



Un leader reconnu mondialement dans le domaine des **logiciels de traduction**, des **dictionnaires électroniques**...

+ de 3 millions d'utilisateurs

+ de 30 millions de traductions par mois

5 paires de langues principales, plus de 200 au total

Des centaines de milliers de mots et expressions, du plus général au plus spécialisé



7 fausses idées sur les outils de traduction

Et en particulier...

...sur le besoin des utilisateurs :

- nous parlons parfaitement anglais
- notre niveau d'anglais est faible mais suffisant

...sur les outils eux-mêmes :

- notre domaine est trop technique → pas adaptés
- ils sont trop imparfaits pour être utiles
- c'est du mot à mot
- les sites gratuits sur le web suffisent
- un dictionnaire papier suffit



7 fausses idées...sur le besoin des utilisateurs (1)

« nous parlons parfaitement anglais »

« notre niveau d'anglais est faible mais suffisant »

Texte source :

- They hammered the market and stocks plummeted.
- We must stop the band-aid approach and get into a root-and-branch reform.
- The robber barons socked him, as he tried to put an end to moonlighting and the black economy.
- They cast lots, as they could not reach an agreement by other means.





Texte source / Texte cible :

- They hammered the market and stocks plummeted.
Ils ont fait baisser les cours et les actions sont tombées.
- We must stop the band-aid approach and get into a root-and-branch reform.
Nous devons arrêter le rafistolage et entrer dans une réforme radicale.
- The robber barons socked him, as he tried to put an end to moonlighting and the black economy.
Les requins de la finance l'ont frappé, comme il a essayé de mettre fin au travail au noir et l'économie parallèle.
- They cast lots, as they could not reach an agreement by other means.
Ils tirent au sort, comme ils ne pouvaient pas atteindre un accord par d'autres moyens.



7 fausses idées...sur les outils eux-mêmes (1)

« *notre domaine est trop technique* »

Texte source (Areva) :

« No forced draught ventilation ducts on the closure head to cool the CRDM coils. No integral antimissile slab. »

Traduction Reverso :

« Aucune gaine de ventilation forcée sur le couvercle de cuve pour refroidir les bobines de commande des mécanismes. Aucune dalle antimissile intégrée. »



7 fausses idées...sur les outils eux-mêmes (2)

« ils sont trop imparfaits pour être utiles »

- Texte en russe :

« Банковский кризис в России разрастается. »

La crise bancaire en Russie grandit.

- Texte en chinois :



Je pars demain matin. Je reviens dans une semaine.



7 fausses idées...sur les outils eux-mêmes (3)

« *c'est du mot à mot* »

- The person I talked to is right here.
La personne à qui j'ai parlé est juste ici.
- I'm impressed by the software I was shown.
Je suis impressionné par le logiciel que l'on m'a montré.
- Should you need a car, let me know !
Si vous avez besoin d'une voiture, faites-le moi savoir !



7 fausses idées...sur les outils eux-mêmes (4)

« *les sites gratuits sur le web suffisent* »

Texte source:

Il est important de prendre en compte la plus-value de cession exceptionnelle de 91,9 millions d'euros réalisée sur le 1^{er} semestre »

=> Reverso.net (site web): It is important to take into account the surplus of exceptional 91,9 million euro transfer realized on the 1st half of the year

=> Reverso Intranet : It is important to take into account the 91,9 million euro extraordinary capital gain realized on the 1st half of the year



7 fausses idées...sur les outils eux-mêmes (4bis)

« les sites gratuits sur le web suffisent »

=> Limites

- Taille de documents, formats de fichiers limités
- Confidentialité
- Précision de la traduction

=> Apports de Reverso Intranet

- Traduction de fichiers Word en conservant la mise en page
- Gestion de différents formats (PDF, Excel,...)
- Adaptation au vocabulaire, à la charte de l'entreprise
- Dictionnaires spécialisés (business, technique, médical...)
- Intégration dans les applications MS Office



Intégration d'une barre d'outil dans Internet Explorer...
...et traduction dans une info-bulle ou dans une fenêtre pop-up

Adresse http://money.cnn.com/2004/06/03/markets/oil_jobs/index.htm

mywebsearch >> Reverso Anglais-Français EFbusiness06

Markets & Stocks

SAVE THIS EMAIL THIS PRINT THIS MOST POPULAR

Oil vs. jobs

The stock market's obsession with oil is distracting it from what's really important -- jobs.

June 3, 2004: 6:24 PM EDT
By [Mark Gongloff](#), CNN/Money senior writer

NEW YORK (CNN/Money) - U.S. stock markets have been

Pétrole(Huile) contre emplois
L'obsession du marché boursier avec le pétrole(l'huile) le distrait de ce qui est vraiment important - des emplois.



7 fausses idées...sur les outils eux-mêmes (5)

« un dictionnaire papier suffit »

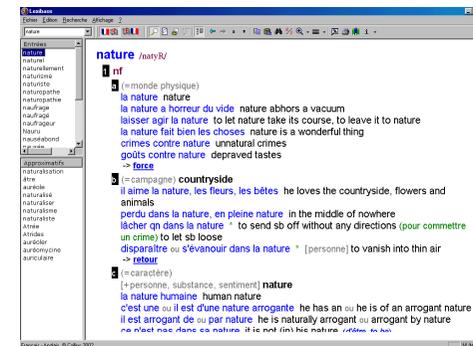
=> Limites:

- Encombrant
- Peu lisible



=> Apports des dictionnaires électroniques:

- Affichage paramétrable
- Ergonomie et lisibilité
- Fonctions de recherches poussées
- Passage d'un dictionnaire à l'autre en 1 clic de souris
- Utilisation depuis vos applications





Recherche

mettre

Entrées

- mettre
- meublant
- meuble
- meublé
- meubler
- meugler
- meule
- meunerie
- meunier
- meurs {etc}
- meurtre
- meurtrier
- meurtrière
- meurtrir
- meurtrissure
- meus {etc}
- Meuse

Suggestions

- mets
- met
- se mettre
- maître
- mètre
- maître à penser
- maître auxiliaire
- maître chanteur
- maître d'armes
- mètre carré/cube

mettre /mɛtʁ/

vt

1 ponere

mettre en bouteille embotellar
 mettre en sac poner en sacos
 mettre en pages compaginar
 mettre qch en terre enterrar algo
 mettre en examen detener (para ser interrogado)
 mettre à la poste echar al correo
 mettre qn debout/assis levantar/sentar a algn

2 (vêtements, revêtir) ponere, (soi-même) ponerse, (installer) ponere

mets ton gilet ponte el chaleco

3 (faire fonctionner, chauffage, réveil) ponere, (lumière) dar, (installer, gaz, e

faire mettre le gaz/l'électricité poner gas/electricidad
 mettre en marche poner en marcha

4 (consacrer)

mettre du temps/2 heures à faire qch tardar tiempo/dos horas en hace

5 (noter, écrire) ponere

qu'est-ce que tu as mis sur la carte? ¿qué has puesto en la postal?
 mettre au pluriel poner en plural

6 (supposer)

mettons que ... pongamos que ...

7 y mettre du sien (dépenser dans une affaire) ponere de su parte

Recherche du mot « mettre » en espagnol



Collins Fr-En Essential

Recherche

Entrées Résultats de la recherche plein texte Page 1/1 12 Résultats

Exemples et phrases idiomatiques

vigueur	être/entrer en vigueur	to be in/come into force
vigueur	en vigueur	current

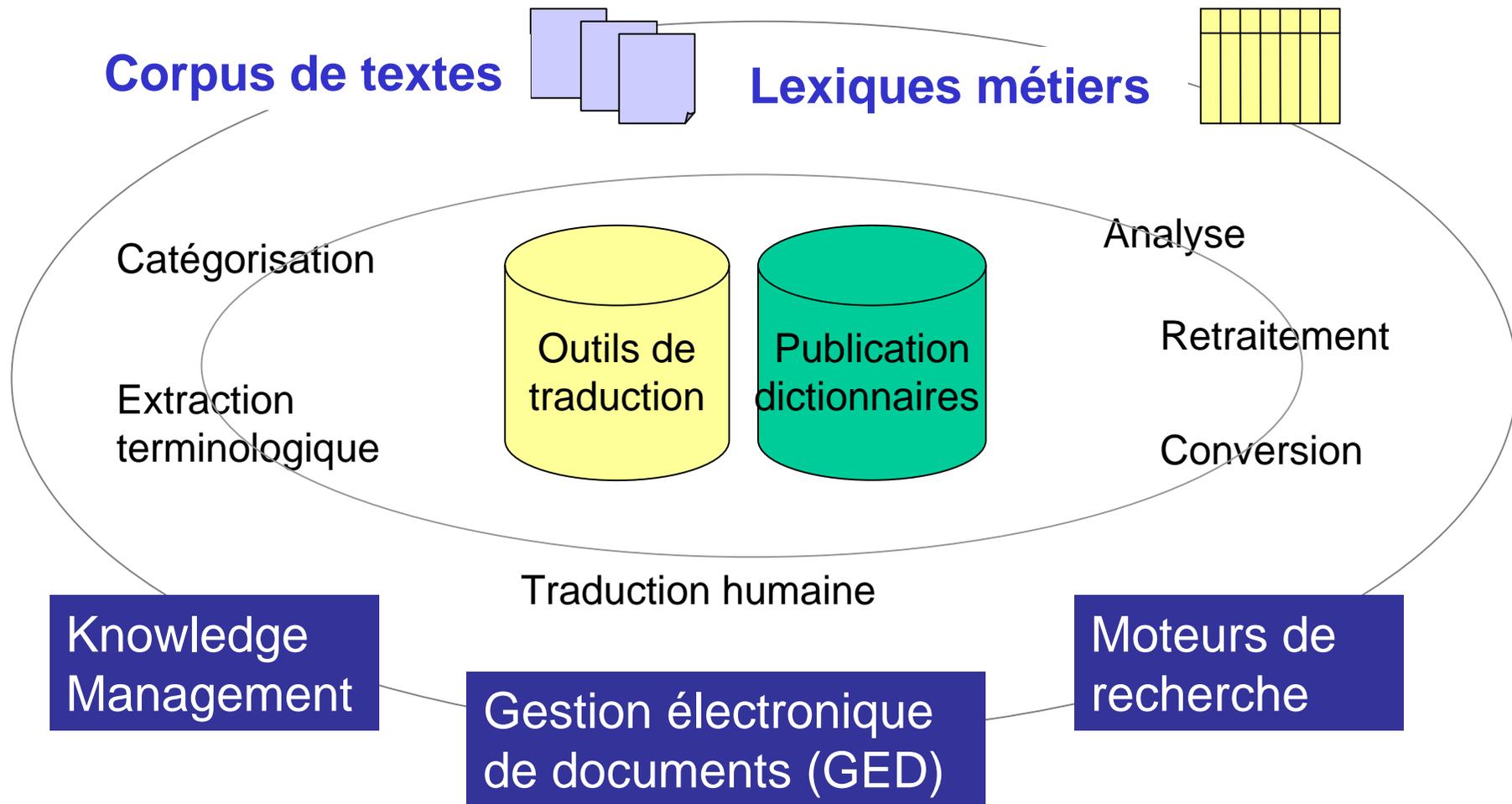
Traductions / Définitions

applicable	the law is applicable from January	la loi entre en vigueur au mois de janvier
effect	to take effect	(law) entrer en vigueur
effective	to become effective	(law) entrer en vigueur
effective	effective date	date f d'effet or d'entrée en vigueur
force	to come into force	entrer en vigueur
going	the going rate	le tarif (en vigueur)
lapse	lapse (law)	cesser d'être en vigueur
operation	to be in operation	[machine, system] être en vigueur
operative	operative	[measure] en vigueur
run	to run sth over	(=pass) [road etc] (=work) [machine, factory, bus, train] (=continue) [play, contract] être valide or en vigueur

Recherche plein texte
du mot « en vigueur »
en anglais
 recherche élargie à
 l'ensemble du dictionnaire
 (exemples, phrases
 idiomatiques, contextes,...)



Plateforme terminologique Softissimo: Enrichir/Stocker/Publier





Une application concrète à forte valeur ajoutée:
l'intégration dans un moteur de recherche
ou
« cross language information retrieval »

- Faire une recherche dans une langue sur un corpus de documents multilingues
- Exploiter les résultats sans maîtriser la langue du document

Verity and Softissimo Integration - Search and Translate - Microsoft Internet Explorer

Fichier Edition Affichage Favoris Outils ? Adresse http://verity.reverso.net/main.asp

Précédente Rechercher Favoris Média 37 bloquée(s) Options Lexibase Collins Fr-En Essential [An-Fr]

Welcome to Verity Softissimo website

1) Requête en anglais

This site will let you submit a query in English and search the search to documents in other languages, and eventually translate results back to English.

