



HAL
open science

HENOCH, un outil d'analyse de corpus d'information scientifique et technique

Luc Grivel

► **To cite this version:**

Luc Grivel. HENOCH, un outil d'analyse de corpus d'information scientifique et technique. L'information scientifique et technique et l'outil Internet, N° spécial du Microbulletin thématique du CNRS, 1999. sic_00000467

HAL Id: sic_00000467

https://archivesic.ccsd.cnrs.fr/sic_00000467

Submitted on 19 Jun 2003

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

HENOCH, un outil d'analyse de corpus d'information scientifique et technique

Luc GRIVEL

Chef de projet

Unité recherche et Innovation

Institut de l'Information Scientifique et Technique (INIST), Centre National de la Recherche Scientifique (CNRS) 2, allée du Parc de Brabois, 54 514 Vandoeuvre-lès-Nancy (France)

email: grivel@inist.fr

Résumé

Le processus d'analyse de l'information est un mélange d'exploration informelle intuitive et d'exploitation méthodique de l'information élaborée par différents outils d'analyse. Cet article montre par un jeu de questions-réponses comment un hypertexte généré à partir de techniques d'analyse de données textuelles permet à ses usagers, par exemple un chercheur, de découvrir les thématiques à la frontière de son domaine de recherche, les équipes qui travaillent sur le même sujet que lui, des revues dans lesquelles publier, des congrès dans lesquels publier et auxquels assister. L'utilisateur dispose ainsi de plusieurs modes de navigation conviviaux lui permettant de satisfaire ses multiples besoins et objectifs, ici, la réalisation d'un rapport de tendance sur les plantes transgéniques.

I Introduction

Henoch est le résultat de travaux menés au sein de l'Unité Recherche et Innovation de l'INIST. Cet article est structuré comme un mode d'emploi. Il définit le public visé par cet outil, son organisation des données selon des principes infométriques pour répondre aux besoins énoncés, puis tente de montrer, par un jeu de questions-réponses, comment se système peut aider des organisations (laboratoire, entreprises) dans leur travail d'exploration et d'analyse de l'information scientifique relative à leur domaine d'activité.

1 A qui s'adresse HENOCH ?

HENOCH s'adresse aux chercheurs, veilleurs, spécialistes d'un domaine scientifique, technique ou économique non nécessairement professionnels de la documentation ou de l'informatique, documentalistes qui, **sur un sujet ou un domaine donné**, veulent, à partir des bases bibliographiques PASCAL et FRANCIS¹ :

- avoir une vue d'ensemble,
- suivre et analyser l'évolution thématique, identifier des relations inter-thèmes non explicites,
- repérer l'émergence de nouveaux thèmes de recherche ,
- identifier et regrouper les acteurs, les institutions, leurs vecteurs de communication (thèses, rapports, monographies, périodiques) par thèmes
- évaluer le positionnement thématique d'un acteur, d'une institution, d'un pays, d'un vecteur de communication (périodique, congrès, ...).

Autrement dit explorer et analyser l'information relative à leur sujet de préoccupation (un corpus bibliographique) pour, par exemple:

¹ HENOCH peut fonctionner à partir de données provenant d'autres bases, mais seules des données provenant de nos bases seront accessibles par Internet.

- avoir une première approche d'un sujet de recherche,
- orienter des recherches,
- identifier des technologies émergentes,
- évaluer les résultats d'une équipe de recherche,
- établir un partenariat, ...

Habituellement un corpus bibliographique sur un sujet ou un domaine donné peut représenter quelques milliers de références qu'il est exclu de parcourir séquentiellement. Dans HENOCH, un tel corpus est structuré selon des principes infométriques de manière à constituer une bases de données dites infométriques, exploitables pour l'analyse de l'information.

2 Qu'est ce qu'une base de données infométriques, à quoi ça sert ?

Pour permettre cette analyse de l'information, HENOCH exploite des indicateurs. Ces indicateurs sont le résultat d'un ensemble de traitements linguistiques et statistiques (classification et cartographie) appliqués à des données structurées de type références bibliographiques ou brevets représentatifs d'un domaine

Ce sont :

1. les mots-clés comme indicateurs de la connaissance véhiculée par le document, associés aux références bibliographiques de façon manuelle ou assistée par ordinateur ;
2. les classes comme indicateurs des thèmes ou centres d'intérêt autour desquels s'agrège l'information (articles, auteurs, institutions, périodiques) ;
3. et enfin, la carte comme indicateur stratégique de la position relative des thèmes dans l'espace de connaissance couvert par les documents analysés.

En conclusion, nous définissons une base de données infométriques comme rassemblant et structurant ces informations en les associant aux données bibliographiques, donnant ainsi la possibilité de mettre en relation tout élément constitutif d'une référence bibliographique (auteur, titre, source, affiliation,...) avec les thèmes (classes ou clusters) obtenus par classification automatique.

Les caractéristiques concernant la classification et la cartographie sont disponibles à l'adresse de l'*Unité Recherche et Innovation* "<http://www.inist.fr/pri/pri.htm>" et dans un article publié dans la revue électronique SOLARIS accessible également sur Internet (<http://www.info.unicaen.fr/bnum/jelec/Solaris/d02/2grivel.html>).

Un compte rendu d'un séminaire tenu à l'ADEST le 9.12.97, <http://www.upmf-grenoble.fr/adept/seminaires/francois.htm> fait le point sur les éléments techniques composant la plateforme infométrique.

