benjazia, Baryum, Anarion9998, Luckas-bot, Julien Stuby, Jotterbot, GrouchoBot, Talabot, Xqbot, MathsPoetry, Rubinbot, Lomita, Randmot, EmausBot, HRoestBot, WikitanvirBot, ChuispastonBot, Peleki, Enrevseluj, VVVF, Addbot, GratusBot, Do not follow et Anonyme : 74
- **File Transfer Protocol** *Source* : <http://fr.wikipedia.org/wiki/File%20Transfer%20Protocol?oldid=113676690> *Contributeurs* : Aoineko, Anthere, Didier, Ryo, Youandme, Virtualblackfox, Mm, Dirac, Popolon, Phido, Orthogaffe, Oz, JidGom, ThomasClavier, HasharBot, Abrahams, R, AlainBb, Zubro, Koyuki, Manchot, Tieno, Okki, Yannig, Alkarex, Archibald, Phe, Marc Mongenet, MedBot, Sam Hocevar, Iznogood, Phe-bot, Liné1, NaSH, Bibi Saint-Pol, Theocrite, SRombauts, Comte0, NeMeSiS, Barthelemy, Jef-Infojef, Dake, T, DD-PAAlphaTiger1, Katsoura, Witoki, Turbant, Caerbannog, Padawane, Maxxyme, GiminiK, V0id, KiaN, DocteurCosmos, Aboumael, Korg, MarcFerrand, Chobot, Aris, Stanlekub, RexGeet, Romanc19s, David Berardan, MagnetiK, Zivax, RobotQuistnix, FlaBot, Nico@nc, Quaternion, Pfv2, YurikBot, LeonardoRob0t, Herve s, Eskimbot, Guillom, Jerome66, CaptainHaddock, MMBot, Ecclecticus, 16@r, Chaps the idol, MagnetiK-BoT, Sitelec, Diego L., Babskwal, Shawn, Dadu, Pautard, Noar, Xofc, SashatoBot, MetalGearLiquid, Genium, Daniel*D, Thijs !bot, OlivierMehani, Kto, TaraO, Jakit, Mrfred, Escarbot, Kyle the bot, JAnDbot, DavidL, Clem23, Nono64, RM77, Elodieleon, Numbo3, Analphabot, Salebot, Nebula38, Speculos, AlnoktaBOT, SharedX, TXiKiBoT, VolkovBot, AmaraBot, Melkor73, -Nmd, Synthebot, Nicmic, AlleborgoBot, BotMultichill, SieBot, Louperibot, JLM, Lucas74330, Nicolargo, Funnycat, DecereBrain, LordAnubisBOT, Bub's, VsBot, DumZiBoT, Loic Ig, GLec, Rushtof, Alexbot, Mro, HerculeBot, SilvononBot, ZetudBot, EivindBot, JackPotte, CarsracBot, Luckas-bot, Julien Stuby, ABACA, GrouchoBot, Olivier d'ALLIVY KELLY, Pdefretin, SassoBot, Xqbot, MathsPoetry, Al Maghi, RibotBOT, Rubinbot, BenzolBot, Lomita, TobeBot, Jno972, Dynamik-bot, Gladir, EmausBot, Kilith, HRoestBot, Keketoby, Hlm Z., WikitanvirBot, ChuispastonBot, MerllwBot, OrlodrimBot, Ayutan, Michelroux, PardusTigris, Reptilien.19831209BE1, CasualNib, Makecat-bot, OriKriBot, Antares7364, Housterdam, Addbot, NexusNet, Raidko, Do not follow et Anonyme : 162

- **File Transfer Protocol Secure** *Source* : <http://fr.wikipedia.org/wiki/File%20Transfer%20Protocol%20Secure?oldid=107853262> *Contributeurs* : Tieno, Alkarex, Bibi Saint-Pol, Notafish, T, Mcannac, FlaBot, YurikBot, Eskimbot, Litlok, MetalGearLiquid, Seero, Thijs !bot, Kto, JAnDbot, IALex, AlnoktaBOT, VolkovBot, -Nmd, Louperibot, Luckas-bot, JackBot, TobeBot, Ver-bot, EmausBot, Hoquei44, Sisqi, ZéroBot, Addbot et Anonyme : 4
- **Secure File Transfer Protocol** *Source* : <http://fr.wikipedia.org/wiki/Secure%20File%20Transfer%20Protocol?oldid=107853235> *Contributeurs* : Tieno, Marc Mongenet, Bibi Saint-Pol, Romainhk, NeMeSiS, T, Nicolas Ray, Mcannac, Mit-Mit, FlaBot, Litlok, Loveless, MetalGearLiquid, Desaparecido, Nono64, Taerendil, MystBot, Bub's wikibot, JackPotte, Addbot et Anonyme : 5
- **SSH File Transfer Protocol** *Source* : <http://fr.wikipedia.org/wiki/SSH%20File%20Transfer%20Protocol?oldid=103088815> *Contributeurs* : HasharBot, Tieno, Nyco, Phe-bot, Liné1, Grimko, NeMeSiS, Dake, T, Bayo, Gratyn, Elg, Aris, David Berardan, Gzen92, Neustradamus, RobotQuistnix, FlaBot, Manuc66, Sequajectrof, Jmax, Manu1400, Liquid-aim-bot, Stephane.lecorne, 5afd4770411ca76c, JAnDbot, MikaelB, Zorrobot, TXiKiBoT, Chicobot, Louperibot, MystBot, STBot, Alecs.bot, LordAnubisBOT, Bub's, DumZiBoT, GLec, Darkicebot, ZetudBot, AkhtaBot, GrouchoBot, Gknor, Xqbot, Rubinbot, Ripchip Bot, Zylabon, Hlm Z., WikitanvirBot, OrlodrimBot, Lydie Noria, Aysheaia, Addbot et Anonyme : 21
- **Applicability Statement 2** *Source* : <http://fr.wikipedia.org/wiki/Applicability%20Statement%20?oldid=106871529> *Contributeurs* : Phe-bot, Benjamin Pineau, Koko90, ZeMeilleur, FlaBot, MMBot, Pautard, Jmax, Ji-Elle, Liquid-aim-bot, Stephane, Chaoborus, Chicobot, ArnaudM, Dsant, Camico, Babelabout, Addbot et Anonyme : 11
- **Axway** *Source* : <http://fr.wikipedia.org/wiki/Axway?oldid=109046932> *Contributeurs* : Panoramix, GL, Pok148, LLM, ManuD, Arnaud 25, Fm790, Copros, Ange Gabriel, Stéphane Veyret, ZetudBot, Guillaume70, Arroser, Skull33, Nezdék, Berdea, MerlIwBot, 3d Communication, Dave Murphy et Anonyme : 9
- **Generix** *Source* : <http://fr.wikipedia.org/wiki/Generix?oldid=111550478> *Contributeurs* : GDC, Inisheer, Loveless, Pautard, Genium, Brejnev, Dauphiné, DumZiBoT, François - Lille, HerculeBot, Centrepompidou, Wikinade, Arroser, Luluberlu22, Mikefuhr, EmausBot, Ltrlg, Universretail, Hfilben, Addbot et Anonyme : 6