Type your query (in English) input device Search

Search sites in:

EN FR SP DE Reverso Verity

Number of results per page: 10 Sort by: language relevance

2) Choix de langue de la requête multilingue

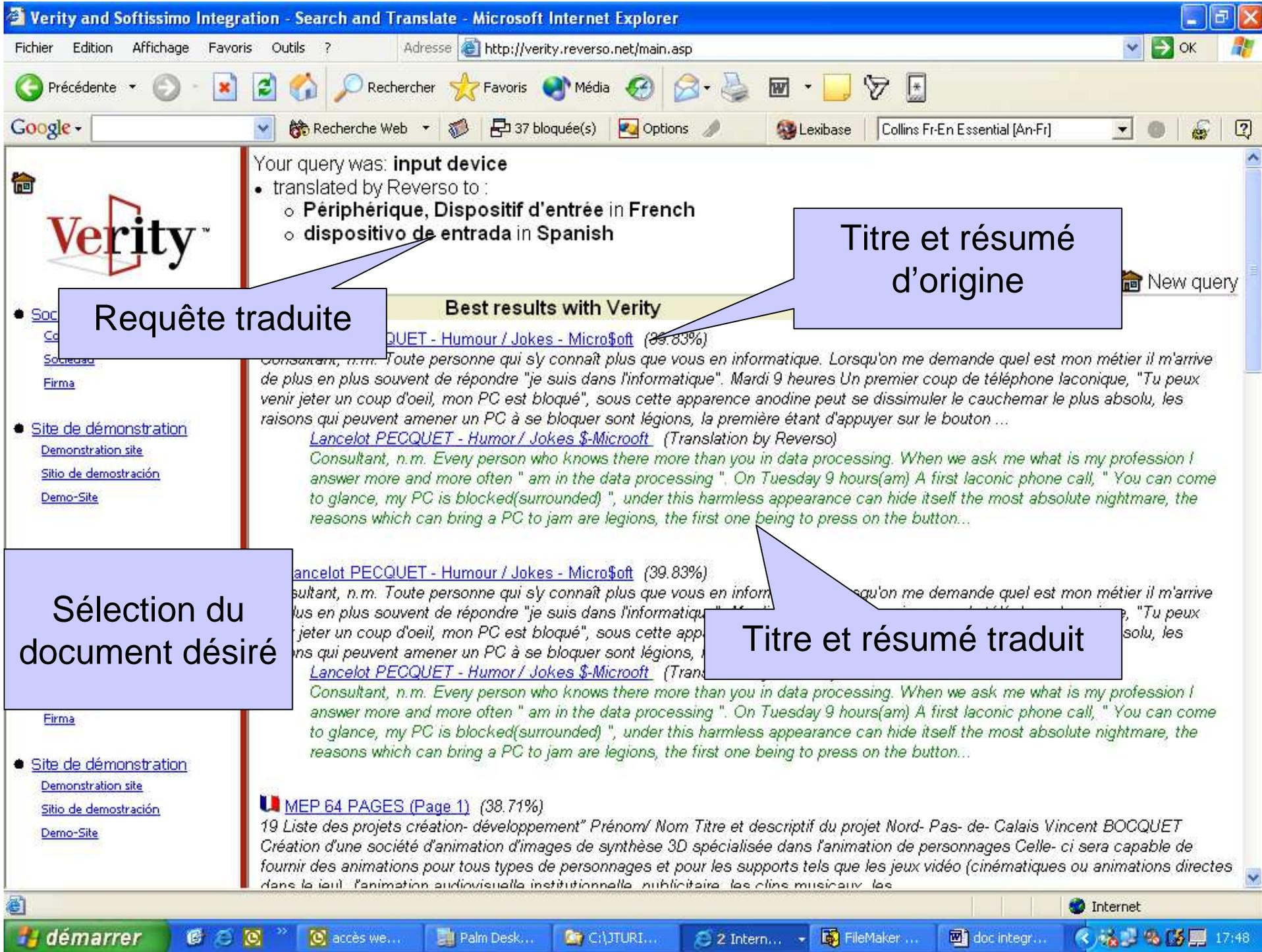
3) Ordre de tri

Some examples of Cross Language Information Retrieval:

- [New York](#) (in English, French and Spanish)
- [input device](#) (in English, French and Spanish)

Verity K2 Enterprise Reverso

démarrer accès we... Palm Desk... C:\JTURI... 2 Intern... FileMaker ... doc integr... 17:47



Verity™

- [Site de démonstration](#)
[Demonstration site](#)
[Sitio de demostración](#)
[Demo-Site](#)

Your query was: **input device**

- translated by Reverso to :
 - **Périphérique, Dispositif d'entrée** in French
 - **dispositivo de entrada** in Spanish

Requête traduite

Titre et résumé d'origine

Sélection du document désiré

Best results with Verity

[Lancelot PECQUET - Humour / Jokes - Micro\\$oft](#) (39.83%)
Consultant, n.m. Toute personne qui s'y connaît plus que vous en informatique. Lorsqu'on me demande quel est mon métier il m'arrive de plus en plus souvent de répondre "je suis dans l'informatique". Mardi 9 heures Un premier coup de téléphone laconique, "Tu peux venir jeter un coup d'oeil, mon PC est bloqué", sous cette apparence anodine peut se dissimuler le cauchemar le plus absolu, les raisons qui peuvent amener un PC à se bloquer sont légions, la première étant d'appuyer sur le bouton ...

[Lancelot PECQUET - Humor / Jokes \\$-Microoft](#) (Translation by Reverso)
Consultant, n.m. Every person who knows there more than you in data processing. When we ask me what is my profession I answer more and more often " am in the data processing ". On Tuesday 9 hours(am) A first laconic phone call, " You can come to glance, my PC is blocked(surrounded) ", under this harmless appearance can hide itself the most absolute nightmare, the reasons which can bring a PC to jam are legions, the first one being to press on the button...

[Lancelot PECQUET - Humour / Jokes - Micro\\$oft](#) (39.83%)
Consultant, n.m. Toute personne qui s'y connaît plus que vous en informatique. Lorsqu'on me demande quel est mon métier il m'arrive de plus en plus souvent de répondre "je suis dans l'informatique". Mardi 9 heures Un premier coup de téléphone laconique, "Tu peux venir jeter un coup d'oeil, mon PC est bloqué", sous cette apparence anodine peut se dissimuler le cauchemar le plus absolu, les raisons qui peuvent amener un PC à se bloquer sont légions, la première étant d'appuyer sur le bouton ...

[Lancelot PECQUET - Humor / Jokes \\$-Microoft](#) (Translation by Reverso)
Consultant, n.m. Every person who knows there more than you in data processing. When we ask me what is my profession I answer more and more often " am in the data processing ". On Tuesday 9 hours(am) A first laconic phone call, " You can come to glance, my PC is blocked(surrounded) ", under this harmless appearance can hide itself the most absolute nightmare, the reasons which can bring a PC to jam are legions, the first one being to press on the button...

[MEP 64 PAGES \(Page 1\)](#) (38.71%)
 19 Liste des projets création- développement" Prénom/ Nom Titre et descriptif du projet Nord- Pas- de- Calais Vincent BOCQUET Création d'une société d'animation d'images de synthèse 3D spécialisée dans l'animation de personnages Celle- ci sera capable de fournir des animations pour tous types de personnages et pour les supports tels que les jeux vidéo (cinématiques ou animations directes dans le jeu), l'animation audiovisuelle institutionnelle publicitaire, les clips musicaux, les

Titre et résumé traduit



- [Société](#)
 - [Company](#)
 - [Sociedad](#)
 - [Firma](#)
- [Site de démonstration](#)
 - [Demonstration site](#)
 - [Sitio de demostración](#)
 - [Demo-Site](#)



- [Société](#)
 - [Company](#)
 - [Sociedad](#)
 - [Firma](#)
- [Site de démonstration](#)
 - [Demonstration site](#)
 - [Sitio de demostración](#)
 - [Demo-Site](#)



**Avant de craquer...
Comparez !**

ACHETER COMPARER



Encyclopedia Reception Look for

The Social Security

- [The French system](#)
 - [The various regimes](#)
 - [The services\(performances\)](#)
 - [The state pension](#)
 - [The welfare](#)
 - [The unemployment insurance](#)
- [The management of the French system](#)
 - [The problem of the deficit](#)
- [The social welfare: choice of company\(social security\)](#)
- [The international agreements](#)

References [Legend](#)

- [Associated subjects](#)
- [See also...](#)
- [Search\(Research\) for Web sites](#)

Affichage du document traduit à la volée

The French system

The law of May 22nd, 1946 proclaimed the generalization of the Social Security in all the categories of the population; in spite of the intentions of this law, the process was long and often difficult (certain liberal occupations, notably, refused what they considered as their "fonctionnarisation"). It was necessary at the same moment to define the conditions of access of the non-employees and to take into account particular advantages acquired before by certain occupations (for example by the minors(miners)).

The various regimes



CONCLUSION

Les applications de nos technologies sont vastes* et nos outils comptent toujours plus d'utilisateurs satisfaits...même si les critiques auront toujours matière à critiquer.

L'implication de nos clients est le plus grand facteur de réussite de nos projets !

*il nous reste à construire un dictionnaire sur la culture des mangues aux Galapagos ou un traducteur homme/femme, enfants/parents, langage SMS/ langue française,...

■ TEXT MINING SOLUTIONS



TEMIS Solutions de Text Mining

■ TEXT MINING SOLUTIONS

Guillaume Mazieres

Vice-président Sales and Marketing



business & technologie
CLUB PROFESSIONNEL EDHEC

INTERMINES
PARIS - SAINT-ETIENNE - RANCY



Do you speak business? - Février 2005

AGENDA 7 FEVRIER

1. **Présentation TEMIS**
2. **Introduction au TEXT MINING**
3. **Solutions technologiques**
4. **Cas Clients**
 1. **TOTAL**
 2. **PSA**
5. **Questions / réponses**



QUI EST TEMIS?

- Contraction de **TE**xt **MI**ning **S**olutions, TEMIS est un éditeur de logiciels créé en septembre 2000.
- Chef de file européen du **Text Mining**, présent en France, Allemagne, Italie, Angleterre et au travers de ses partenaires dans le reste de l'Europe et aux États-Unis.
- 50 personnes, 5 localisations.
- **Équipe de chercheurs, développeurs et dirigeants** issus d'**IBM** France, USA, Allemagne et Italie.
- Acquisition de **Xerox Linguistics** en Juillet 2003.
- **Maturité technologique** d'une équipe qui dispose de plus de 15 ans d'expérience dans le domaine du Text Mining.



GET 15 COMMISSION-FREE TRADES
 when you open an Ameritrade® account.

PLUS GET \$100 CASH BONUS IF YOU QUALIFY FOR APEX. **GO>**

A Reuters Service
multexinvestor®
 financial research and information

[Help](#) [Login](#) [Join](#) [Free Offers](#)

- Quick Info
- Stock Overview
- Quote
- Chart
- Key Developments
- News**
- Company Profile
- Financial Info
- Snapshot
- Growth
- Ratios
- Financial Statements
- Profile Report **NEW!**
- Market Sentiment
- Performance
- Shares Shorted
- Insider Trading
- Institutional Holders
- Professional Analysis
- Earnings Estimates
- Recommendations
- Research Reports

News

sponsored by **\$7 TRADES**

XEROX CORPORATION (NYSE)
 Sector: [Technology](#) Industry: [Office Equipment](#)

- 7 Stocks to Buy Now!
- \$7 Trades, NO Fee IRAs

Headlines List



September 04, 2003 05:35:04 AM ET

Sep. 04 | M2 Communications

Sep 04, 2003 (TELECOMWORLDWIRE via COMTEX) --Text mining software tools provider TEMIS has acquired the linguistic product operations of Xerox Corp, a digital imaging company.

Under the terms of the agreement TEMIS will obtain the business operations that produce the XeLDA and XTS software products. Both of these offerings will be incorporated into TEMIS' existing product portfolio. TEMIS will also take on the nine staff members that worked in the

7 Stocks to buy now
[Click here for info.](#)

QUI EST TEMIS?

- ***TEMIS, Europe's leading Text Mining company opens for business in the United States***

Paris, November 23, 2004 – TEMIS, Europe's leading provider of Text Mining solutions has opened a sales office in Washington DC in response to the sustained demand on the US market.