3 Architecture informatique

D'un point de vue informatique, HENOCH réalise une passerelle entre trois éléments:

- un système infométrique (extracteur terminologique, moteurs de classification et de cartographie),
- un système de gestion de bases de données relationnel (ici Oracle),
- un navigateur sur le Web,

en assurant deux fonctions principales :

- alimenter le SGBD à partir de documents SGML produits par le système infométrique,
- établir une interface WWW-SGBD par une passerelle qui sait donc se connecter au SGBD, soumettre des requêtes SQL à partir d'un modèle de page HTML incluant des requêtes SQL, récupérer le résultat et le mettre au format HTML conformément au modèle, et enfin se déconnecter.

Le cadre informatique de ce développement a été plus précisément décrit dans [1,2,3].

II Comment HENOCH permet d'explorer et d'analyser l'information scientifique et technique sans avoir à faire l'apprentissage de commandes complexes ?

L'information est organisée sous la forme d'un hypertexte basée sur une métaphore cartographique. Ainsi HENOCH dispose d'outils de navigation qui permettent d'éviter le phénomène de désorientation commun aux hypertextes. Pour naviguer, l'utilisateur dispose d'une carte, d'une "boussole" pour orienter sa carte (sa connaissance du domaine) et de méthodes pour faire le point, connaître son positionnement et celui des autres. C'est le rôle joué par les indicateurs infométriques.

HENOCH propose deux types de navigation complémentaires en exploitant les indicateurs infométriques :

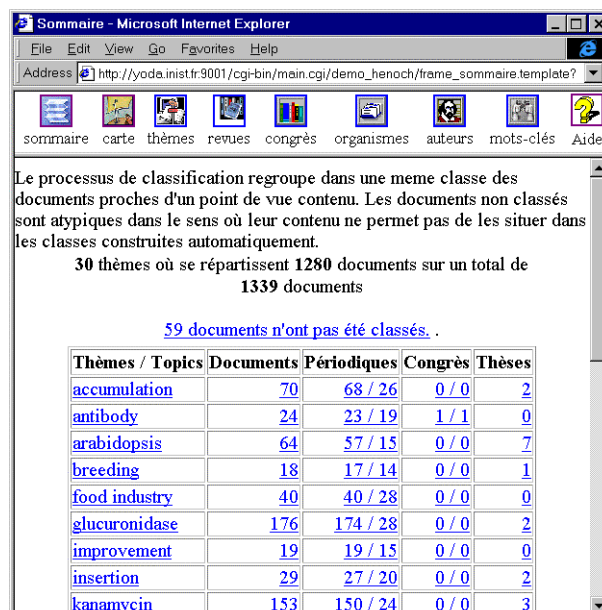
- une exploration intuitive basée sur la carte thématique permettant d'accéder rapidement à des listes pondérées de mots-clés, auteurs, affiliations, sources pour chaque thème, puis de naviguer vers les documents associés à chaque élément de ces listes.
- des fonctions de recherche basées sur ces indicateurs permettent par exemple de savoir dans quelles thèmes un organisme est positionné, le nombre de documents qui est à l'origine de ce positionnement dans le corpus pour chaque thème, puis de naviguer vers ces documents.

L'interface d'HENOCH obéit au principe des interfaces métaphoriques, c'est à dire qu'elle permet à l'utilisateur de travailler sans nécessiter l'apprentissage fastidieux de procédures et de commandes.

Nous prendrons comme exemple un corpus issu de la base Pascal (1 339 enregistrements (production Pascal 1995-96) qui a été utilisé par Le Bureau Van Dijk et l'INIST afin de réaliser un rapport de tendance sur les plantes transgéniques.

L'outil de classification et cartographie qui a été employé est NEURODOC. HENOCH, en organisant les résultats du programme précédent dans un SGBD, a permis aux équipes BVD et INIST de réaliser l'analyse de l'information collectée à partir de son interface WWW dont nous illustrons ici les fonctionnalités.

Après s'être connecté à HENOCH et avoir choisi son corpus de travail, l'utilisateur se trouve face à une page contenant une barre de menu principal et la liste de thèmes qui ont été constitués automatiquement.



The screenshot shows a web browser window titled "Sommaire - Microsoft Internet Explorer". The address bar contains the URL: http://yoda.inist.fr:9001/cgi-bin/main.cgi/demo_henoch/frame_sommaire.template?. The browser's menu bar includes File, Edit, View, Go, Favorites, and Help. Below the menu bar is a toolbar with icons for Home, Back, Forward, Stop, Refresh, Print, and Help. A main menu bar contains buttons for "sommaire", "carte", "thèmes", "revues", "congrès", "organismes", "auteurs", "mots-clés", and "Aide". The main content area contains the following text:

Le processus de classification regroupe dans une meme classe des documents proches d'un point de vue contenu. Les documents non classés sont atypiques dans le sens où leur contenu ne permet pas de les situer dans les classes construites automatiquement.