- **International Business Machines** *Source* : <http://fr.wikipedia.org/wiki/International%20Business%20Machines?oldid=113678032> *Contributeurs* : Alno, Greudin, HasharBot, Cœur, Koyuki, Symac, Deelight, Phe, Proton, MedBot, Phe-bot, Turb, Romainhk, ADM, ~Pyb, Urban, Fylyp22, Escaladix, Notafish, François-Dominique2, Cr0vax, O. Morand, Darkoneko, Maïeul, YannickPatois, Pixeltoo, Vincnet, Leag, GastelEtwane, Tiennel, Thibault Taillandier, Pseudomoi, Koko90, Poil, Yurik, MrFahrenheit, Maxxyme, Gdgourou, Diderobot, L'amateur d'aéroplanes, Epommate, Gede, Chobot, Elvin, Gribeco, GôTô, RobotE, Zetud, Romanc19s, Ddardenne, Probot, Arnaud.Serander, Pok148, Jydidier, Gzen92, Jeanot, Zwobot, RobotQuistnix, Gpvsobot, FlaBot, Arnaudh, Necrid Master, Bcutiere, Pfv2, YurikBot, LeonardoRob0t, Eskimbot, MistWiz, Ico, BMR, MMBot, Toutoune25, Mi Ga, Bobochan, Oziris, Loveless, Explopulator, Arrakis, Megateuf, Benjism89, Oxo, Chlewbob, Tatoute, Polmars, Pautard, Blone, Mecreant, Démocrite, NGPixel, Barraki, SashatoBot, Hokkos, Pld, MetalGearLiquid, Larrousiney, Liquid-aim-bot, GaMip, Pso, Morus, Eloworld, Do it, Inicky, Ironbrother, PhilFree, Rhadamante, Daniel*D, CK, Thijs !bot, En passant, Escarbot, Axou, Brunodesacacias, RémiH, Deep silence, Arnaud 25, Le Pied-bot, JAnDbot, Fm790, Épiméthée, Nono64, Bilbax, Sebleouf, Urelianu, Jmcib, Philippe rogez, CommonsDelinker, Verbex, Beeper, Erabot, RigO-Luche, NassWiki, LeGéantVert, Jerometron, Wikig, Salebot, Yodaspirine, MyBot, Akeron, Speculos, Zorrobot, Pythakos, TheRookie, Carolebergara, Isaac Sanolnacov, Ardu Petus, AlnoktaBOT, Enguerrand VII, Idioma-bot, TXiKiBoT, Aibot, VolkovBot, Ducloy, 20ce, AmaraBot, Chicobot, Pierre.fischof, Synthebot, Ptbotgourou, Crom1, Orthomaniaque, Boutouil, Antistress, Gz260, Doalex, Yannledu, SieBot, YonaBot, Sylfred1977, Olevy, Durifon, Kyro, Udufruduhu, Alecs.bot, LordAnubisBOT, Vlaam, Hercule, Tambourlaine, Fitz-Sai, DumZiBoT, GLec, PicsLogoLoader, DragonBot, Orphée, Jace stknights, Fanfwah, Cymbella, Rinaku, Alexbot, Darkicebot, Mro, Mayayu, Oyp, Efka, HerculeBot, WikiCleanerBot, WikiDreamer, G-37, ZetudBot, Ggal, WikiDreamer Bot, Bub's wikibot, Fiaaasco, Blogbreather, NjardarBot, Fabienamnet, Highrock, Luckas-bot, Amirobot, Micbot, Jotterbot, FrançoisA, Zorglube, Anne Bauval, Copyleft, Cantons-de-l'Est, Jacques Ballieu, NobelBot, Xqbot, Obersachsebot, RibotBOT, Tango Panaché, JackBot, Nouill, Kanabiz, Alex-F, Benji76, Orbitale, XeBot, TobeBot, Botozor, Hippo75, Duncanogi, OgilvyOne, Ysabel01, Ripchip Bot, VanBrooken, Toto Azéro, Frakir, Zorglub27, EmausBot, Kilith, TulipeJaune, JackieBot, Henri75, LD, Jules78120, Daehan, MerlIwBot, Eric-vrcl, OrlodrimBot, Elliott006, Sergelucas, AvicBot, Pano38, Raptor.cbre, Minsbot, Skrillex75016, MatthieuCroissant, Tibault35, Joel.milgram, Jf380, LectriceDuSoir, Addbot, AméliorationsModestes, Sismarinho, Echekémat, LOLOLOL52, Cchaudey, JohnnyMilliard, Thomas Négrier, FloozMaster et Anonyme : 151
- **Seeburger** *Source* : <http://fr.wikipedia.org/wiki/Seeburger?oldid=104449535> *Contributeurs* : Phe-bot, Chris93, Thierry Bingen, Litlok, Vivarés, Pautard, Aaker, Jmcib, Gothmarilyne, ZetudBot, RatBOT, ZéroBot, Addbot et Anonyme : 3
- **Tibco Software** *Source* : <http://fr.wikipedia.org/wiki/Tibco%20Software?oldid=103342851> *Contributeurs* : Criric, Jmax, Speculos, Bub's, Mabifixem, Nezdék, EmausBot, OrlodrimBot, Rahl43bot, Addbot et Anonyme : 3
- **Théorie de l'information** *Source* : <http://fr.wikipedia.org/wiki/Th%C3%A9orie%20de%20l'information?oldid=111817064> *Contributeurs* : Erwan, Dtcube, R4f, Ryo, Mm, Orthogaffe, Moala, HasharBot, Aldoo, Haypo, Tieno, Marc.lasson, MrTout, Phe, MedBot, Sam Hocevar, Xmlizer, Phe-bot, Dromygo, Turb, François-Dominique, Smily, Romainhk, Markadet, Fylyp22, Escaladix, Bap, Gus2, Theon, MickaëlG, JeanClem, Xavier Combelle, Yurik, Vicmackey, Sherbrooke, Christophe.moustier, Mirgolth, Michel Volle, Epommate, Wart Dark, Chobot, Eusebius, Seb35, RobotE, Stanlekub, Wikipedia-fox, Jean Sreng, Boréal, Pok148, Thedreamtree, Philagui, RobotQuistnix, YurikBot, Gene.arboit, Eskimbot, Medium69, Jerome66, Jean-Christophe BENOIST, Lyapounov, Chlewbob, Alain.Darles, Boretti, Akiry, Pautard, Sylenius, SashatoBot, Pld, Drust, Soilac, Paf, Fisheye, WartBot, Thijs !bot, Gmt, Rogilbert, RémiH, Deep silence, PierreSelim, JAnDbot, Rhizome, Xiawi, Zawer, Nipou, FlashX, TouristeCatégorisant, PimpBot, Speculos, Zorrobot, Dewal argan, Arouet, VolkovBot, SebYarrette, Orthomaniaque, SieBot, Ambigraphe, MystBot, JLM, Zil, Alecs.bot, Lepsyleon, Kelam, Lilyu, GLec, Estirabot, Expertom, Kolossus, WikiCleanerBot, ZetudBot, Luckas-bot, Micbot, Pgautioneuz, GrouchoBot, Thierry169, Anne Bauval, Simard, Cantons-de-l'Est, Olivier d'ALLIVY KELLY, Xqbot, Norrin TR, Coyote du 57, PAC2, TakniX, Nezdék, EmausBot, Abdelkarim Fourati, Pilteurforichon, ZéroBot, WikitanvirBot, Jules78120, MerlIwBot, Bertol, OrlodrimBot, Haroche, Hosterdam, Mariesyr, Addbot, Foudebassans, Corbel et Anonyme : 49