- ***TEMIS, Leader in Text Mining in Europe, raises 3.6 million euros.***

Paris, November 9th 2004 - TEMIS, provider of corporate Text Mining solutions, has just completed a 3.6 million euro round of financing with ACE Management and Crédit Lyonnais Private Equity (CLPE).



QUELLE EST LA PROBLEMATIQUE?

Données



Les collaborateurs passent jusqu'à **60% de leur temps** à chercher et récupérer de l'information, temps qui est perdu pour l'analyse

Les données textuelles, par définition hétérogènes et multi-sources représentent **80% des données** de l'Entreprise.



Information



75% du temps des analystes de l'Entreprise (économiques, financiers, qualité, marketing, relation client, ...) est consommé par la lecture, le tri et l'extraction de l'information pertinente

La diffusion manuelle engendre **coûts, délais et commissions.**



Connaissance



Le surplus d'information tue la **connaissance...** et son partage.

Elle seule améliore les meilleures pratiques, facilite la prise de décision



Action



Simplifier les processus de décision

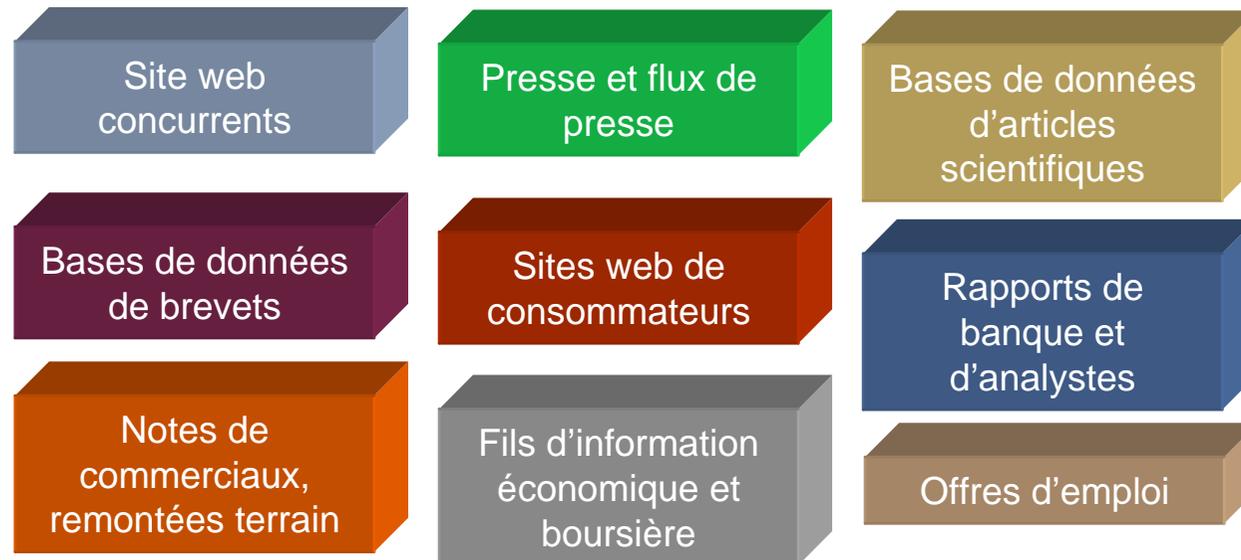
Réduire les risques et améliorer la qualité des décisions

Gagner du temps



UN EXEMPLE...

▪ *Les sources de données textuelles concernant vos concurrents et votre marché sont variées, multi-formats, multi-lingues.*

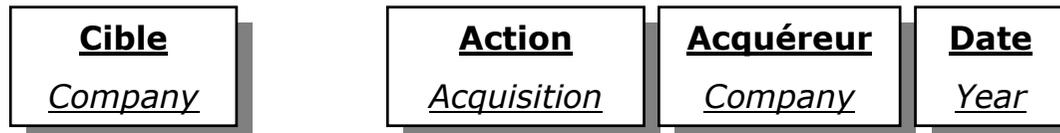


INTRODUCTION AU TEXT MINING

Transformer du texte brut en information à valeur ajoutée

Extrait

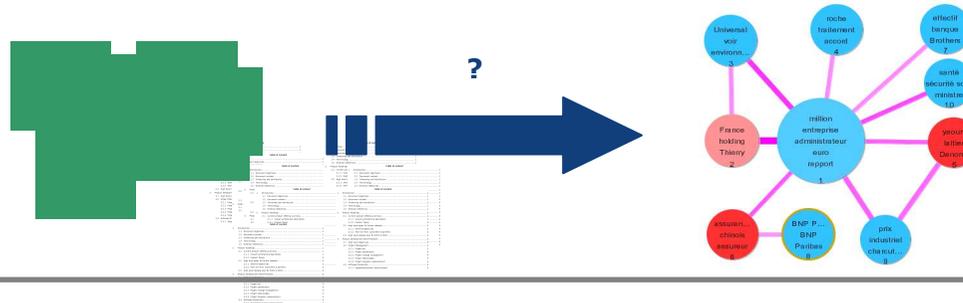
Pharmacia Corp. was acquired by Pfizer Inc. in 2000



Rachat

- Cible
- Acquéreur
- Lieu
- Date
- Montant
- ...

Organise



Catégorise



S KILL CARTRIDGE™

Skill
Cartridge™

- Modules d'analyse d'information spécialisés pour des domaines spécifiques de l'entreprise:
 - Surveillance de la concurrence / competitive intelligence
 - Lecture de documents scientifiques
 - Analyse de la relation client
 - Gestion des ressources humaines
- Les Skill Cartridges™ sont composées de dictionnaires et de règles grammaticales.



S KILL CARTRIDGE™

Skill
Cartridge™

Skill Cartridge™ Competitive Intelligence

- *Identifie les concepts économiques et stratégiques*
- **Financials** : turn-over, profitability, growth
- **Sales & marketing** : market share, pricing, communication, product launch
- **Research & Development** : R&D strategy, clinical trials, product efficacy, product dosing
- **Regulatory** : FDA approvals, letter of intent, reimbursement
- **Strategy** : licensing agreement, co-development, co-marketing, merger, acquisition, joint-venture
- **Organization** : appointments, position changes.



P

RODUCT AT A GLANCE

	APPLICATIONS	GAINS
	<p>Analyse stratégique (données financières, économiques, commerciales, boursières) ...sur un flux de presse ou le web</p>	<p>L'analyste va à l'essentiel Agit comme un filtre, une grille de lecture</p>
	<p>Analyse et navigation dans des fonds documentaires (articles économiques ou scientifiques, brevets...)</p>	<p>Transformation d'une masse de documents en information exploitable</p>
	<p>Catégorisation automatique d'articles économiques ou scientifiques, de brevets...</p>	<p>Gain de temps/productivité</p>

QUI L'UTILISE?

IPSEN
Innovation for patient care

NOVARTIS

MDL
an Elsevier company

Roche

Boehringer
Ingelheim

LION

CONOCO

TOTAL

EDF
Electricité
de France

DAIMLERCHRYSLER

PSA PEUGEOT CITROËN

Ford

RENAULT

AIRBUS

EADS

TRENITALIA

CREDIT LYONNAIS

HSBC

Dresdner Bank
Advice you can bank on

coface

French flag

European Union flag

CAS

CLIENT



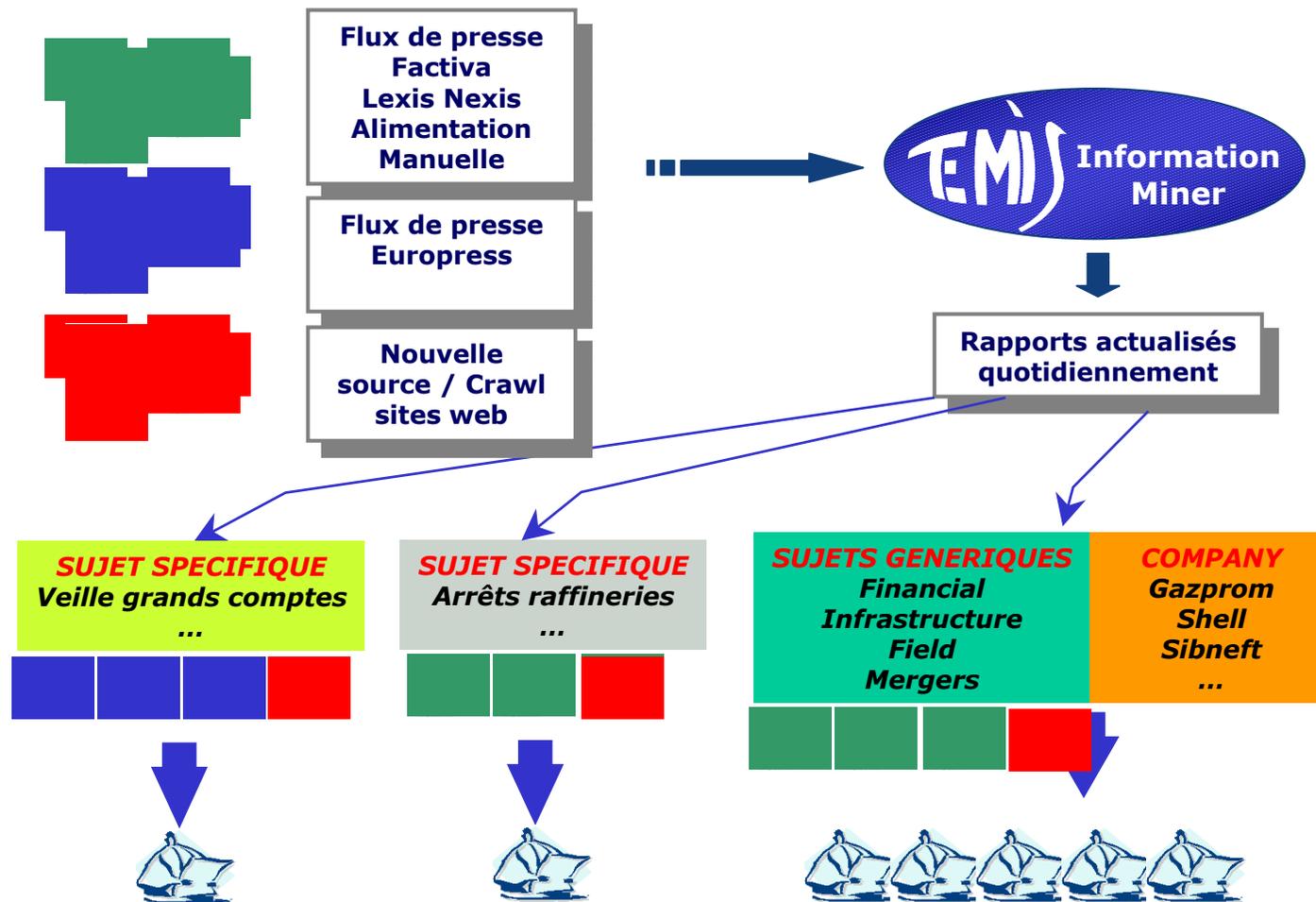
CAS CLIENT



- **Contexte**
 - ✓ Surveiller la concurrence et les marchés Pétrole / Gaz
- **Sources**
 - ✓ Reuters-Factiva, Lexis-Nexis, Crawl automatisé de sites web, Rapports d'analyse
- **Client**
 - ✓ Département VIA (Veille Information Archive, rattaché TOTAL Holding, en charge de distribuer/valoriser l'information pour le Groupe)
 - ✓ Direction Systèmes d'Information – DSIT
 - ✓ Projet « Information Miner »



INFORMATION MINER



Organisations | People | Categories | Documents

- Acquisition and Selling (133)
- Company Relation (139)
- Competition Regulation (44)
- Customer (15)
- Export Import (169)
- Field (75)
- Financial (77)
- Infrastructure (332)
- Investment (108)
- JointMarketing (2)
- Market Share (9)
- Mergers (9)
- Partnership (196)
- Privatization (15)
- Stake Information (101)
- Stock Index Information (4)
- Stock Information (71)
- Person - Head (320)





Organisations | People | Categories | Documents

- Acquisition and Selling (133)
- Company Relation (139)
- Competition Regulation (44)
- Customer (15)
- Export Import (169)
- Field (75)
- Financial (77)
- Infrastructure (332)
- Investment (108)
- JointMarketing (2)
- Market Share (9)
- Mergers (9)
- Partnership (196)
- Privatization (15)
- Stake Information (101)
- Stock Index Information (4)
- Stock Information (71)
- Person - Head (320)

Customer

- [Distriqaz Nord](#) <=> 1.3 million clients
- [Distriqaz Sud](#) <=> around 2.5 million clients
- [EconGas](#) <=> two former clients
- [EconGas](#) <=> two former clients
- [Endesa](#) said has gained 95,000 household customers for its joint power
- [Enel](#) <=> 28 million electricity customers in Italy
- [Essent's](#) <=> customer base is presently just below 2 million
- [Gazexport's](#) <=> three customers
- [German Wingas](#) says 40 domestic customers

Organisations | People | Categories | Documents

Distrigaz Nord

Graph (36)

Acquisition and Selling

- CSFB
<=>
 - Romgaz
 - Distrigaz Sud
 - Distrigaz Nord
- Romgaz.The
<=>
 - Distrigaz Sud
 - Distrigaz Nord

Company Relation

- CSFB
<=>
 - Romgaz
 - Distrigaz Sud
 - Distrigaz Nord
- <=>
 - Distrigaz Sud
 - Distrigaz Nord

Customer

- Distrigaz Nord
<=>
1.3 million clients

Financial

- Distrigaz Nord
<=>
turnover of \$380 million in 2002 .