30 thèmes où se répartissent 1280 documents sur un total de 1339 documents

[59 documents n'ont pas été classés.](#)

Thèmes / Topics	Documents	Périodiques	Congrès	Thèses
accumulation	70	68 / 26	0 / 0	2
antibody	24	23 / 19	1 / 1	0
arabidopsis	64	57 / 15	0 / 0	7
breeding	18	17 / 14	0 / 0	1
food industry	40	40 / 28	0 / 0	0
glucuronidase	176	174 / 28	0 / 0	2
improvement	19	19 / 15	0 / 0	0
insertion	29	27 / 20	0 / 0	2
kanamycin	153	150 / 24	0 / 0	3

Figure 1

La barre de menu principale est composée de 8 boutons plus une aide en ligne. Les trois premiers boutons permettent d'accéder respectivement à la liste des bases de données infométriques

disponibles, à la carte thématique, au tableau des thèmes (Fig. 1), les 5 suivants correspondent aux fonctions de positionnement. Nous allons voir leur utilisation, en commençant par la carte thématique qui est un peu le poste de pilotage de la navigation tandis que le tableau des thèmes présenté en premier lieu a surtout pour objectif de fournir à l'utilisateur des informations quantitatives sur les résultats de classification.

Le lecteur peut notamment se faire rapidement une idée de la distribution des documents dans les thèmes et par type de document. Par exemple, le thème « Accumulation » rassemble 70 documents dont 68 répartis dans 26 périodiques et 2 thèses.

1 Comment naviguer depuis la carte thématique ?

Depuis la carte de thématique (Fig.2), l'analyste peut accéder à différents types d'informations pertinentes pour un thème et visualiser très rapidement les éléments les plus représentatifs du thème, son organisation.

Pour l'utilisateur, la procédure est la suivante:

1) Choisir, dans la table de boutons radio à gauche de la carte, un type destination

- MC --> une liste triée de mots-clés,
- TI --> une liste triée de titres des documents associés,
- AU --> une liste triée des auteurs,
- AF --> une liste triée des affiliations des auteurs,
- SO --> une liste triée des des modes de publication des auteurs (les sources),

2) Choisir un thème sur la carte (en cliquant sur le nom du thème ou sur le cercle noir positionné avant le nom.)

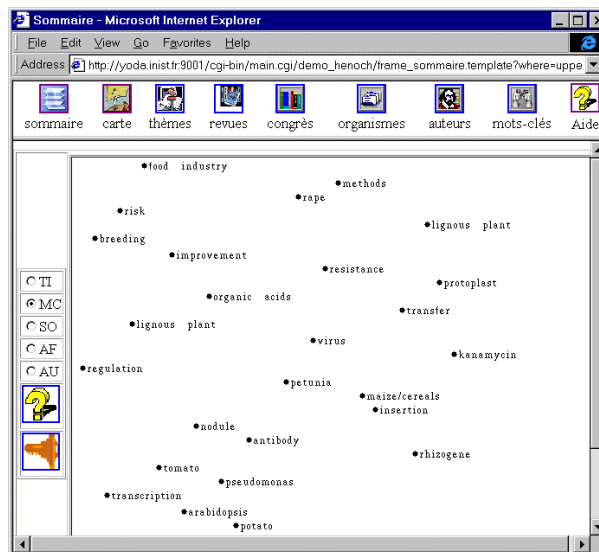


Figure 2

2 Comment analyser la carte ?

L'analyse de la carte dépend de la méthode de cartographie employée. Dans le cas présent, l'algorithme de cartographie, l'analyse en composante principale, réduit le nombre de dimensions de données statistiques multidimensionnelles, de telle manière que deux thèmes au contenu similaire soient relativement proches sur la carte. L'interprétation de la carte est délicate car le lecteur doit s'appuyer sur les connaissances qu'il a du domaine et sur une observation attentive de chaque thème (opération illustrée dans 4.3) L'interprétation de la carte correspondant à ce jeu de données est disponible dans le rapport de tendance qui est commercialisé.

3 Comment observer l'organisation thématique ?

Observer l'organisation d'un thème suppose de pouvoir décrire son contenu, les auteurs qui travaillent sur ce thème, leur organisme d'appartenance, leur modes de publication, les relations avec les autres thèmes.

L'organisation du thème est décrite par :

- une liste de mots-clés ordonnés selon leur importance pour la définition du thème, le mot-clé de plus fort poids donnant son nom au thème
- une liste de titres de documents ordonnés selon le même critère,
- une liste des affiliations, ensemble des affiliations des auteurs des documents du thème, ces derniers sont triés selon leur fréquence dans le thème,
- une liste d'auteurs : ensemble des auteurs des documents du thème, ces derniers sont triés selon leur fréquence dans le thème et
- une liste de sources : ensemble des titres de revues où sont édités les documents du thème, ces dernières sont triés selon leur fréquence dans le thème.

On peut accéder à chacune de ces informations par la carte. Il existe une fenêtre par type d'information (mots-clés, titres des documents, auteurs, affiliations, sources) associé à chaque thème. Une barre de menu commune à toutes ces fenêtres et locale au thème permet également d'accéder à ces différents types d'informations sans repasser par la carte et informe sur la quantité d'information qui est agrégée autour du thème.