- **Liste de normes ISO par domaines** *Source* : <http://fr.wikipedia.org/wiki/Liste%20de%20normes%20ISO%20par%20domaines?oldid=113755272> *Contributeurs* : Maggic, BTH, Phe-bot, FredPraca, Jef-Infojef, Bob08, STyx, Gotty, Yanngéffrotin, MMBot, Litlok, Tcharvin, GFDL fan, Pautard, Gonioul, GaMip, PieRRoBoT, Titou42000, Escarbot, Jleroulley, ThbtGrrd, Caristo 2, Analphabot, Salebot, Speculos, Awk, HerculeBot, Pachycephale, Cantons-de-l'Est, Gedman2, Pinji, Jules78120, Agatino Catarella et Anonyme : 16
- **ISO 8601** *Source* : <http://fr.wikipedia.org/wiki/ISO%208601?oldid=106860903> *Contributeurs* : Ryo, CharlesNepote, Vargenau, Fab97, Dirac, HasharBot, Alain Caraco, Serged, Jusjih, Parmenion, Verdy p, Phe, MedBot, TigH, Phe-bot, Domsau2, Tegu, Foxandpotatoes, Dar-

koneko, Clio64, Pixeltoo, Piku, Yurik, Sherbrooke, Idarvol, Peter17, Gzen92, Gpvsobot, DC2, Palpalpal, EDUCA33E, CaptainHaddock, Néfermaât, Tcharvin, Adrian de Lausanne, Bruno1807, Lolando, Barraki, Liquid-aim-bot, Benoît Fabre, Lacrymocéphale, Keindrech, A Pirard, Thijs !bot, Chaoborus, Apc005, Escarbot, Deep silence, Sebleouf, Beeper, Erabot, Numbo3, PimpBot, Rei-bot, Salebot, VolkovBot, Dalb, AmaraBot, Bâmiyân, DaBot, Ange Gabriel, Vlaam, Dhatier, Bub's, GLec, Leontrague, Bastien Sens-Méryé, Alexbot, Mro, JeanHenri79, Mayayu, Belugo, HerculeBot, Luckas-bot, ChenzwBot, Xtremejames183, Xqbot, Marchash, JackBot, Vincent.vaquin, TobeBot, Lola d'AFNOR, JackieBot, WikitanvirBot, Jules78120, Bof-bof, Zebulon84, Vagobot, Orikrin1998, Makecat-bot, Fomtg, Addbot, Heddewin, Kétella, Snowflake Fairy, 61valdhuisne et Anonyme : 42

- **Code lieu-fonction** *Source* : <http://fr.wikipedia.org/wiki/Code%20lieu-fonction?oldid=90351284> *Contributeurs* : BrightRaven, Chtfn, MastiBot, Jules78120 et Addbot
- **Codage des caractères** *Source* : <http://fr.wikipedia.org/wiki/Codage%20des%20caract%C3%A8res?oldid=113119882> *Contributeurs* : Med, Vargenau, Cdag, Robbot, Haypo, Verdy p, BenoîtL, Marc BERTIER, Archibald, Marc Mongenet, MedBot, ChrisJ, Phe-bot, Dromyolo, Bibi Saint-Pol, Mickaël Delahaye, Wikic, Seb35, Zetud, Arnaud.Serander, Gzen92, Gene.arboit, MagnetiK-BoT, Jerome Charles Potts, Pautard, Xofc, MetalGearLiquid, Liquid-aim-bot, Itheo, Jarfe, A2, Circular, JAnDbot, Nono64, Sebleouf, Cbigorgne, Fyergeau, Salebot, Speculos, Zorrobot, Dimdamdom, Critias, Ardu Petus, Synthebot, SieBot, DaBot, Dhatier, AFAccord, Filipvansnaeskerke, Wuyouyuan, Mro, Sacamol, Barakafrit, Luckas-bot, TaBOT-zerem, Xqbot, Loreleil, Vincent.vaquin, Styxyn, Botozor, EmausBot, WikitanvirBot, Jules78120, MerllwBot, Rezabot, Krys74, Rene1596, Soboky, Addbot, Do not follow et Anonyme : 48
- **ISO 8859** *Source* : <http://fr.wikipedia.org/wiki/ISO%208859?oldid=109927739> *Contributeurs* : Med, Ryo, Vargenau, Dirac, Haypo, Yggdras, Verdy p, Helldjinn, Archibald, Phe, MedBot, Sam Hocevar, VIGNERON, Phe-bot, Woww, Romary, Bap, Moyogo, Carlb, Jef-Infojef, Darkoneko, Sbrunner, Pixeltoo, Leag, Bob08, Ork, Celui, Nias, Chobot, Stanlekub, Lmaltier, Kilom691, Pok148, Gzen92, Zwobot, Gpvsobot, FlaBot, Ash Crow, YurikBot, Tcharvin, Loveless, MagnetiK-BoT, Bouktin, MetalGearLiquid, PieRRoBoT, Azerty-Fab, Thijs !bot, Chaoborus, Escarbot, RémiH, Nono64, Sebleouf, Salebot, DodekBot, TXiKiBoT, Ptbotgourou, SieBot, Hxhbot, Vlaam, Dhatier, BenoniBot, GLec, Philippe.petrinko, HerculeBot, ZetudBot, Bub's wikibot, Jean-Louis Hens, Wikivirgile, Xqbot, Rubinbot, Nouilh, Garsd, Midas02, Addbot et Anonyme : 21