Infrastructure

- Distrigaz Nord
 - Transylvania
 - Distrigaz Sud
<=>
 - its 17,000 kilometer pipeline network
 - pipes

Investment

- Distrigaz Sud

Customer

- Distrigaz Nord
<=>
1.3 million clients
- Distrigaz Sud
<=>
around 2.5 million clients
- EconGas
<=>
two former clients
- EconGas
<=>
two former clients
- Endesa
said
has gained 95,000 household customers for its joint power
- Enel
<=>
28 million electricity customers in Italy
- Essent's
<=>
customer base is presently just below 2 million
- Gazexport's
<=>
three customers
- German Wingas
says
40 domestic customers

Poste de travail

Organisations | People | Categories | Documents

- Acquisition and Selling (133)
- Company Relation (139)
- Competition Regulation (44)
- Customer (15)
- Export Import (169)
- Field (75)
- Financial (77)
- Infrastructure (332)
- Investment (108)
- JointMarketing (2)
- Market Share (9)
- Mergers (9)
- Partnership (196)
- Privatization (15)
- Stake Information (101)
- Stock Index Information (4)
- Stock Information (71)
- Person - Head (320)

Customer

Distrigaz Nord	<=>	1.3 million clients
Distrigaz Sud	<=>	around 2.5 million clients
EconGas	<=>	two former clients
EconGas	<=>	two former clients
Endesa	said	has gained 95,000 household customers for its joint power
Enel	<=>	28 million electricity customers in Italy
Essent's	<=>	customer base is presently just below 2 million
Gazexport's	<=>	three customers
German Wingas	says	40 domestic customers

Map Navigator
 Drag Rectangle to change extent in main map

E.ON's : Stake Information
 E.ON : Company Relation
 E.ON : Partnership
 E.ON : Company Relation
 E.ON : Acquisition and Selling
 E.ON/Ruhrgas : Acquisition and Selling
 E.ON/Ruhrgas : Company Relation
 E.ON/Ruhrgas : Partnership
 Enel : Acquisition and Selling
 GDFPp.PA : Partnership
 GDFPp.PA : Company Relation
 Latvijas gaze : Company Relation
 OMV : Partnership
 SPP's : Company Relation

Gazprom Ruhrgas Gaz de France and Enel were also shortlisted for the sale of Distrigaz Sud

Gazprom, Ruhrgas, Gaz de France and Enel were also shortlisted for the sale of Distrigaz Sud.

The ministry did not specify the deadline for submitting non-binding bids, but an official from the Economy Ministry said last year that final bids must be lodged by February 29.

A company can bid for both utilities but can only buy one of them, Economy Ministry officials said.

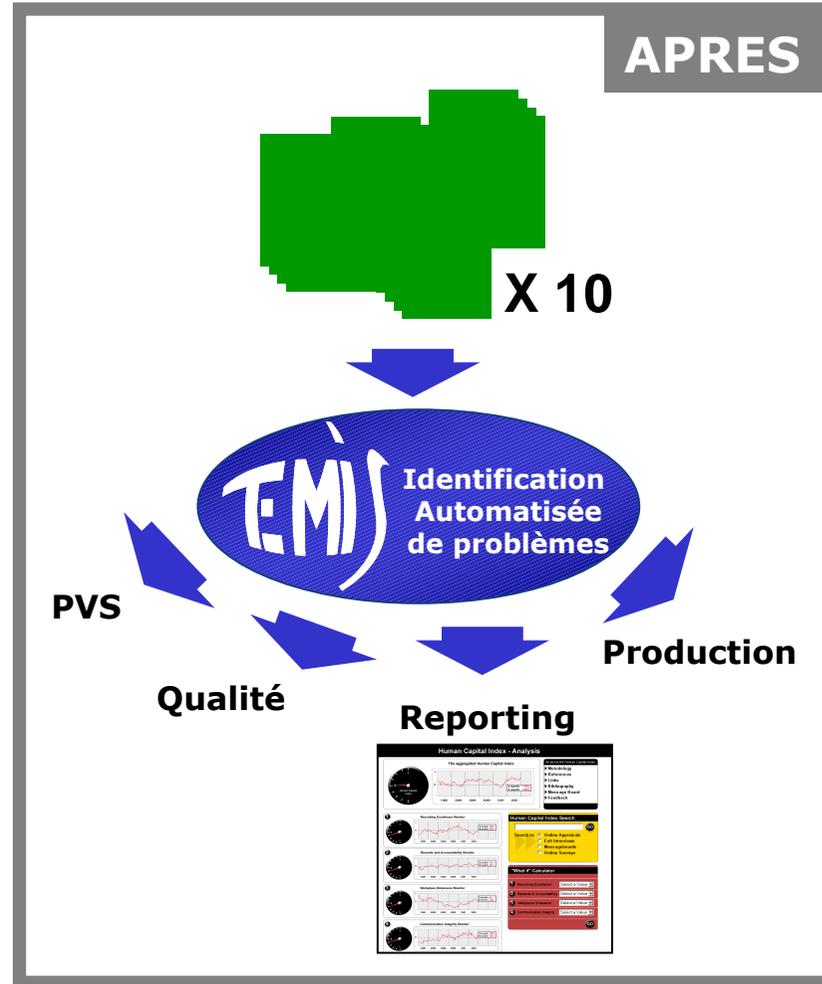
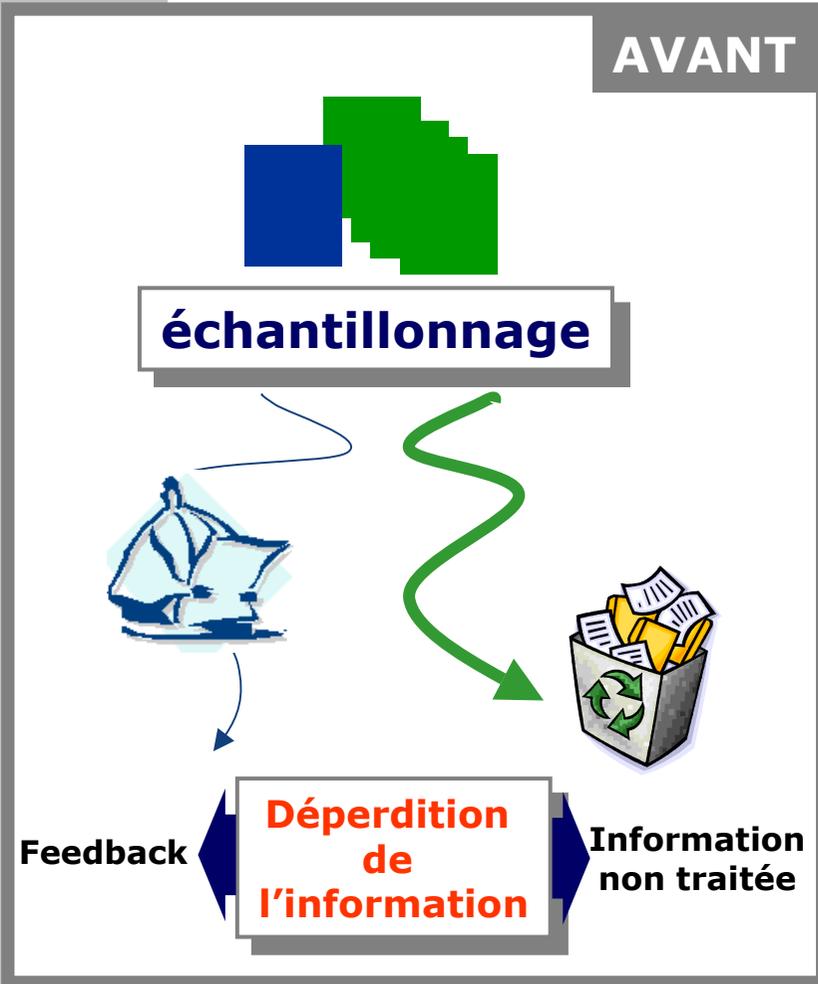
The government will sell 30 percent of Distrigaz Nord and Distrigaz Sud outright and at the same time the bidder must buy newly issued shares to raise its stake to 51 percent.

Poste de travail

CAS CLIENT

- **Contexte**
 - ✓ Surveiller la qualité perçue, en recueillant les sentiments des clients
- **Sources**
 - ✓ Des enquêtes SOFRES stockées dans INFOQUA
- **Interlocuteurs**
 - ✓ Direction Systèmes d'Information – DSIN
 - ✓ Pôle S-Do-Q : Systèmes documentaires Qualité
 - ✓ Projet ADEC, Analyse Des Enquêtes Clients
 - ✓ Intégré à INFOQUA, portail qualité groupe.
- **Clients**
 - ✓ Les PVS
 - ✓ DEQUA

CAS CLIENT



CAS CLIENT



- **Contexte**
 - ✓ Simplifier l'accès à l'information scientifique pour les chercheurs, identifier automatiquement les nouvelles pistes de recherche, les traitements potentiels et les méthodes animales présents dans la littérature.

- **Sources**
 - ✓ PubMed-Medline, IDDB3, Proceedings de conférence

- **Client**
 - ✓ Direction Informatique Recherche
 - ✓ Département Discovery au sein de la Recherche, à l'Institut Henri Beaufour
 - ✓ Projet « **Ipsen Text Mining** » (ITeM)

Search

In all fields

AND In Texts

(Combine Textual searches : at least one all)

<input type="text"/>	in field Title
<input type="text"/>	in the Skill Cartridge™ Item
<input type="text"/>	in the Skill Cartridge™ Mesh
<input type="text"/>	in field Authors

AND In Sources

[Uncheck All](#)

- Neuroscience
- Mouvement Disorders

Specify date

After : Before :
(DD/MM/YYYY) (DD/MM/YYYY)

Limit number of doc extracted

Max. Docs

Submit

Search Results

1000 documents found.

1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19 20 > >> Last

News	Rank	Date
The management of tremor.	100.0	01/01/200
Origins of tremor.	100.0	01/01/200
[Postural or orthostatic tremor: a variant of static tremor]	97.9	01/01/201
Clinical measurement of tremor.	97.2	01/01/199
Tremor in children.	95.1	01/01/200

rom : /Mouvement_disorders/PubMed

The many forms of tremor. Precise classification guides selection of therapy.

The many forms of **tremor**. Precise classification guides selection of therapy.

Tremor is a common, potentially disabling symptom that occurs with a variety of

neurologic conditions. [Rational therapy is enhanced by correct identification of](#)

[the tremor type](#) and, when possible, of its cause. In this review, the basic

phenomenology of **tremor** is presented, followed by an examination of the most

common **tremor types** and their treatment. [Surgical therapy for medically](#)

[refractory tremor](#) is discussed.