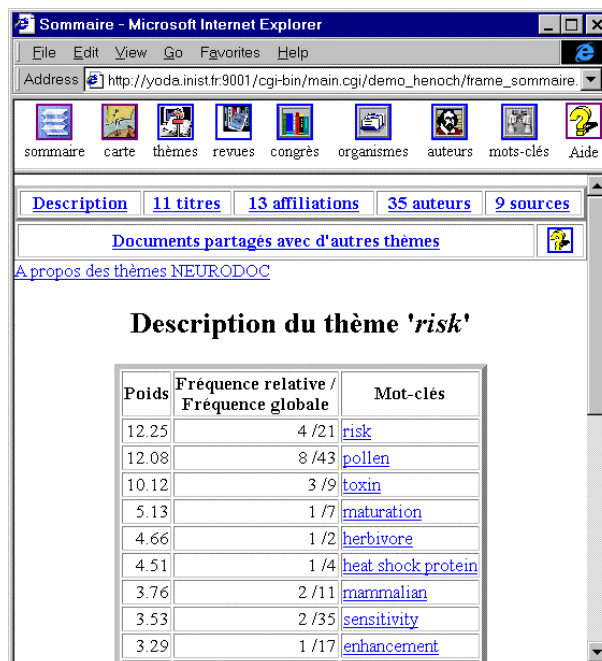
3.1 Comment se faire une idée du contenu d'un thème ?

Deux moyens différents sont mis à disposition de l'utilisateur via le menu local au thème:

- la liste triée de mots-clés,
- la liste triée des documents du thème.

3.1.1 La liste triée de mots-clés (Fig.3)

Le lien **Description** renvoie donc à la liste triée de mots-clés ordonnés selon leur importance pour le thème. Chaque mot-clé est précédé de son poids et de sa fréquence locale et globale. Le mot de poids le plus élevé donne par défaut son nom au thème.



Poids	Fréquence relative / Fréquence globale	Mot-clés
12.25	4 / 21	risk
12.08	8 / 43	pollen
10.12	3 / 9	toxin
5.13	1 / 7	maturation
4.66	1 / 2	herbivore
4.51	1 / 4	heat shock protein
3.76	2 / 11	mammalian
3.53	2 / 35	sensitivity
3.29	1 / 17	enhancement

Figure 3

La colonne Fréquence globale donne le nombre total de documents indexés par chaque mot du thème dans le corpus. Il ne s'agit pas de la fréquence du mot dans les documents relatifs au thème, fréquence dite locale. Leur rapport donne le pourcentage relatif de documents indexé par le mot dans le thème.

L'utilisateur peut donc visualiser rapidement les mots-clés liés à ce thème triés par degré de pertinence par rapport au thème et utiliser chacun des mots pour effectuer des recherches locales au thème « risk ». En effet chaque mot-clé donne accès à la liste des titres des documents du thème qui sont indexés par ce mot dans le thème (Fig.4). Ce qui peut lui donner des idées sur des termes liés à « risk » et l'aider à formuler de nouvelles hypothèses. Exemple : «Pollen » et « heat shock protein ». Sur 21 documents indexés par « risk », 4 seulement figurent dans le thème, alors que le thème ne comporte en tout que 11 documents correspondant globalement à ce profil thématique. Ainsi donc le mot « risk » a été employé dans différents contextes. Pour les 4 documents en question, il s'agit plutôt des risques pour l'environnement à travers les pollens.

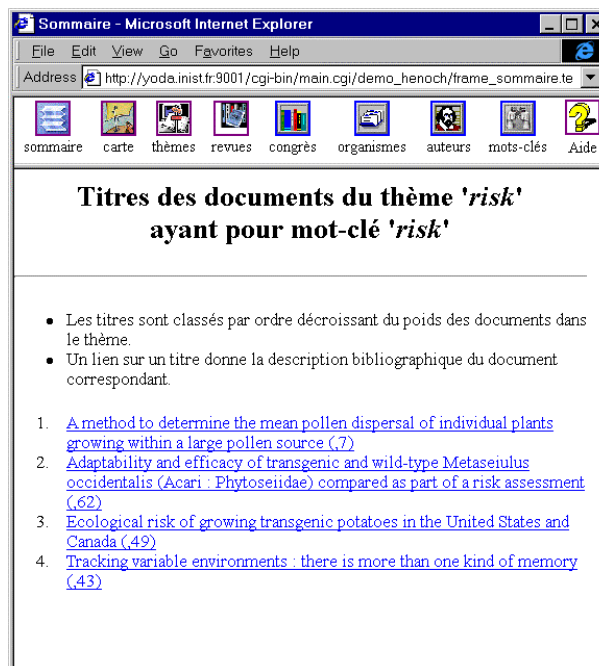


Figure 4

3.1.2 La liste triée des titres de document (Fig. 5)

Le lien **titres** renvoie à la liste de titres documents qui est ordonnée selon la pertinence (le poids) du document pour le thème. La date de publication et le poids de chaque document encadrent le titre en donnant accès à la description bibliographique du document correspondant.

Ces deux éléments (date et poids) permettent au lecteur de se faire une idée de l'âge moyen d'un thème, puisque les documents les plus pertinents (d'un point de vue statistique) pour le thème sont classés en tête. A partir de ces indicateurs, l'utilisateur peut faire des hypothèses sur les thèmes en vogue (nombre important de documents, date récente pour l'essentiel des documents), en perte de vitesse (faible nombre de documents, date ancienne), en émergence (faible nombre de documents, date récente). Bien entendu, c'est l'expertise du lecteur, ses recoupements avec d'autres informations qui lui permettront d'affirmer ou d'infirmer la validité de ses hypothèses.