- **American Standard Code for Information Interchange** *Source* : <http://fr.wikipedia.org/wiki/American%20Standard%20Code%20for%20Information%20Interchange?oldid=111628995> *Contributeurs* : Aoineko, Olrick, Calo, Valéry Beaud, Christian List, Yann, Ryo, Alvaro, Vargenau, Nataraja, Orthogaffe, Traroth, Céréales Killer, Hephaestos, Ske, Cdag, Herman, Sarrazip, HasharBot, Cham, FoeNyx, Robbot, Haypo, Davux, Verdy p, Archibald, Marc Mongenet, MedBot, Sam Hocevar, VIGNERON, Oblic, Phe-bot, Bibi Saint-Pol, LeTeXnicien de surface, Escaladix, Moyogo, Darkoneko, Sbrunner, MisterMatt, Voodoo, Marcel.c, Leag, Witoki, Koko90, Poulos, Jon207, Emirix, Chobot, Artichaut, Stéphane33, Gribeco, RobotE, Jona, Romanc19s, Lgd, Mandrak, Arnaud.Serander, Cherry, Gzen92, Testtools, Solensean, Coyau, RobotQuistnix, FlaBot, Tavernier, EDUCA33E, Ash Crow, YurikBot, LeonardoRob0t, Zelda, Xavier Combelle Bot, Askywhale, Mi Ga, 16@r, Loveless, MagnetiK-BoT, Le gorille, Grecha, Amine Brikci N, Croquant, Pautard, Dup libre, Torpii, SashatoBot, Maitre So, IP 84.5, MetalGearLiquid, Liquid-aim-bot, Grondin, NicoV, Diti, Vspaceq, Escarbot, JAnDbot, Arkanosis, IALex, Eybot, Rei-bot, Salebot, Loonies, Dimdamdom, DorganBot, AlnoktaBOT, WarddrBOT, TXiKiBoT, Jbvioix, Moa18e, AmaraBot, Lipcare, -Nmd, Synthebot, Fabrice Dury, SieBot, Punx, Abalg, Dsant, Mavati, LordAnubisBOT, Vlaam, François Melchior, Pyerre, Michel421, Bub's, DumZiBoT, Estirabot, Dominiko, Rinaku, Taranzef, Alexbot, Stackas, Darkicebot, Mro, BodhisattvaBot, SilvonenBot, ZetudBot, Lorangeo, LaaknorBot, CarsracBot, Arsaal, LinkFA-Bot, Luckas-bot, Bellatrix Black, Penjo, Joachimpr, Xqbot, Rubinbot, JackBot, Lomita, Orlodrim, DixonDBot, Miacix, TjBot, Toto Azéro, EmausBot, Ediacara, Kilith, Zizomis, Ltrlg, Franz53sda, ZXX 2006, SenseiAC, Oblomov2, OrlodrimBot, DG-IRAO, Vagobot, Automatik, FDo64, Cgsyannick, Reptilien.19831209BE1, DualGoeS-lime, OrikrinBot, Razibot, DiliBot, Lyréa, Addbot, Zebulon84bot et Anonyme : 120
- **Extended Binary Coded Decimal Interchange Code** *Source* : <http://fr.wikipedia.org/wiki/Extended%20Binary%20Coded%20Decimal%20Interchange%20Code?oldid=107747465> *Contributeurs* : Med, David Latapie, Traroth, Dilbert, HasharBot, Haypo, Verdy p, Marc Mongenet, MedBot, Sam Hocevar, Francois Trazzi, Phe-bot, Romainhk, Sherbrooke, Epommate, RobotE, Jfheche, RobotQuistnix, Gpvsobot, FlaBot, EDUCA33E, YurikBot, Loveless, MagnetiK-BoT, Briling, GCloutier, Captainm, Corbeil, anacondabot, Nono64, Rei-bot, Speculos, DodekBot, Isaac Sanolnacov, DorganBot, Masti, SieBot, Laddo, Cépey, Ken123BOT, GLec, HerculeBot, Luckas-bot, RibotBOT, Rubinbot, Lomita, WikitanvirBot, Addbot et Anonyme : 15
- **UTF** *Source* : <http://fr.wikipedia.org/wiki/UTF?oldid=102831607> *Contributeurs* : Ske, NicoV, Bub's, Eric-92, Zizomis et Addbot
- **UTF-8** *Source* : <http://fr.wikipedia.org/wiki/UTF-8?oldid=111789539> *Contributeurs* : Meszigues, Hashar, Calo, Med, COLETTE, Vargenau, Looxix, Popolon, Orthogaffe, DominiqueM, Highlander, Cdag, Romanm, HasharBot, Zubro, Haypo, Davux, Anakin, Caouecs, Verdy p, Alkarex, Archibald, Sanao, Phe, Scullder, Marc Mongenet, MedBot, Sam Hocevar, Xmlizer, Skamp, Phe-bot, Smily, Bibi Saint-Pol, Romainhk, Nocmahr, BernardM, JeanMichel, François-Dominique2, Lyondif02, Doch54, Pixeltoo, Obsidian, Xavier Combelle, YolanC, Wart Dark, Elg, Chobot, LuisMenina, Romanc19s, Arnaud.Serander, Zwobot, RobotQuistnix, FlaBot, YurikBot, MMBot, Litlok, Toutoune25, Pautard, Bech, Jmax, Liquid-aim-bot, Bbruet, Emmanuel Vallois, S.dubourg, RasqualTwilight, KooK, Jarfe, A2, Peetah, JAnDbot, Oliricha, Arkanosis, TomT0m, Nono64, Sebleouf, Verbex, AdQ Bot, Fyergeau, Nicolas.le-guen, Wikig, Dimdamdom, Isaac Sanolnacov, Zakke, Baldodo, TXiKiBoT, Rabatakeu, SuperHeron, Clépey, Love Sun and Dreams, DUPAYSHAUT, Bebeth, Dhatier, DumZiBoT, BodhisattvaBot, Oyp, WikiCleanerBot, ZetudBot, Glide, Herr Satz, Luckas-bot, FrancoisA, Crocxx, TaBOT-zerem, DSisyphBot, Cantons-de-l'Est, Xqbot, Poltuiu, RibotBOT, Schlum, Chauoistes, RedBot, Sbasso, Kilith, Ltrlg, Mastergreg82, Lholivier, Lgfd, Ppff, Addbot, Snowflake Fairy, Lomkimarsh, Gmalauk et Anonyme : 78
- **Adresse IP** *Source* : <http://fr.wikipedia.org/wiki/Adresse%20IP?oldid=113189385> *Contributeurs* : Hashar, R4f, Ryo, Orthogaffe, Céréales Killer, Hémant, Oz, SimonP, Shakti, (:Julien :), Cdag, Herman, HasharBot, Abrahami, R, Raph, Zubro, NucleoS, Jusjih, P-e, Robbot, Haypo, Tieno, Verdy p, Jastrow, Phe, CaptainBeat, Marc Mongenet, MedBot, Oblic, Nico 88, Nyco, Gibux, Xmlizer, Francois Trazzi, Phe-bot, Louis-garden, Eberkut, Deniss, La pinte, JB, Romainhk, Ollamh, Hégésippe Cormier, Jean-no, Lyondif02, Hbbk, Theon, Thesaurus, T, Ork, JujuTh, Poil, MasterMatt, Kornfr, Guiz, Padawane, Eiktee, Anierin, Klemen Kocjancic, Florent(84), Docteur-Cosmos, JKHST65RE23, Chobot, Jorunn, Stanlekub, Taguelmoust, Romanc19s, David Berardan, Lgd, ArséniureDeGallium, Inisheer, EyOne, Slasher-fun, Gzen92, MagnetiK, Liquid 2003, Plyd, Coyau, RobotQuistnix, EDUCA33E, NicolasHairon, Ico, Bortzmeier, MM-Bot, Purodha, Litlok, Misdre, Raphounet, Spungly, MagnetiK-BoT, Et caetera, Shawn, Grecha, Zertrin, Heureux