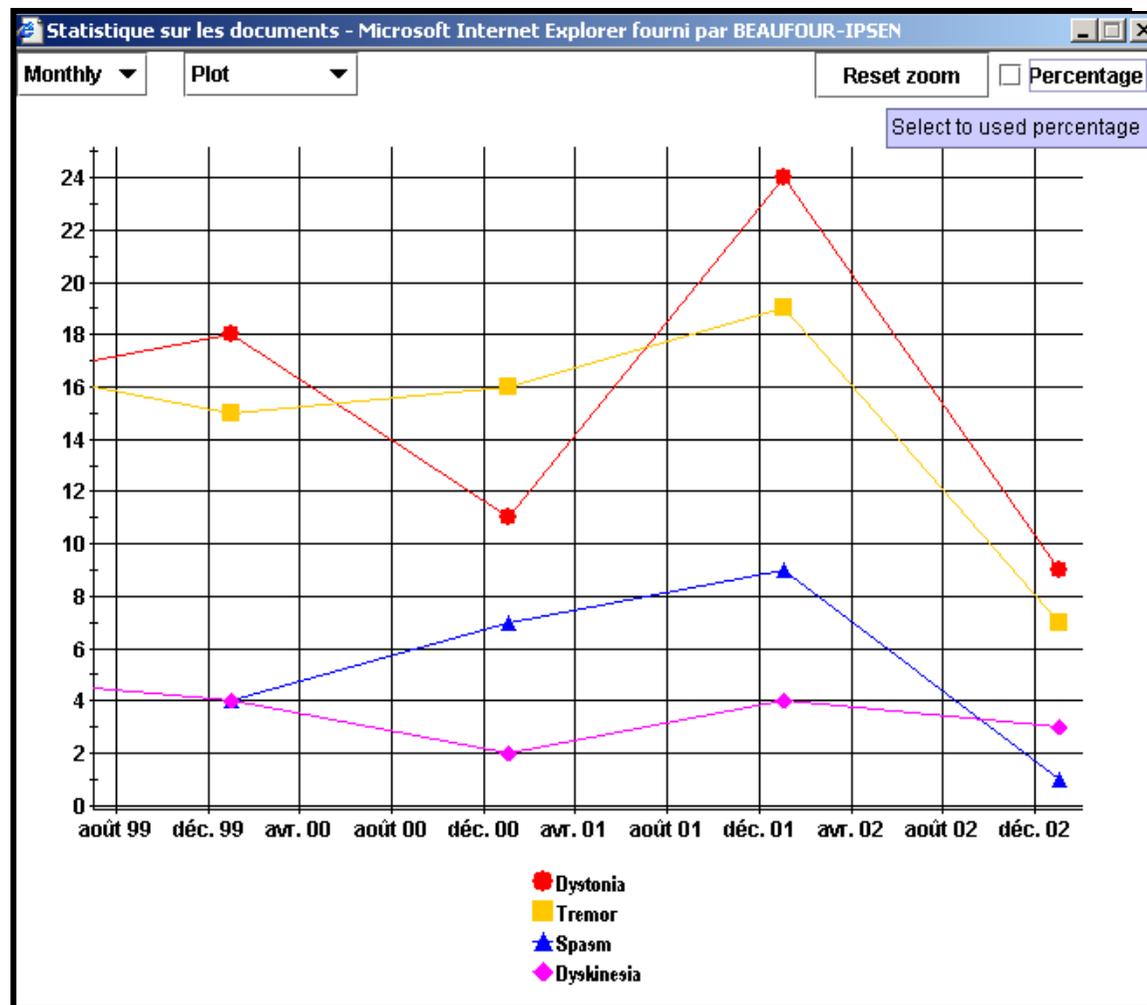
/Tremor Treatment/rational therapy be enhance by correct identification of the tremor type

-  **Ipsen**
 - [refractory tremor](#)
 - [tremor](#)
 - [tremor type](#)

- [rational therapy](#)
- [surgical therapy](#)

- [rational therapy be enhance / identification of the tremor type](#)
- [surgical therapy for medically tremor](#)
-  **Mesh**
-  **Author**

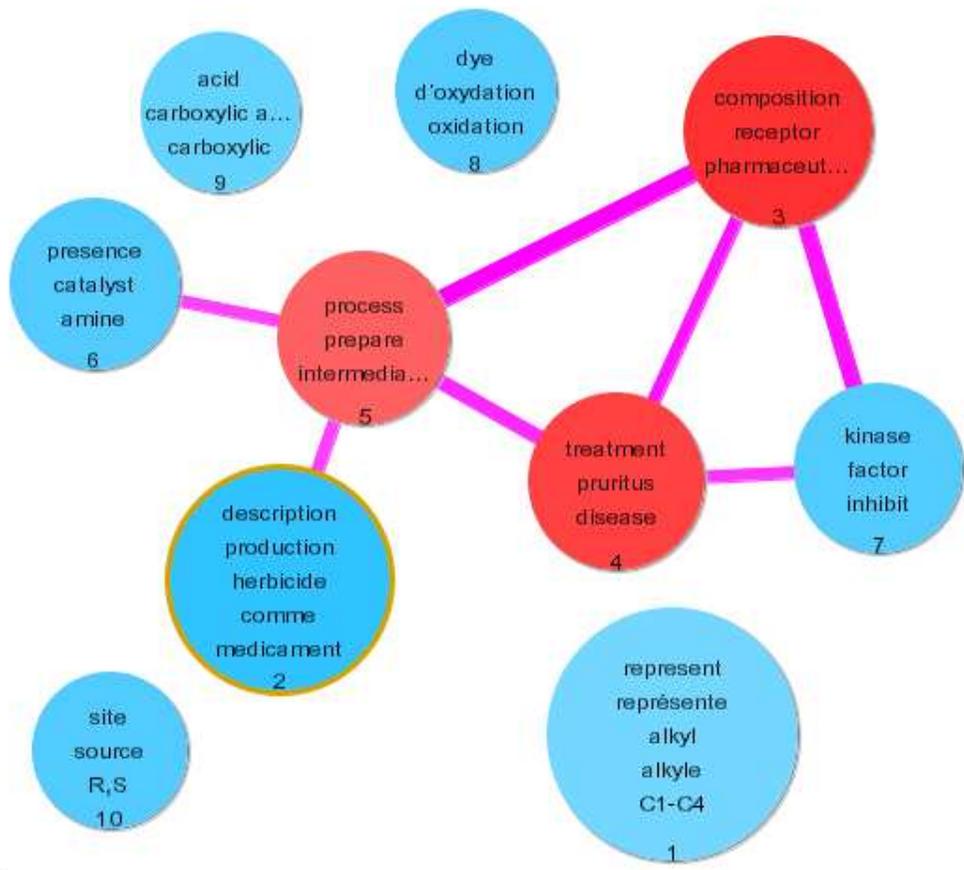
CAS CLIENT



[Full screen](#) | [Match this size](#) | [Close Window](#)

Displaying cluster : /

[Upper Cluster](#) [Animation : Off](#) [Print](#)



- 44 Field : Patent Assignees
- 40 ● Aventis Pharma Limited
 - 36 ● Aventis Pharma S.A.
 - 36 ● AVENTIS PHARMACEUTICAL PRODUCTS INC.
 - 32 ● Aventis Pharmaceuticals Inc.
 - 32 ● AVENTIS PHARMACEUTICALS PRODUCTS INC.
 - 28 ● AVENTIS PHARMACEUTICALS, INC.
 - 28 ● Aventis Research & Technologies GmbH & Co KG
 - 28 ● Aventis Research & Technologies GmbH & Co. KG.
 - 24 ● BASF AKTIENGESELLSCHAFT
 - 24 ● BAYER AG
 - 20 ● Bayer Aktiengesellschaft
 - 20 ● Bayer Corporation
 - 20 ● BAYER YAKUHIIN, LTD.
 - 16 ● Eif Atochem S.A.
 - 16 ● GLAXO GROUP LIMITED
 - 12 ● GÖDECKE AKTIENGESELLSCHAFT
 - 12 ● L'OREAL
 - 8 ● NIHON BAYER AGROCHEM K.K.
 - 8 ● ONYX PHARMACEUTICALS
 - 8 ● **PFIZER INC.**
 - 8 ● **Pfizer Limited**
 - 8 ● **PFIZER PHARMACEUTICALS INC.**
 - 8 ● **Pfizer Products Inc.**
 - 4 ● SANOFI-SYNTHELABO
 - 0 ● YEDA RESEARCH AND DEVELOPMENT CO. LTD.

Cluster : 2
 Name **description ; production ; herbicide ; comme ; medicament ; pesticide ; signification ; produce ; utilisation ; relate**
 Size : 97
 Frequency : 0 %

Name	Support	Complete name
descriptinn	0.078	/COMMON-NOUN/descriptinn

[Close Window](#)

C ONCLUSION

QUESTIONS REponses



Do you speak " BUSINESS " ?

Impact des logiciels de traduction et d'aide au multilinguisme sur l'efficacité dans l'Entreprise et dans la vie privée

Documentation technique, Notices techniques, Formulaires, Présélection de CV.....

Maison des Arts et Métiers
7 février 2005

Do you speak " BUSINESS " ?

Usage des logiciels de traduction

Un enjeu pour l'Entreprise d'aujourd'hui :

lever la barrière des langues

ou

Comment des innovations dans le domaine de la linguistique apportent des outils stratégiques aux entreprises confrontées à la globalisation et au traitement rapide de l'information multiculturelle ?



business &
technologie
CLUB PROFESSIONNEL EDEC

INTERMINES
PARIS - SAINT-ETIENNE - NANCY



L'Informatique peut-elle supprimer la barrière des langues ?

- Historique
- La Situation aujourd'hui
- Les Outils mis à disposition pour le Personnel EADS
- L'usage de ces Outils
- Améliorations souhaitées
- Bilan général

Historique

Mise en place des outils et des méthodes associées



Dans les années 90, Aerospatiale (40000 Salariés) a défini des solutions pour améliorer :

- la communication interne et entre partenaires
- la compréhension entre les experts de différents pays
- le travail collaboratif
- l'efficacité de toute la chaîne documentaire.

(contraintes et exigences contractuelles: (Simplified English (SE) de AECMA))

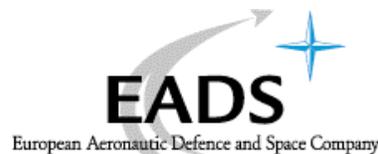
1999



56000 Salariés

EXTENSION DES GROUPES DE TRAVAIL

2000



110000 Salariés

BESOIN URGENT D' INTEGRER DES EQUIPES MULTINATIONALES:
mise en place d'un large dispositif l'aide à la traduction pour tout le Personnel :

2001 IMPLEMENTATION DE LA SOLUTION REVERSO INTRANET



business &
technologie
CLUB PROFESSIONNEL EDEC

INTERMINES
PARIS - SAINT-ETIENNE - NANCY



Historique

Mise en place des outils et des méthodes associées



Des débuts difficiles:

- **Les Allemands très critiques sur ces Outils**
 - choix français et mis en place par les français!
 - couple Allemand ↔ Anglais sans doute perfectible!
 - meilleure maîtrise de la langue étrangère
- **Le niveau de maîtrise de la langue étrangère impacte très fortement le jugement de l'utilisateur:**

Les « bons » en anglais jugent ces outils sur la qualité de la traduction ...



Maintenant



Cette application est aujourd'hui la plus utilisée, parmi toutes les applications accessibles via l'Intranet et différents Portails existants :

400000 requêtes en octobre 2004:

- Anglais-Français : 33%
- Français-Anglais : 21%
- Anglais-Allemand: 18%
- Allemand-Anglais: 13%
- Autres : 15% (veille techno)

et tendance à la baisse en fin 2004

ce qui démontre:

- l'utilité de ces outils dans un Groupe qui est né de la fusion d'Entreprises allemandes, espagnoles, et françaises, et qui maintenant comprend des Anglais, Italiens ...
- un ROI très court : de l'ordre de 2 à 3 jours
- intégration en bonne voie des équipes multinationales chez EADS



business &
technologie
CLUB PROFESSIONNEL EDHEC

INTERMINES
PARIS - SAINT-ETIENNE - NANCY



Les Outils à la disposition du Personnel Du Groupe EADS



**business &
technologie**
CLUB PROFESSIONNEL EDEC

INTERMINES
PARIS - SAINT-ÉTIENNE - NANCY





TRADUCTION DE TEXTES COURTS

Tapez directement ou collez le texte à traduire :
(le nombre de caractères est limité à 1024, soit environ 15 lignes)

Caractères spéciaux
à é î ô ù Autre

L'Informatique peut-elle supprimer la barrière des langues ?

Direction de traduction :
Français-Anglais

Domaine :
Optronics
Techno + Business
Technology (Missiles)
OK CLEAR

Texte traduit :

Can the Data processing delete the barrier of the languages?

- Textes
- Pages URL
- Fichiers
- Lexibase COLLINS
- Accueil
- Contact

Mon Reverso

Glisser-déposer le lien du pop-up de traduction Mon Reverso dans votre barre des liens pour traduire des parties de textes à tout moment.