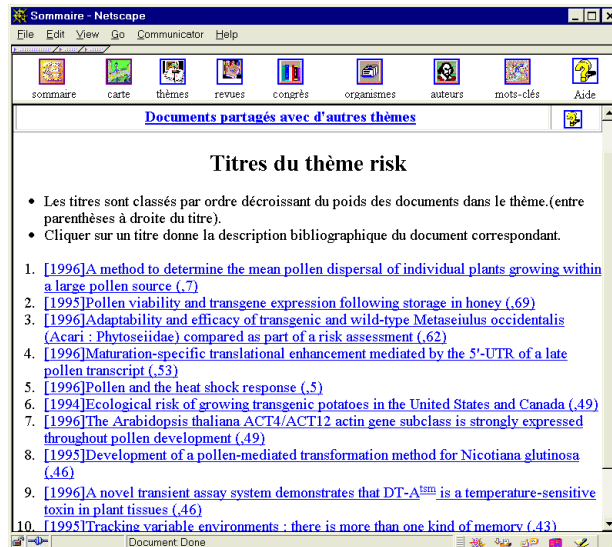


Figure 5

3.2 Comment se faire une idée de l'environnement institutionnel d'un thème ?

Le lien **affiliations** renvoie à la liste des affiliations des auteurs des documents du thème (Fig. 6), triées selon leur fréquence dans le thème. Ceci donne une indication de la productivité de l'institution dans le thème.

Chaque affiliation est précédée de sa fréquence dans le thème et donne accès à la liste des titres des documents du thème écrits par des auteurs membres de cette affiliation.

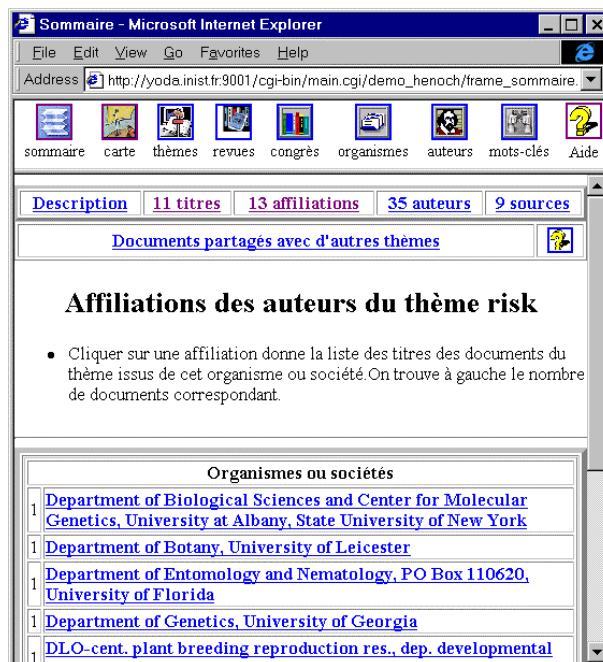


Figure 6

3.3 Comment se faire une idée des acteurs principaux du thème ?

Le lien **Auteur** renvoie à la liste des auteurs des documents du thème, triés selon leur fréquence dans le thème (Fig. 7).

Chaque auteur est précédé de sa fréquence dans le thème (sa productivité dans le thème) et donne accès à la liste des titres des documents du thème écrits par l'auteur.

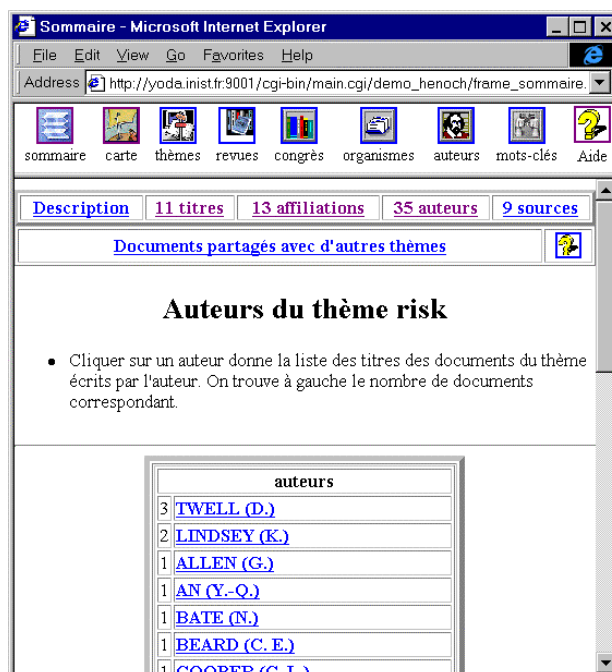


Figure 7

3.4 Comment se faire une idée des modes de publications des auteurs du thème ?

Le lien **source** renvoie aux listes des titres de revues, congrès ou universités de soutenance pour les thèses, triées selon leur fréquence dans le thème, où sont publiés les documents du thème (Fig. 8). Les sources sont précédées de leur fréquence et donnent accès la liste des titres des documents du thème publiés par cette source. La distribution selon le type de source (revue, congrès), puis pour chaque type de source selon le journal ou le nom du congrès permet de qualifier les modes de communications privilégiés des auteurs.