qui comme ulysse, Alain.Darles, Pautard, ObiWan Kenobi, Noar, Cpartiot, DainDwarf, Xofc, MJ, Olmec, Barraki, Serein, François, Malost, Pld, Manu1400, MetalGearLiquid, Moumousse13, Ahbon ?, Liquid-aim-bot, Pingui-King, Guérin Nicolas, PieRRoBoT, NicoV, Thijs !bot, Bourrichon, Flamby, Chaoborus, Grimlock, Jarfe, Escarbot, Albinflo, Axou, Kyle the bot, Laurent Nguyen, RémiH, JeanMarcFontaine, Deep silence, JAnDbot, IALex, Nono64, Sebleouf, Alchemica, Adrille, AdQ Bot, Pamplelune, Rei-bot, Salebot, Akeron, Yann Lejeune, Speculos, Sebf, DorganBot, Idioma-bot, TXiKiBoT, Vifsorbier, VolkovBot, Cdiot, BlueGinkgo, Mikayé, Jeremah76, Mitch-mitch, Chicobot, Synthebot, Ptbotgourou, AlleborgoBot, SieBot, Louperibot, Punx, Enedal, Abalg, Kyro, Deziles, Zil, Dhatier, Croc, Vladoulianov, SniperMaské,

GLec, Alphas, Rabany, DragonBot, Shib Mc Ne, Quentin57, Skippy le Grand Gourou, Chrono1084, Rinaku, Wcorrecor, Superjuju10, Krizald, Mro, David536, HerculeBot, Letartean, SilvonenBot, WikiDreamer Bot, Sylock, Elfix, Kisskat, Nycolas, MakiZen, Morgalion, JackPotte, JeanBono, Nallimbot, ABACA, Manu5891, GrouchoBot, Nakor, Tobyasse, Moipaulochon, Pic-Sou, Tubamirum, Rubinbot, Alex-F, AnneJea, BenzolBot, Omarcoz, KLR FreeLance, Lomita, Xiglofre, TobeBot, Botozor, Super Bazooka, Dinamik-bot, GrrrrBot, AviaWiki, Ripchip Bot, Grominet12345, Frakir, EmausBot, Jejerap, ZéroBot, Mastergreg82, BioPupil, Opendoc, Jules78120, Surdox, Méphisto38, Movses-bot, MerlIwBot, Zebulon84, OrlodrimBot, Pano38, HiW-Bot, Klipe, Minsbot, Mattho69, Chadouming, Haitam122, OrikriBot, Lyréa, Francool50, Addbot, Toreto, Inter photographer, Asimov1966, Cristof32, Do not follow, Antonix, Mathieu11b et Anonyme : 324

- **Port (logiciel)** *Source* : [http://fr.wikipedia.org/wiki/Port%20\(logiciel\)?oldid=113656913](http://fr.wikipedia.org/wiki/Port%20(logiciel)?oldid=113656913) *Contributeurs* : Olrick, R4f, Ryo, Alvaro, Orthogaffe, Oz, Ploum's, AlainBb, Casimir, Sanao, Marc Mongenet, Francois Trazzi, Phe-bot, Bibi Saint-Pol, Cédric, Romainhk, Hbbk, Jef-Infojef, Darkoneko, Nicolas Ray, Smolny, Marcel.c, Piku, ILux, MAX-k, Sherbrooke, Mirgolth, Xfigpower, Romanc19s, Remy34, Dereckson, Slasher-fun, Coyau, Bortzmeier, MMBot, Klipper, Heureux qui comme ulyse, Noar, Manu1400, Rajive, GaMip, PieR-RoBoT, Ig0r, JAnDbot, Jihaim, Salebot, Prométhée33, Crashkoe, Isaac Sanolnacov, SharedX, VolkovBot, Melkor73, -Nmd, Hegurka, SieBot, Fflgt, DumZiBoT, Babssss, Du43l, PixelBot, WikiCleanerBot, ZetudBot, JackPotte, GrouchoBot, Byakuya, Elmander, Jackoc, Xqbot, Laomai Weng, *SM*, Lomita, Eria, WikitanvirBot, MerlIwBot, AvicBot, Pano38, Airatique, Deasyzor, KLBot2, YANN92340, Teabo, Mo5ul, Addbot et Anonyme : 63

7.2 Images

- **Fichier:2011_IBM_Business_Partner_logo.jpg** *Source* : http://upload.wikimedia.org/wikipedia/fr/7/7b/2011_IBM_Business_Partner_logo.jpg *Licence* : marque déposée *Contributeurs* : <http://www.oceanus.co.uk/index.aspx> *Artiste d'origine* : oceanus
- **Fichier:2011_IBM_France_Bois_Colombes.jpg** *Source* : http://upload.wikimedia.org/wikipedia/commons/8/82/2011_IBM_France_Bois_Colombes.jpg *Licence* : GFDL *Contributeurs* : Travail personnel *Artiste d'origine* : Daniel Rodet
- **Fichier:ASCII-Table-wide.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/1/1b/ASCII-Table-wide.svg> *Licence* : Public domain *Contributeurs* :
- **ASCII-Table.svg** *Artiste d'origine* : ASCII-Table.svg : ZZT32
- **Fichier:Adresse_Ipv4.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/c8/Adresse_Ipv4.svg *Licence* : Public domain *Contributeurs* : Indeterminate *Artiste d'origine* : Star Trek Man
- **Fichier:Ambox_globe_content.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/b/bd/Ambox_globe_content.svg *Licence* : Public domain *Contributeurs* : Travail personnel, using File:Information icon3.svg and File:Earth clip art.svg *Artiste d'origine* : penubag
- **Fichier:Applications-internet.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/7/70/Applications-internet.svg> *Licence* : Public domain *Contributeurs* : The Tango! Desktop Project *Artiste d'origine* : The people from the Tango! project
- **Fichier:Baudot_Tape.JPG** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/17/Baudot_Tape.JPG *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Ricardo Ferreira de Oliveira
- **Fichier:Blason_Nord-Pas-De-Calais.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/1c/Blason_Nord-Pas-De-Calais.svg *Licence* : CC BY-SA 3.0 *Contributeurs* : Le code de ce fichier SVG est <a data-x-rel="nofollow" class="external text" href="http://validator.w3.org/check?uri=http%3A%2F%2Fcommons.wikimedia.org%2Fwiki%2Fspecial%3AFilepath%2FBlason_Nord-Pas-De-Calais.svg,&&,,ss=1#source">valide.