LEXIBASE INTRANET



Collins Fr-En

Recherche

barrière

Entrées

- barrière
- barrière (de passage à niveau)
- barrière de corail
- barrière de dégel
- barrière de sécurité
- barrière douanière
- barrière naturelle
- barrique
- barrier
- barrissement
- bartavelle
- barrycentre
- barryton
- barryum
- bas {1}
- bas {2}

Suggestions

- barde-barrière
- barrier

barrière /bar.je.r/

1 nf (= clôture) fence, (= porte) gate, (lit fig) (= obstacle) barrier, (Hist) (= octroi) tollgate
 dresser une barrière to put up a barrier (**entre** between)
 franchir la barrière de la langue to break through the language barrier

2 mots composés

- ♦ **barrière de corail** coral reef
la Grande Barrière (de corail) the Great Barrier Reef
- ♦ **barrière de dégel** roadsign warning of dangerous road conditions during a thaw
- ♦ **barrière douanière** trade ou **tariff barrier**
- ♦ **barrière naturelle** natural barrier
- ♦ **barrière (de passage à niveau)** level (Brit) ou **grade (US)** crossing gate
- ♦ **barrière de sécurité** (dans les rues) crowd barrier, (pour un bébé) safety gate



Collins Robert



L'usage de ces Outils

Recommandations à l'Usage

- Les Juniors ne pensent pas et ne réagissent pas comme les Séniors : beaucoup plus à l'aise avec les langues !!!
- Les Français sont atypiques : nuls en langues étrangères !!!
- Les outils sont et restent des outils d'aide à la traduction, voire d'aide à la compréhension : une traduction précise et réutilisable directement après traitement par ces outils automatiques, nécessite des retouches et l'aide de dictionnaires spécifiques:
 - Optronics
 - Techno+Business
 - Technology (Missiles)
- Accessibilité (Intranet , pop-up,...)
possibilité de traduire à tout moment

Différentes utilisations

1. Communication (besoin de)

- **Prise de connaissance globale d'un texte, document ou message**
- **Comprendre et se faire comprendre**
- **Echanger plus rapidement.**
- **Casser la barrière des langues : moins de retenue, de peur....**

2. Compréhension :

- **Récupération de textes en langues non maîtrisées et complètement incompréhensibles**
(très utile pour la veille technologique: comprendre du Russe, Chinois,...)
- **Utilité de dictionnaires spécifiques.**

Différentes utilisations

3. Traduction :

- Fourniture rapide d'un draft
- Remplacement de termes avec du bon vocabulaire
- Enrichissement de dictionnaires
- Tri qui permet de ne traduire que l'essentiel
(par exemple : ne traduire exactement que 3 pages d'un document de 20 pages)
- Complément à une lecture directe en langue étrangère

Les fonctionnalités nécessaires des Outils



- Doivent permettre la mémorisation de traduction
- Doivent signaler les mots non traduits et proposer plusieurs possibilités
- Doivent avoir une dimension sémantique
- Doivent être évolutifs et « s'auto enrichir ».
- Doivent être adaptés à différents formats.
- Doivent permettre de créer ses propres dictionnaires
- Doivent permettre le pré-traitement orthographique et grammatical (amélioration de la qualité du texte source)



Les limites

- Certains formats sont des freins à la communication et à l'échange : antagonisme entre intégrité d'un document et sa compréhension
- Peu intégrés à la GED
- L'administrateur n'a pas accès aux règles (modifications, corrections)
Trop de dépendance vis à vis du Fournisseur
- Les dictionnaires spécialisés

Recommandations :

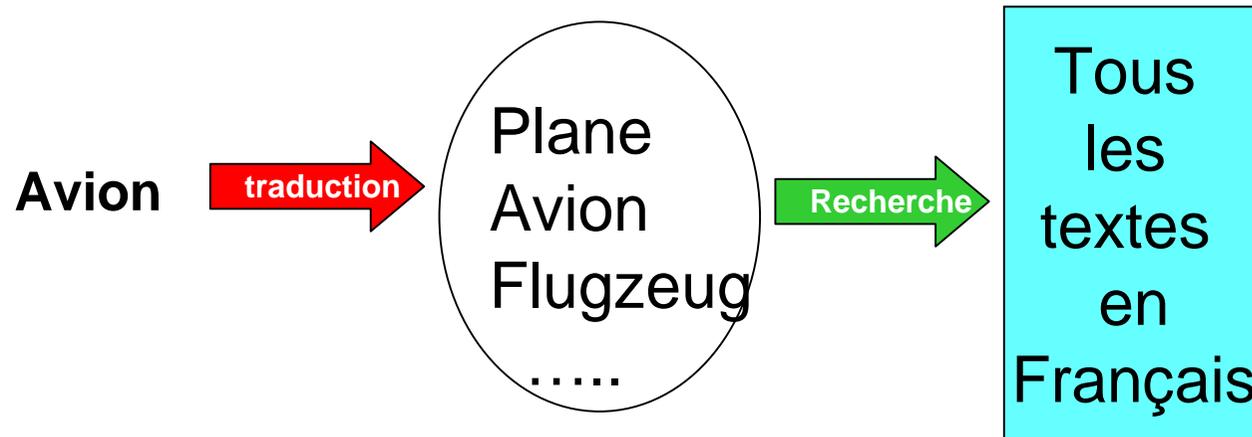
Toutes ces limites pourraient sans doute être atténuées, voire supprimées, si une solide structure projet existait pour assurer la mise en place, le déploiement, l'évolution et la maintenance de ces applications.

Il faut faire vivre ces Outils !

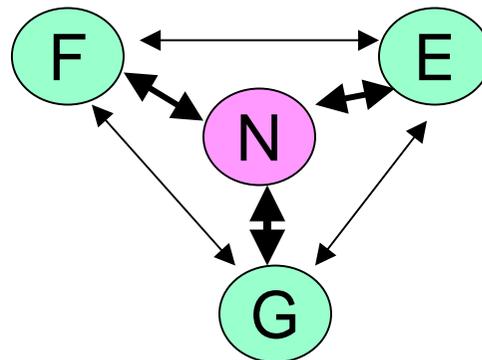
Les améliorations souhaitées

Les fonctionnalités

➤ Recherche multi-langues



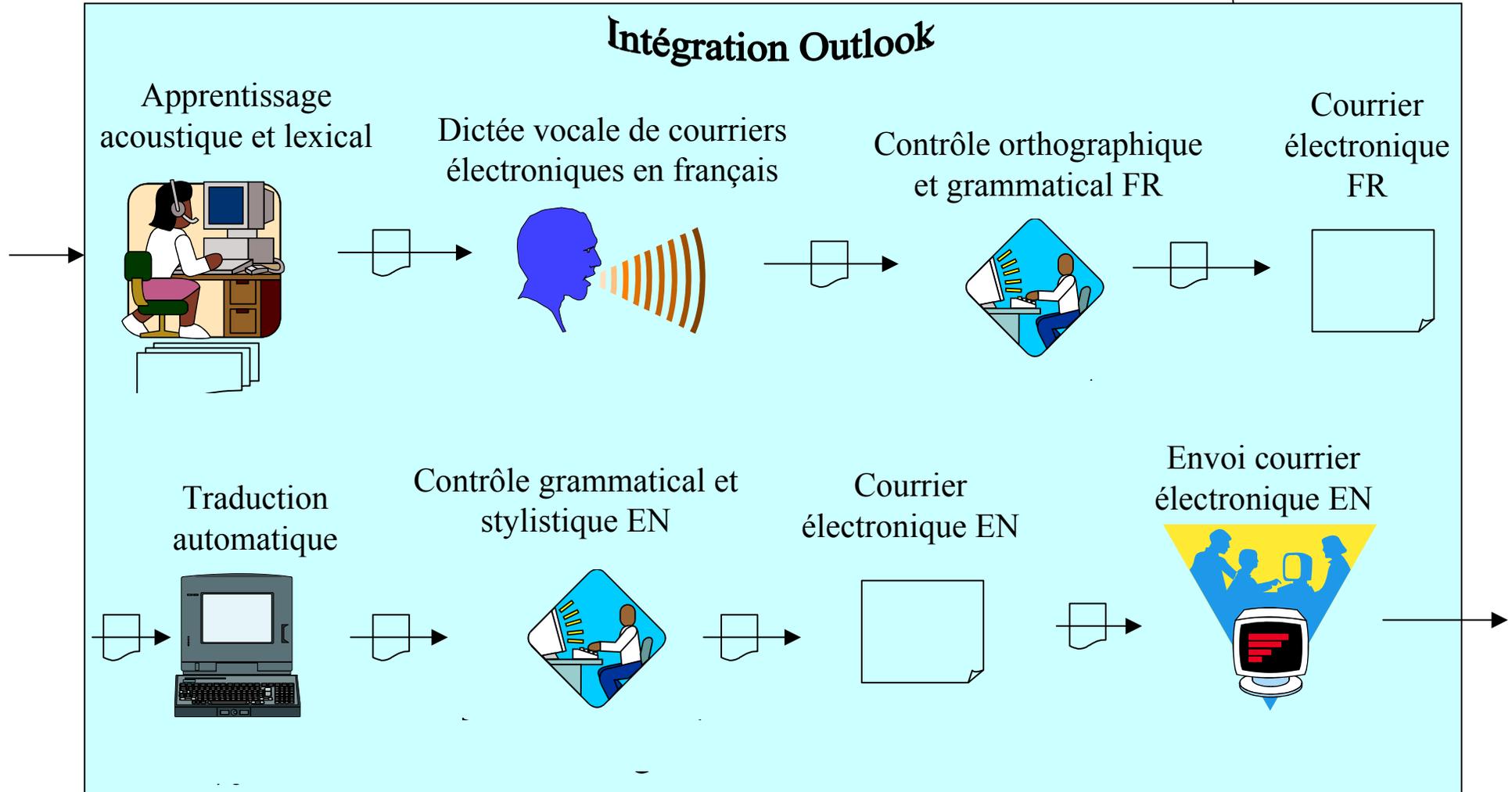
➤ Traduction en utilisant un « noyau » commun



Une chaîne documentaire complète



Une chaîne documentaire complète



En guise de conclusion

Différentes vues sur Le multilinguisme



Souvent jugé comme un handicap, une entreprise ayant besoin de maîtriser le multilinguisme interne possède un atout face aux Entreprises Américaines:

notre culture nous pousse au multilinguisme et à l'expérience internationale: nécessité économique ou simple sentiment naturel. <http://www.insert-export.com>

Considérant que, de manière générale, le client préfère être servi dans sa langue, il est possible de déduire de cette évolution que, à court terme, l'anglais verra sa place relative diminuer en matière de commerce électronique.

Dans ce contexte, il sera risqué pour une entreprise québécoise faisant des affaires à l'échelle internationale de seulement mener ses activités de commerce électronique en anglais, tout comme il serait dangereux pour elle de se limiter au français.

Cela signifie que **l'internationalisation des sites Web** québécois, c'est-à-dire leur **adaptation** aux besoins culturels et **linguistiques** de leurs clients, fournisseurs, partenaires ou autres, **constitue un enjeu important.** <http://www.cefrio.qc.ca/projets/Documents/Enjeuxint.html>



Du vécu !

- **Oui les Technologies de l'Information ont levé les barrières des Langues**
- **Dans les équipes multinationales habituées à communiquer, un véritable langage commun avec ses termes propres s'impose : la nécessité de communiquer et se comprendre prime sur la qualité de " l'Anglais »**
- **Les handicaps subsistent dans des situations difficiles où l'oral prédomine comme des négociations, démonstrations, règlements de conflits ...**
- **Les Anglo-Saxons ne nous comprennent pas toujours et réciproquement!!**





business &
technologie
CLUB PROFESSIONNEL EDHEC

INTERMINES
PARIS - SAINT-ETIENNE - NANCY



Do you speak " BUSINESS " ?

Impact des logiciels de traduction et d'aide au multilinguisme sur l'efficacité dans l'Entreprise et dans la vie privée

Documentation technique, Notices techniques, Formulaires, Présélection de CV.....

Maison des Arts et Métiers
7 février 2005



Do you speak Business?

Le traitement automatique des langues, au delà de la traduction

Patrick Mazeau

Xerox Research Centre Europe

G9+ 7 février 2005



La société XEROX

- ❑ CA de 15,7 Milliards de \$
- ❑ 61,000 personnes
- ❑ Créée en 1948 (Chester Carlson applique ses brevets chez Haloid)



Office Group
Work Centre 2128
Phaser 8400



new

Production Printing
DocuColor 8000
iGen3

XEROX® GLOBAL SERVICES

Technology | Document Management | Consulting Services

Xerox Innovation Group: une présence globale



Sleepy Hollow, NY



Webster, NY



Mississauga, Canada



Grenoble, France

❑ **La recherche est le moteur de la croissance**

- 717 brevets en 2004, au total plus de 16000 brevets US (sans Fuji Xerox)
- 9 chercheurs ont déposé chacun plus de 100 brevets

❑ **Investissements en R&D**

- 850 millions de \$ pour Xerox
- 500 millions de dollars pour Fuji Xerox



Palo Alto, CA



El Segundo, CA

Historique de la recherche linguistique

□ Années 70

- Recherches au PARC: langage naturel, correcteurs orthographiques, outils de recherche rapide dans les dictionnaires

□ Années 90

- Création du Xerox Research Centre Europe
- Automates a états finis, méthodes statistiques

□ Années 2000

- Retour de l'apprentissage machine en conservant les autres méthodes



L'analyse syntaxique incrémentale (Xerox Incremental Parser)



Le concept

□ L'analyseur syntaxique

■ *Il réunit des mots correspondant à une unité linguistique appelée syntagme*

◆ *« le beau chien » est un syntagme nominal*

◆ *« a mangé » est un syntagme verbal*

■ *Il renvoie les liens grammaticaux entre les mots les plus importants de la phrase*

◆ *SUJET(manger, chien)*

Comment ça marche?