Exemple

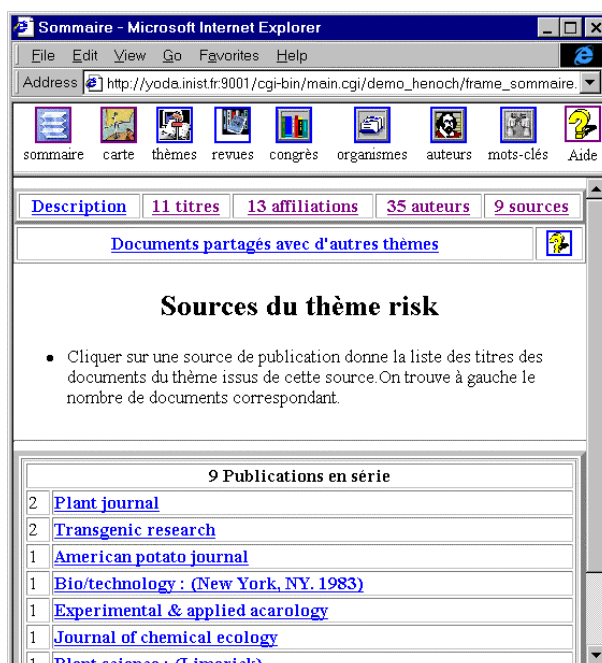


Figure 8

3.5 Comment se faire une idée des relations qu'un thème entretient avec les autres thèmes ?

Le lien **Documents partagés** permet d'accéder en premier lieu à la distribution des documents dans les autres thèmes, puis pour chaque thème, à la liste des titres des documents communs (Fig. 9). Le titre renvoie à la description bibliographique du document correspondant.

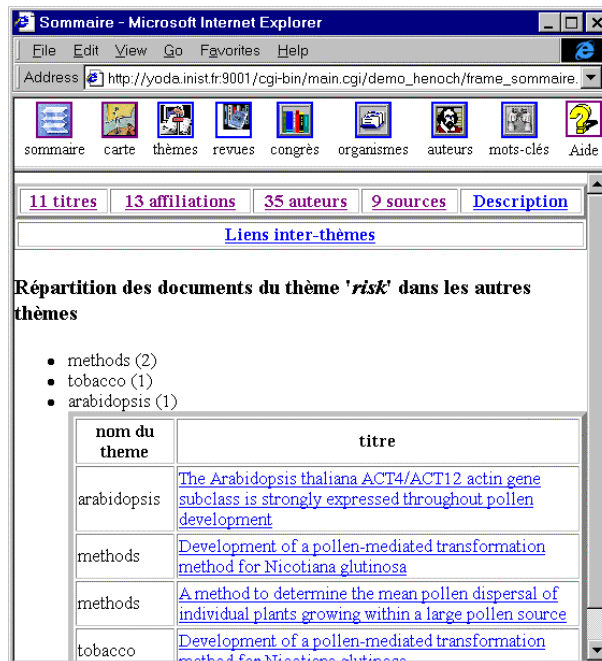


Figure 9

4 Comment utiliser la description bibliographique d'un document ?

La figure 10 montre un exemple de référence sélectionné à partir de la liste des documents d'un thème. Les flèches de navigation donnent la possibilité de naviguer transversalement vers des documents voisins d'un poids immédiatement supérieur ou inférieur dans le thème et la possibilité d'accéder aux mots clés décrivant le thème ainsi que de revenir à la liste des documents du thème. Un document peut en effet se trouver dans plusieurs thèmes, comme dans le cas ci-dessous.

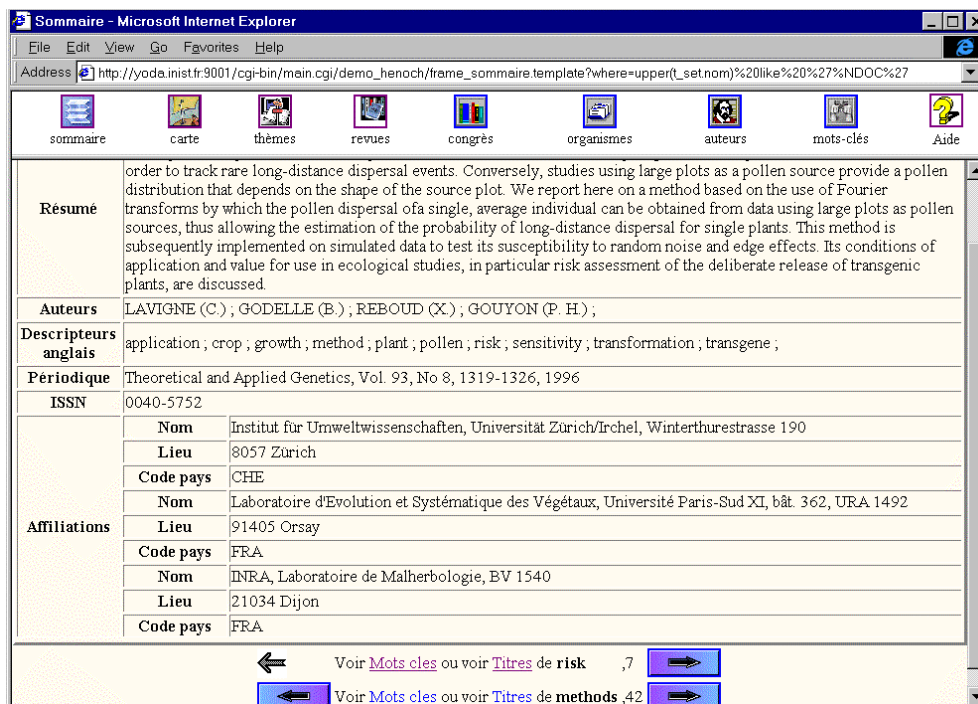


Figure 10

L'accès à la description bibliographique du document permet de compléter les observations effectuées et peut suggérer de nouvelles voies de navigation. Ici, supposons que l'utilisateur

s'interroge sur le positionnement thématique de la revue « Theoretical and Applied Genetics » puis sur le positionnement des 3 organismes qui coopèrent et enfin se faire une idée des contextes dans lequel le mot-clé « risk » est employé. La section suivante illustre comment ce besoin peut être satisfait.