Artiste d'origine : Fhiv
- **Fichier:Blason_duche_fr_Savoie.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/33/Blason_duche_fr_Savoie.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Syryatsu
- **Fichier:Blue-punch-card-front-horiz.png** *Source* : <http://upload.wikimedia.org/wikipedia/commons/4/4c/Blue-punch-card-front-horiz.png> *Licence* : Public domain *Contributeurs* :
- **Blue-punch-card-front.png** *Artiste d'origine* : Blue-punch-card-front.png : Gwern
- **Fichier:Collaboration_logo.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/d/d8/Collaboration_logo.svg *Licence* : CC BY-SA 3.0 *Contributeurs* :
- **Wikinews_collaboration_logo_2.svg** *Artiste d'origine* : Wikinews_collaboration_logo_2.svg : Masur
- **Fichier:Communication_shannon-weaver2.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/37/Communication_shannon-weaver2.svg *Licence* : CC-BY-SA-3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Einar Faanes
- **Fichier:Crankshaftrendering.png** *Source* : <http://upload.wikimedia.org/wikipedia/commons/b/b8/Crankshaftrendering.png> *Licence* : CC-BY-SA-3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Boroski,
- **Fichier:Crypto_key.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/5/5b/Crypto_key.png *Licence* : LGPL *Contributeurs* : Originally from fr.wikipedia ; description page is/was here. *Artiste d'origine* : Original uploader was Dake at fr.wikipedia Later versions were uploaded by Croquant at fr.wikipedia.
- **Fichier:Crystal_Clear_app_linneighborhood.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/d/d0/Crystal_Clear_app_linneighborhood.png *Licence* : LGPL *Contributeurs* : All Crystal icons were posted by the author as LGPL on kde-look *Artiste d'origine* : Everaldo Coelho and YellowIcon

- **Fichier:Crystal_energy.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/14/Crystal_energy.svg *Licence* : LGPL *Contributeurs* : Own work conversion of Image:Crystal_128_energy.png *Artiste d'origine* : Dhatfield
 - **Fichier:Crystal_kpackage.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/4/40/Crystal_kpackage.png *Licence* : LGPL *Contributeurs* : All Crystal icons were posted by the author as LGPL on kde-look *Artiste d'origine* : Everaldo Coelho (YellowIcon)
 - **Fichier:Crystal_mycomputer.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/e/e3/Crystal_mycomputer.png *Licence* : LGPL *Contributeurs* : All Crystal icons were posted by the author as LGPL on kde-look *Artiste d'origine* : Everaldo Coelho (YellowIcon)
 - **Fichier:Decrease2.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/e/ed/Decrease2.svg> *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Sarang
 - **Fichier:Disambig_colour.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/3e/Disambig_colour.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Bub's
 - **Fichier:Emblem-money.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/f/f3/Emblem-money.svg> *Licence* : GPL *Contributeurs* : <http://www.gnome-look.org/content/show.php/GNOME-colors?content=82562> *Artiste d'origine* : perfectska04
 - **Fichier:FTP_mode_actif.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/2/29/FTP_mode_actif.png *Licence* : CC BY-SA 3.0 *Contributeurs* : Dérivé par travail personnel de "Diagramme des flux de FTP.png" *Artiste d'origine* : MathsPoetry
 - **Fichier:FTP_mode_passif.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/c5/FTP_mode_passif.png *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel à partir de "Diagramme des Flux de FTP.png" *Artiste d'origine* : MathosPoetry
 - **Fichier:Factory.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/a/a2/Factory.svg> *Licence* : Public domain *Contributeurs* : Self-made, taken from Image : 1 9 2 9.svg *Artiste d'origine* : Howard Cheng
 - **Fichier:Farm-Fresh_curriculum_vitae.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/4/46/Farm-Fresh_curriculum_vitae.png *Licence* : CC BY 3.0 us *Contributeurs* : <http://www.fatcow.com/free-icons/> *Artiste d'origine* : FatCow Web Hosting
 - **Fichier:Flag_of_France.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/c3/Flag_of_France.svg *Licence* : Public domain *Contributeurs* : http://www.diplomatie.gouv.fr/de/frankreich_3/frankreich-entdecken_244/portrat-frankreichs_247/die-symbole-der-franzosischen-republ-260/trikolore-die-nationalfahne_114.html *Artiste d'origine* : This graphic was drawn by SKopp.