- ❑ Un analyseur syntaxique découpe d'abord le texte en mots
- ❑ Puis il groupe ces mots en phrases ou en paragraphes, que nous appellerons nos unités linguistiques.
- ❑ L'analyseur syntaxique applique alors sur ces groupes de mots un ensemble de règles, la grammaire.
- ❑ La grammaire effectue alors les tâches suivantes:
 - Elle **simplifie** les mots ambiguës. Ainsi « ferme » peut être *nom*, *verbe* ou *adjectif*.
 - Elle groupe ensemble les mots en **syntagme**
 - Elle extrait les **relations** que les têtes de ces syntagmes entretiennent les unes avec les autres.

Exemple de simplification

- La dame a acheté une ferme.
 - *dame* est un verbe ou un nom.
 - *ferme* est un verbe, une nom et un adjectif.
 - *La* est un pronom ou un article

- La grammaire contient des règles qui permettent de choisir la bonne interprétation:
 - *dame* est suivi d'un auxiliaire, c'est donc un nom
 - *La* est suivi d'un nom, c'est donc un article
 - *ferme* est précédé d'un article, c'est donc un nom.

Exemple de syntagme

- ❑ La dame a acheté une ferme
 - La dame
 - a acheté
 - une ferme

- ❑ Chaque syntagme est défini par sa tête:
 - *La dame* c'est un syntagme **nominal** dont la tête est *dame*
 - *a acheté* c'est un syntagme **verbal** dont la tête est *acheté*
 - *une ferme* c'est un syntagme **nominal** dont la tête est *ferme*

Exemple de relation

- ❑ La dame a acheté une ferme
 - Nous pouvons extraire les relations suivantes:
 - ◆ Sujet(acheter, dame)
 - ◆ Objet(acheter, ferme)
 - Ces relations ont une grande importance, c'est elle que l'on peut utiliser pour déterminer le sens d'une phrase, ce qui peut améliorer la recherche d'information.

Spécialisation d'une grammaire

- Le formalisme de XIP est très riche. Il permet de spécialiser une grammaire pour des cas particuliers:
 - On peut par exemple extraire des entités d'un texte:
 - ◆ Les noms propres
 - ◆ Les noms de compagnie
 - ◆ Les noms de lieu
 - ◆ Les dates
 - ◆ Les quantités et autres mesures
 - On peut détecter des tournures stylistiques particulières

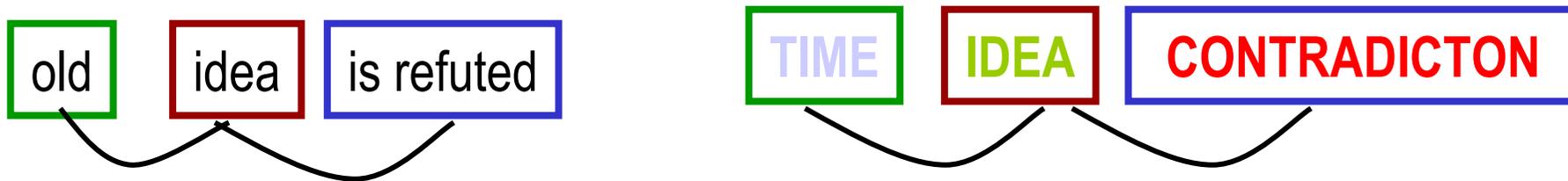
Extraction d'entités

- ❑ **Margaret Sinclair Trudeau**, born **September 10 ,1948** in **Vancouver, British Columbia, Canada** , was the wife of the late Canadian **Prime Minister Pierre Trudeau**. The daughter of **James Sinclair**, a former Liberal member of the **Parliament of Canada** and fisheries minister, she attended **Simon Fraser University** where she obtained a degree in English literature

- ❑ **Sortie du parser incrémental**
 - PERSON (Margaret Sinclair Trudeau)
 - TEMPEXPR_DATE (September 10 , 1948)
 - LOCATION_CITY (Vancouver)
 - LOCATION (British Columbia)
 - LOCATION_COUNTRY (Canada)
 - PERSON (Prime Minister Pierre Trudeau)
 - PERSON (James Sinclair)
 - ORGANISATION (Parliament of Canada)
 - ORGANISATION (Simon Fraser University)

Repérer des configurations particulières

Détection de contradiction



■ *... challenged many of our previous notions ...*

■ *... In contrast to the traditional view ...*

- Ce que ces deux expressions ont en commun:
 - Mots clefs spécifiques
 - Relié par une relation grammaticale.

Vitesse d'analyse syntaxique

❑ Élément crucial: la vitesse

- Pour utiliser un analyseur dans le monde industriel, il faut au moins pouvoir traiter 1000 mots/s.
- La complexité d'un analyseur syntaxique est exponentielle
- La vitesse des machines croît de façon linéaire. Il est difficile dès lors espérer résoudre les problèmes d'efficacité en tablant sur une plus grande vitesse des machines.



Catégorisation / Clustering



Clustering / Catégorisation

❑ Le Clustering et la Catégorisation sont des principes cognitifs de base

- Une machine qui produit des copies couleur → Copieur Xerox

❑ La Catégorisation

- Assignation d'un document dans une ou plusieurs catégories

- ◆ Manuelle: précise mais chère pour les hauts volumes, risques d'erreurs importants (15% à l'Office Européen des Brevets)
- ◆ Automatique: bon marché, rapide et cohérent, mais pas toujours précis
- ◆ Semi-automatique: avantages des deux approches

❑ Le Clustering

- Découverte des catégories en analysant un texte (apprentissage machine)

❑ Quels sont les besoins en catégorisation?

- *Aider à retrouver des documents: nous passons 20 à 30 % de notre temps à rechercher des documents qui devraient être plus faciles à localiser (Delphi Group)*

Clustering / Catégorisation

□ Les méthodes

■ Clustering

- ◆ Apprentissage machine
- ◆ Analyse d'un corpus de documents pour en extraire les mots-clés communs
- ◆ Création de catégories, hiérarchie possible

■ Catégories existantes (ex: Docushare, DMOZ, EPO)

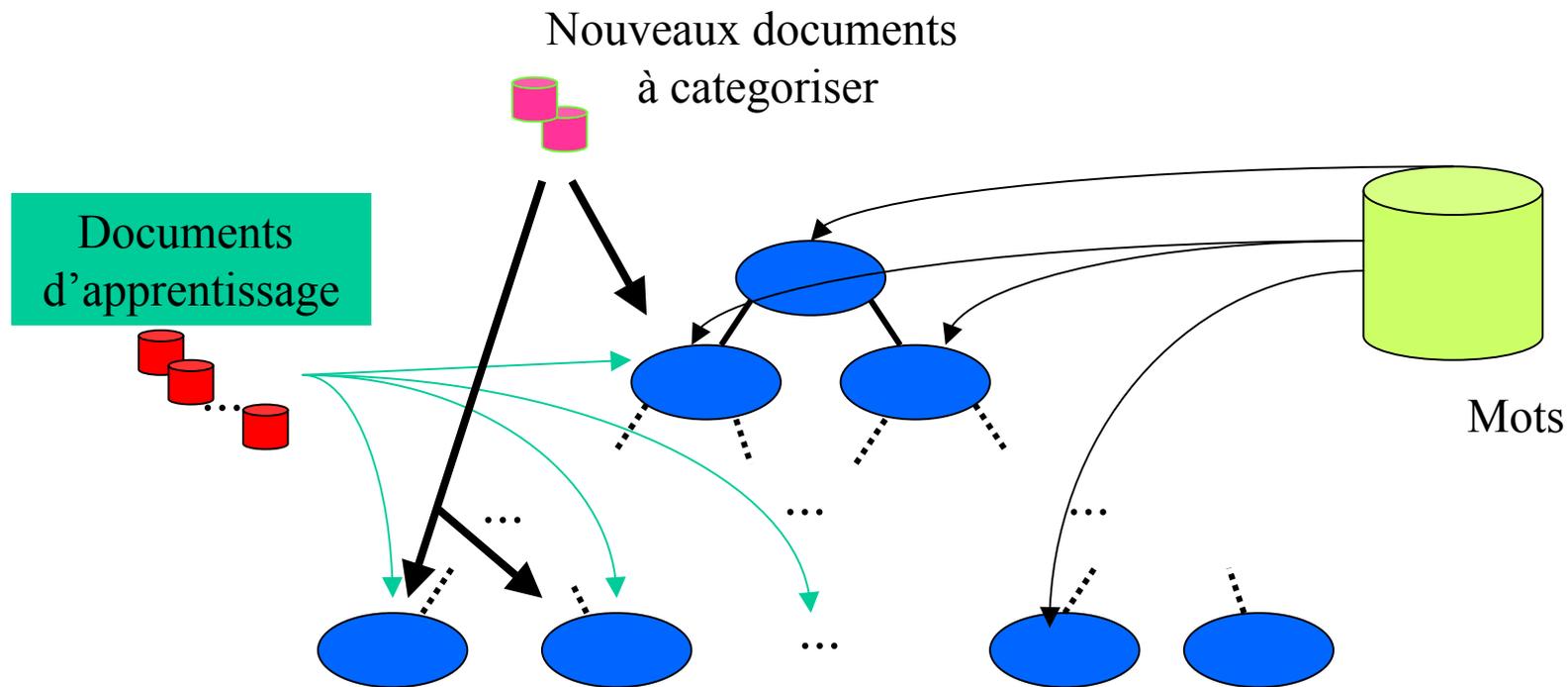
- ◆ Extraction des mots-clés des documents et association avec les répertoires (catégories)
- ◆ Utilisation de ces listes pour ranger de nouveaux documents

■ Utilisation conjointe

- ◆ De technologies linguistiques (extraction de mots-clés)
- ◆ **De techniques d'apprentissage machine**

Clustering / Catégorisation

CATEGORISATION



Catégorisation

□ Nos forces?

- Modules génériques (multi-class, multi-label, hierarchies)
- Performances excellentes (haute précision, grande vitesse)
- Moins de données d'entraînement nécessaires que les autres systèmes
- Précision obtenue grâce à notre analyse linguistique

□ Technologie appliquée à différentes collections de documents

- Reuters, DMOZ, El País, biologie, Topster (TDT-1), Knowledge Pump, Swiss-Prot/MedLine...