5 Comment effectuer le positionnement d'un périodique (d'un auteur, d'une affiliation, d'un mot-clé) dans les thèmes ?

Objectif

Le but est de savoir dans quelles thèmes un périodique (un congrès, un auteur, une affiliation, un mot-clé) est positionné, le nombre de documents qui est à l'origine de ce positionnement dans le corpus pour chaque thème, puis de naviguer vers ces documents.

Procédure à suivre

Il faut d'abord savoir si le périodique (le congrès, l'auteur, l'affiliation, le mot-clé) dont on cherche le positionnement thématique est bien dans la liste des périodiques (des congrès, des auteurs, des affiliations, des mots-clés) et sous quelle(les) forme(s) il a été saisi.

Chacun des boutons 4 à 8 qui figurent de manière permanente dans le menu principal, a pour effet de d'afficher une boîte de sélection composée d'une zone de saisie et d'un bouton intitulé "filtrer" qui permet de faire une recherche sur la liste correspondant à l'intitulé du bouton (revues, congrès, organismes, auteurs, mots-clés).

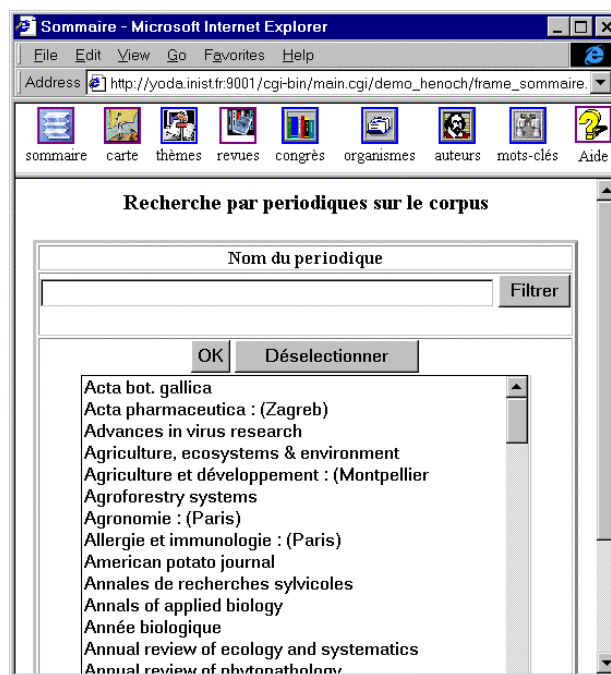


Figure 11

Quand le nombre d'éléments est important, les lister comme dans la figure 11 peut prendre du temps. C'est pourquoi, dans le cas d'une liste de plus de 1000 éléments, HENOCH n'affiche pas directement tous les éléments de la liste. A la place, est proposé la possibilité de filtrer par une expression entrée dans une zone de saisie. Par exemple en entrant les premières lettres de l'élément en utilisant la troncature à droite (le caractère *). Les minuscules et les majuscules ne sont pas différenciés.

Par exemple, *Genetics* signifie tout élément contenant "genetics".

a* donnera tous les éléments de la liste commençant par a.

Si le lecteur veut malgré tout afficher toute la liste, il doit taper * dans la zone de saisie puis cliquer sur le bouton "filtrer".

Ensuite, il faut sélectionner le ou les éléments intéressants dans la liste proposée (titre de périodique, congrès, auteur, affiliation, mot-clé) et valider en cliquant sur le bouton OK.

Le résultat est une distribution des documents relatifs à ce ou ces éléments par thème.

En sélectionnant un nom de thème et en validant par OK, on accède à la liste des titres des documents relatifs à ce thème pour le ou les éléments sélectionnés.

Dans le cas de la revue qui nous intéresse (« Theoretical and Applied Genetics »), le lecteur peut voir (Fig. 12) qu'elle se positionne en premier lieu dans le thème « Protoplasts » qui correspond au transfert de gènes dans les protoplastes et régénération des plantes à partir de cultures de protoplastes.

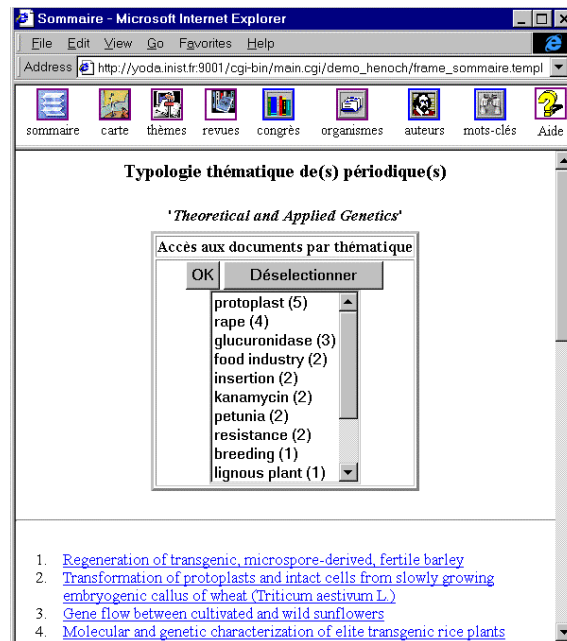


Figure 12

Pour des raisons diverses, les titres de périodique, les noms de congrès ou d'auteurs ou les affiliations peuvent se trouver sous différentes formes lexicographiques. En effet, il n'est pas possible de normaliser de manière fiable ces entités sans utiliser de fichiers d'autorité. Si on prend l'exemple du Laboratoire de Malherbologie, on peut observer qu'il figure dans la base de données sous 3 formes différentes (Fig.13). Ce qui ne pose pas de problème pour HENOCH, car on peut sélectionner plusieurs formes lexicographiques en les considérant comme constituant un objet unique.