 - **Fichier:Flag_of_Germany.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/b/ba/Flag_of_Germany.svg *Licence* : Public domain *Contributeurs* : ? *Artiste d'origine* : ?
 - **Fichier:Flag_of_Italy.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/0/03/Flag_of_Italy.svg *Licence* : Public domain *Contributeurs* : There has been a long discussion on the colors of this flag. *Please read the talk page before editing or reverting this image.* Pantone to RGB performed by <http://www.pantone.com/pages/pantone/colorfinder.aspx> *Artiste d'origine* : see below
 - **Fichier:Flag_of_Romania.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/7/73/Flag_of_Romania.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : AdiJapan
 - **Fichier:Flag_of_Russia.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/f/f3/Flag_of_Russia.svg *Licence* : Public domain *Contributeurs* : Государственный флаг Российской Федерации. Цвета флага : (Blue - Pantone 286 C, Red - Pantone 485 C) взяты из [1][2][3][4] *Artiste d'origine* : Zscout370
 - **Fichier:Flag_of_San_Francisco.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/a/af/Flag_of_San_Francisco.svg *Licence* : Public domain *Contributeurs* : <http://vector-images.com/image.php?epsid=6058> *Artiste d'origine* :
 - Converted from EPS to SVG by : Oren neu dag
 - **Fichier:Flag_of_Spain.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/9/9a/Flag_of_Spain.svg *Licence* : CC0 *Contributeurs* : ["Sodipodi.com Clipart Gallery". Original link no longer available] *Artiste d'origine* : Pedro A. Gracia Fajardo, escudo de Manual de Imagen Institucional de la Administración General del Estado
 - **Fichier:Flag_of_Sweden.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/4/4c/Flag_of_Sweden.svg *Licence* : Public domain *Contributeurs* : ? *Artiste d'origine* : ?
 - **Fichier:Flag_of_Switzerland.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/f/f3/Flag_of_Switzerland.svg *Licence* : Public domain *Contributeurs* : PDF Colors Construction sheet *Artiste d'origine* : User:Marc Mongenet
- Credits :
- **Fichier:Flag_of_Turkey.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/b/b4/Flag_of_Turkey.svg *Licence* : Public domain *Contributeurs* : Turkish Flag Law (Türk Bayrağı Kanunu), Law nr. 2893 of 22 September 1983. Text (in Turkish) at the website of the Turkish Historical Society (Türk Tarih Kurumu) *Artiste d'origine* : David Benbennick (original author)
 - **Fichier:Flag_of_the_Netherlands.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/2/20/Flag_of_the_Netherlands.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Zscout370
 - **Fichier:Flag_of_the_United_Kingdom.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/a/ae/Flag_of_the_United_Kingdom.svg *Licence* : Public domain *Contributeurs* : Travail personnel per data at <http://flagspot.net/flags/gb.html> *Artiste d'origine* : Original flag by Acts of Union 1800
 - **Fichier:Flag_of_the_United_States.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/a/a4/Flag_of_the_United_States.svg *Licence* : Public domain *Contributeurs* : SVG implementation of U. S. Code : Title 4, Chapter 1, Section 1 [1] (the United States Federal "Flag Law"). *Artiste d'origine* : Dbenbenn, Zscout370, Jacobolus, Indolences, Technon.
 - **Fichier:Generix.jpg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/d/db/Generix.jpg> *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Katanja
 - **Fichier:Gtk-dialog-info.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/b/b4/Gtk-dialog-info.svg> *Licence* : LGPL *Contributeurs* : <http://ftp.gnome.org/pub/GNOME/sources/gnome-themes-extras/0.9/gnome-themes-extras-0.9.0.tar.gz>. *Artiste d'origine* : David Vignoni

- **Fichier:Hollerith_punched_card.jpg** *Source* : http://upload.wikimedia.org/wikipedia/commons/f/f2/Hollerith_punched_card.jpg *Licence* : Public domain *Contributeurs* : Railroad Gazette *Artiste d'origine* : Herman Hollerith
- **Fichier:Hourglass_drawing.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/c2/Hourglass_drawing.svg *Licence* : Public domain *Contributeurs* : ? *Artiste d'origine* : ?
- **Fichier:IBM-Denia.JPG** *Source* : http://upload.wikimedia.org/wikipedia/commons/2/2b/IBM_Denia.jpg *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Orrling
- **Fichier:IBM2314DiskDrivesAndIBM2540CardReaderPunch.jpg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/c/c3/IBM2314DiskDrivesAnd.jpg> *Licence* : Public domain *Contributeurs* : Scott Gerstenberger *Artiste d'origine* : Scott Gerstenberger
- **Fichier:ICS_November.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/2/21/ICS_November.svg *Licence* : Public domain *Contributeurs* : ? *Artiste d'origine* : ?
- **Fichier:ICS_Oscar.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/c9/ICS_Oscar.svg *Licence* : Public domain *Contributeurs* : ? *Artiste d'origine* : ?
- **Fichier:ICS_Repeat_One.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/6/61/ICS_Repeat_One.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Created by Denelson83
- **Fichier:IPv4_blocs.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/1b/IPv4_blocs.svg *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Mro
- **Fichier:Increase2.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/b/b0/Increase2.svg> *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Sarang
- **Fichier:Information_icon.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/35/Information_icon.svg *Licence* : Public domain *Contributeurs* : en:Image:Information icon.svg *Artiste d'origine* : El T
- **Fichier:Ipv6_address.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/15/Ipv6_address.svg *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Indeterminate
- **Fichier:MM1PaperTapeEncoding.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/7/7b/MM1PaperTapeEncoding.svg> *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Malleus Fatuorum
- **Fichier:Market.png** *Source* : <http://upload.wikimedia.org/wikipedia/commons/9/9f/Market.png> *Licence* : CC-BY-SA-3.0 *Contributeurs* : ? *Artiste d'origine* : ?