Les applications de la catégorisation

- Aide à l'analyse des problèmes rencontrés sur le matériel Xerox
 - 1000 à 1500 anomalies au niveau 3 de support technique par mois
 - Analyse manuelle d'un échantillon (100 à 200 documents)
 - Analyse et répercussion sur le design et la maintenance de produits
 - **Problème:** *tous les problèmes découverts ne sont pas traités*

- Utilisation de la catégorisation
 - Définition de catégories (matériel, logiciel, scanner, impression, autres ...)
 - Utilisation du Xerox Categorizer pour affecter les documents décrivant les problèmes dans une catégorie
 - Analyse des catégories, groupement de problèmes, action sur le design et la maintenance
 - ◆ 90 % des documents sont bien classés, malgré des documents mal rédigés
 - ◆ Application aux Document Centre 555/545/535 , extension aux WorkCentre et Work Centre Pro 35/45/55/65/75/90 et WorkCentre 32/40
 - ◆ 2000 problèmes analysés en 10 heures, contre 200 en 40 heures



Applications futures



Le Copieur Intelligent

- ❑ Aide a la compréhension
 - Traduction en contexte des mots les plus difficiles à comprendre
- ❑ Résumé automatique de texte
 - Extraction des groupes nominaux (parsing)
 - Analyse statistique, conservation des phrases contenant les mots les plus fréquemment utilisés

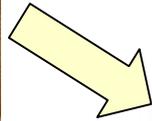


CopyFinder

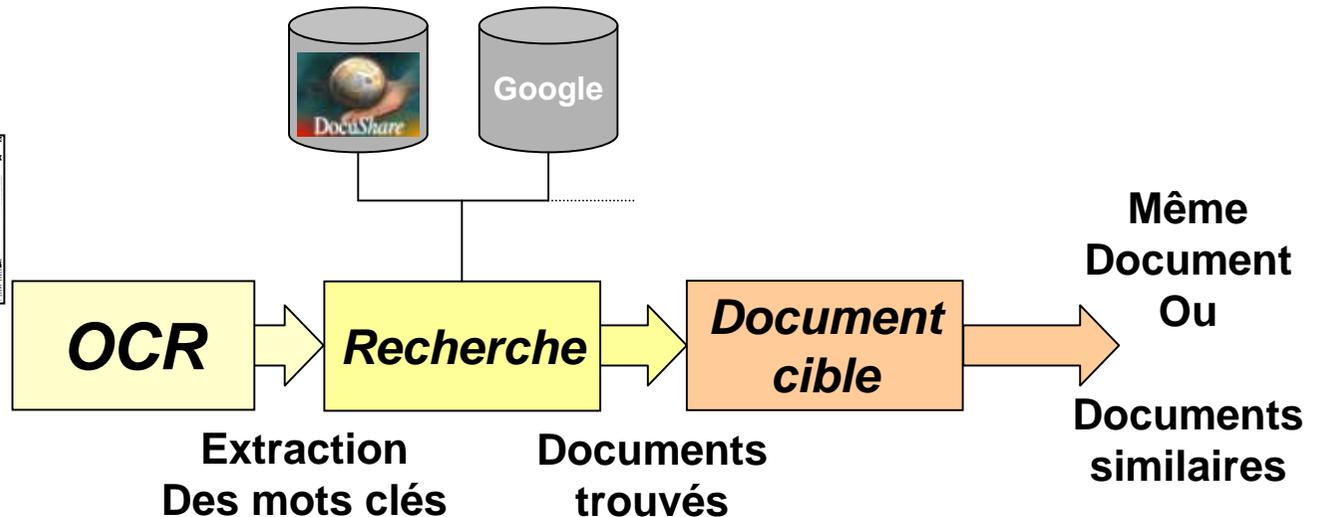
“Où se trouve la version électronique de ce document?”



**Document
Papier**



Image

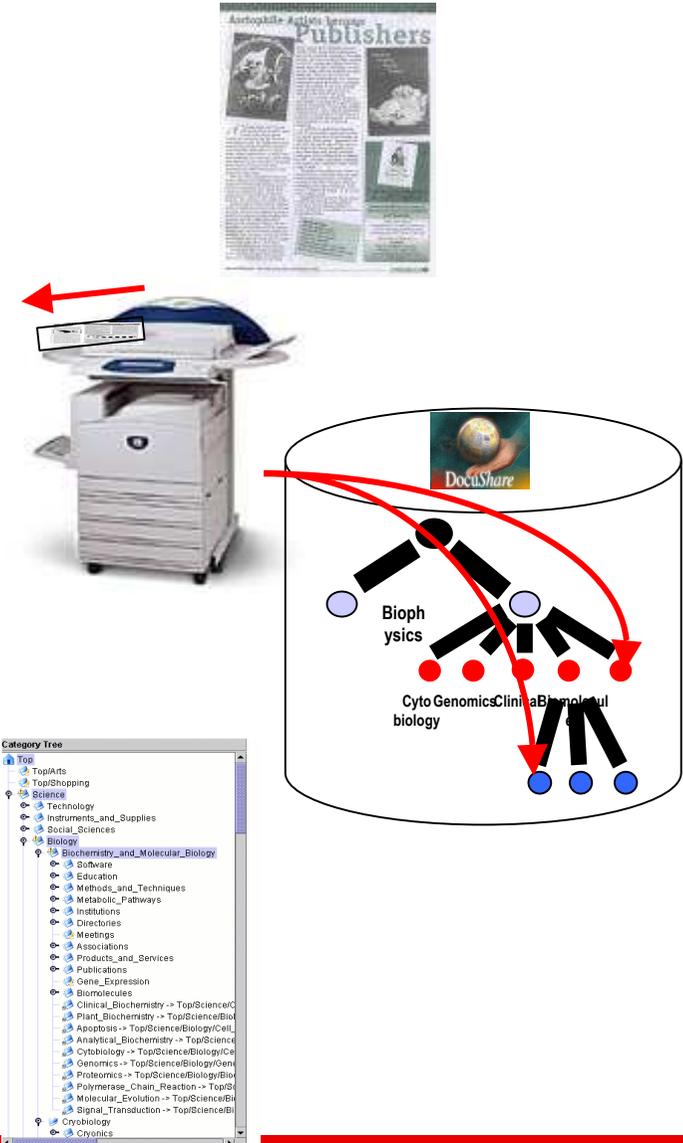


Exemples de scénarios

- Trouver la dernière version électronique d'un document (présentation, formulaire, ...)
- Trouver des documents similaires (article, brevets, ...)

Categorizer

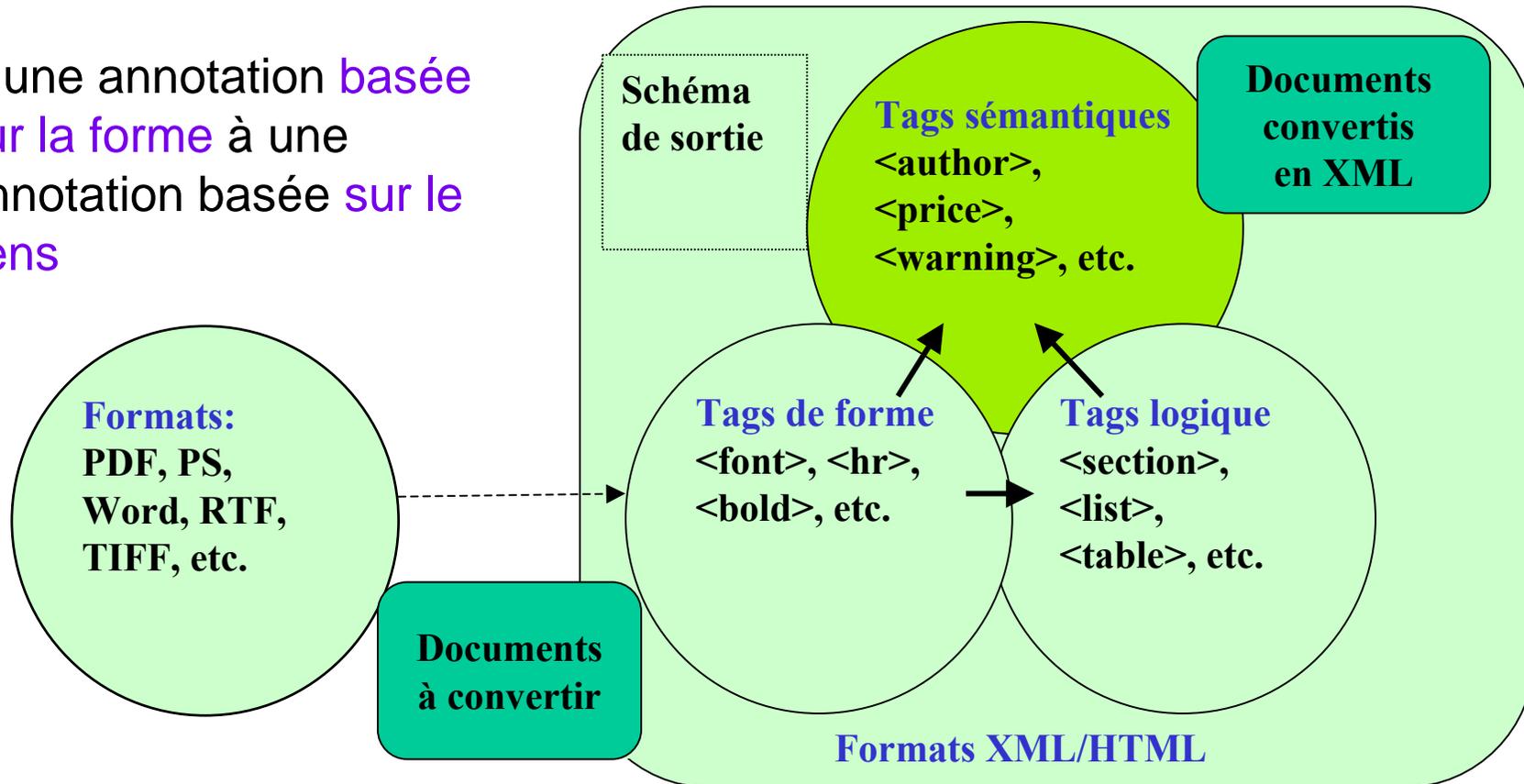
- ❑ **“Je veux organiser mes documents et ceux de ma société efficacement”**
- ❑ **Organisation simple des document en collections: indexation, catégorisation et routage des documents de manière automatique**
- ❑ **Applications**
 - Catégorisation de brevets et des documents associés
 - Catégorisation de documents CRM
- ❑ **Technologies clés**
 - Apprentissage à partir de documents déjà classifiés
 - Proposition de nouvelles catégories
 - Catégorisation hiérarchique et multi-classe



Conversion de documents structurés

Nous nous intéressons aux 3 facettes d'un document structuré
Mise en page, structure logique et sémantique

D'une annotation basée sur la forme à une annotation basée sur le sens



SmartTagger

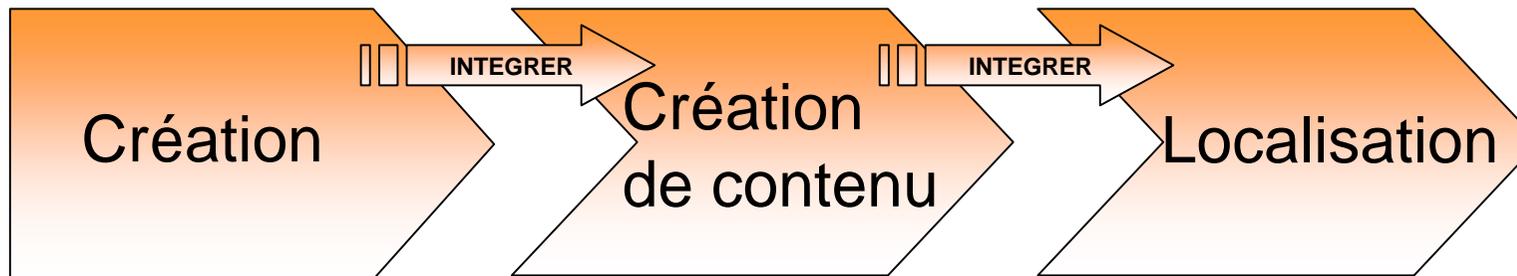
- ❑ Les challenges
 - Transformer rapidement des documents existant (texte ou papier) en **XML**
 - Réduire le coût d'un travail long et fastidieux

- ❑ Les modules Smart Tagger
 - Tagging semi-automatique
 - ◆ Des documents déjà taggés sont analysés et produisent des règles de marquage
 - ◆ Ces règles sont appliquées pour transformer les documents en XML, en taggant les données reconnues
 - ◆ Ces deux actions se font sous le contrôle de l'auteur
 - Aide a la rédaction
 - ◆ Des documents existant sont analysés et produisent des règles de suggestion
 - ◆ Les règles sont utilisées pour suggérer à l'auteur d'utiliser certaines parties de l'arbre XML plutôt que d'autres lors de la rédaction du document

- ❑ Intérêt de SmartTagger
 - Le tagging semi-automatique permet d'accélérer la conversion de documents en XML tout en limitant les erreurs
 - L'aide à la rédaction accélère la rédaction de documents, en proposant à l'auteur les règles de structures les plus probables dans le document



Création, création de contenu et localisation



Conception du document

- Conformité avec la marque
- Conception ciblée
- Mondialisation et besoins en localisation
- Guide de style modèle
- Images numériques
- Outils d'application

Création de contenu

- **Structure**
- **Guide du rédacteur**
- Développement du glossaire
- Décompte mot réduit
- **Réduction traduction**
- Cohérence
- **Anglais simplifié**
- Réutilisation des informations
- Mise à jour facile des informations patrimoniales
- **Contrôle des versions**

Traduction

- **+ de 40 langues**
- Traducteurs locaux
- Connaissance du domaine
- **Gestion de mémoires de traduction, pour un taux optimal de réutilisation**
- **Terminologie cohérente**

Édition

- Conventions spécifiques selon le pays
- Mise en forme
- Création de documents PDF et eBook