- **Fichier:Max-cut.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/c/cf/Max-cut.svg> *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Miym
- **Fichier:Nuvola_apps_kgpg.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/a/a2/Nuvola_apps_kgpg.png *Licence* : LGPL *Contributeurs* : <http://icon-king.com> *Artiste d'origine* : David Vignoni / ICON KING
- **Fichier:Nuvola_apps_ksim.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/8/8d/Nuvola_apps_ksim.png *Licence* : LGPL *Contributeurs* : <http://icon-king.com> *Artiste d'origine* : David Vignoni / ICON KING
- **Fichier:Nuvola_apps_package_wordprocessing.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/36/Nuvola_apps_package_wordprocessing.png *Licence* : LGPL *Contributeurs* : <http://icon-king.com> *Artiste d'origine* : David Vignoni / ICON KING
- **Fichier:P_parthenon.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/f/f4/P_parthenon.svg *Licence* : CC-BY-SA-3.0 *Contributeurs* : ? *Artiste d'origine* : ?
- **Fichier:Passive_FTP_Verbindung.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/8/84/Passive_FTP_Verbindung.svg *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Michael Lorer
- **Fichier:Punched_card.jpg** *Source* : http://upload.wikimedia.org/wikipedia/commons/f/f3/Punched_card.jpg *Licence* : CC BY 2.5 *Contributeurs* : Travail personnel *Artiste d'origine* : Mutatis mutandis
- **Fichier:Question_book-4.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/6/64/Question_book-4.svg *Licence* : CC-BY-SA-3.0 *Contributeurs* : Created from scratch in Adobe Illustrator. Originally based on Image:Question book.png created by User:Equazcion. *Artiste d'origine* : Tkgd2007
- **Fichier:Racine_carrée_bleue.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/1/1f/Racine_carr%C3%A9e_bleue.svg *Licence* : LGPL *Contributeurs* : Image:Nuvola apps edu mathematics-p.svg *Artiste d'origine* : historicair 17 :50, 4 June 2007 (UTC)
- **Fichier:SMTP-transfer-model.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/6/69/SMTP-transfer-model.svg> *Licence* : CC BY-SA 3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Ale2006-from-en
- **Fichier:Server-based-network.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/f/fb/Server-based-network.svg> *Licence* : LGPL *Contributeurs* : derived from the Image:Computer n screen.svg which is under the GNU LGPL *Artiste d'origine* : User:Mauro Bieg
- **Fichier:Sierpinski_arrowhead_3d_stage_5.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/3a/Sierpinski_arrowhead_3d_stage_5.png *Licence* : CC BY-SA 3.0 *Contributeurs* : Own work, created with Mathematica 6 *Artiste d'origine* : Robert Dickau
- **Fichier:Tangopub.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/7/78/Tangopub.svg> *Licence* : CC BY-SA 2.5 *Contributeurs* : Image:Software-update-available.svg + self made *Artiste d'origine* : Garfieldairlines & the tango desktop team
- **Fichier:TheStructorr_Lamborghini_Gallardo.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/8/8a/TheStructorr_Lamborghini_Gallardo.svg *Licence* : CC0 *Contributeurs* : OpenClipArt.org *Artiste d'origine* : Michał Pecyna
- **Fichier:To_validate.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/c/ca/To_validate.svg *Licence* : LGPL *Contributeurs* : Image:Nuvola apps korganizer.svg, Image:Nuvola apps error alt.svg, Image:Sdm.svg *Artiste d'origine* : en:David Vignoni, User:Stannered
- **Fichier:Transmission.png** *Source* : <http://upload.wikimedia.org/wikipedia/commons/d/d1/Transmission.png> *Licence* : MIT *Contributeurs* : Transferred from it.wikipedia; transfer was stated to be made by User:Vajotwo. *Artiste d'origine* : Original uploader was Krdan at it.wikipedia

- **Fichier:TwoCoins.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/2/29/ThreeCoins.svg> *Licence* : Public domain *Contributeurs* : Originally from en.wikipedia ; description page is/was here. User:Ysangkok added shadows and silhouettes from Image:Lars_Gustaf_Tersmeden.svg, Image:Caspar Friedrich Wolff.svg and Image:Sieveking-Silhouette.svg. *Artiste d'origine* : Original uploader was Busy Stubber at en.wikipedia, effects : User:Ysangkok
- **Fichier:Wiki_letter_w.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/6/6c/Wiki_letter_w.svg *Licence* : CC-BY-SA-3.0 *Contributeurs* : Travail personnel *Artiste d'origine* : Jarkko Piironen
- **Fichier:X25-network-diagram-0a.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/5/5c/X25-network-diagram-0a.svg> *Licence* : Public domain *Contributeurs* : Travail personnel *Artiste d'origine* : Adamantios
- **Fichier:_Crystal_Clear_action_run.svg** *Source* : http://upload.wikimedia.org/wikipedia/commons/3/35/Crystal_Clear_action_run.svg *Licence* : Public domain *Contributeurs* : All Crystal icons were posted by the author as LGPL on kde-look *Artiste d'origine* : Everaldo Coelho and YellowIcon Vectorizer : User:HereToHelp
- **Fichier:_Crystal_mycomputer.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/e/e3/Crystal_mycomputer.png *Licence* : LGPL *Contributeurs* : All Crystal icons were posted by the author as LGPL on kde-look *Artiste d'origine* : Everaldo Coelho (YellowIcon)
- **Fichier:_Emblem-money.svg** *Source* : <http://upload.wikimedia.org/wikipedia/commons/f/f3/Emblem-money.svg> *Licence* : GPL *Contributeurs* : <http://www.gnome-look.org/content/show.php/GNOME-colors?content=82562> *Artiste d'origine* : perfectska04
- **Fichier:_Nuvola_apps_kgpg.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/a/a2/Nuvola_apps_kgpg.png *Licence* : LGPL *Contributeurs* : <http://icon-king.com> *Artiste d'origine* : David Vignoni / ICON KING
- **Fichier:_Nuvola_apps_password.png** *Source* : http://upload.wikimedia.org/wikipedia/commons/7/7c/Nuvola_apps_password.png *Licence* : LGPL *Contributeurs* : <http://icon-king.com> *Artiste d'origine* : David Vignoni / ICON KING
- **Fichier:_Ébaucheass.gif** *Source* : <http://upload.wikimedia.org/wikipedia/commons/b/b5/%C3%89baucheass.gif> *Licence* : Public domain *Contributeurs* : Transferred from fr.wikipedia ; transferred to Commons by User:Bloody-libu using CommonsHelper. *Artiste d'origine* : Original uploader was Liquid 2003 at fr.wikipedia

7.3 Licence du contenu

- Creative Commons Attribution-Share Alike 3.0