Figure 13

Le lecteur peut donc observer le positionnement thématique de ce laboratoire (Fig.14) et lister les documents qu'il a produit. Dans le cas présent, ce laboratoire est à l'origine de trop de documents dans le corpus pour en tirer des conclusions.

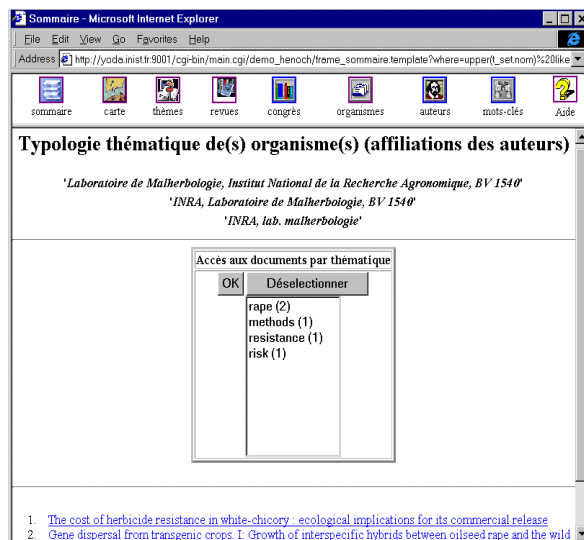


Figure 14

Le même type d'opération peut être effectué pour un mot-clé (Fig.15).

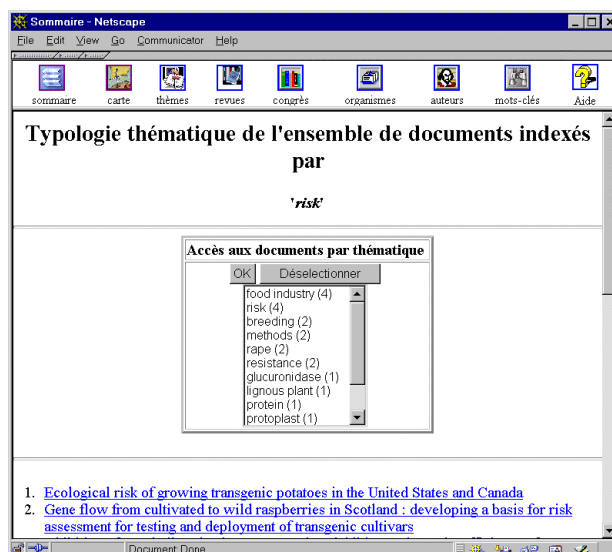


Figure 15

Les documents indexés par le mot-clé « risk » se répartissent entre différents thèmes : « risk » (les risques encourus par les consommateurs et l'environnement), « food industry » (l'agro-alimentaire) ... Le lecteur peut ensuite lister les documents thèmes par thèmes.

III Conclusion et perspectives

Dans sa version prochainement accessible sur Internet, HENOCH, offrira l'accès à un ensemble de corpus bibliographiques constitués à partir des bases PASCAL OU FRANCIS sur quelques sujets sélectionnés par l'INIST comme représentatifs de préoccupations de recherche actuelles.

De l'observation des réactions des utilisateurs dépendront les orientations que prendra l'INIST concernant l'exploitation d'un système tel qu'HENOCH.

Par ailleurs, depuis plusieurs années, l'Unité Recherche et Innovation de l'INIST est sollicité par divers organismes publics et privés qui souhaitent que leurs données (scientifiques et techniques pour

la plupart) soient traitées par des outils linguistiques et statistiques. Ils souhaitent également que ces données et ces résultats soient rendus accessibles, via le réseau, à la communauté de leurs membres à travers une interface conviviale comme celle d'HENOCH.

Parallèlement, on voit se mettre en place, ici ou là, la notion d'observatoire thématique produisant de l'information élaborée partagée par un cercle d'utilisateurs-souscripteurs, les thèmes répondant aux besoins exprimés par un groupe d'utilisateurs, (les membres d'un organisme, une association, ...), ou à l'initiative d'un producteur d'information sur un thème porteur.

On peut donc imaginer dans le futur un lieu virtuel (village Internet ?) où l'information partagée autour d'une thématique serait exclusivement réservée aux membres du cercle, lesquels pourraient également commander les documents en ligne, demander une étude complémentaire ou échanger des informations via un forum spécialement mis à leur disposition. Un système comme HENOCH pourrait donc tout à fait s'intégrer dans un bouquet de services.

BIBLIOGRAPHIE RELATIVE à HENOCH

1. Grivel L., X. Polanco, A. Kaplan "A computer System for Big Scientometrics at the Age of the World Wide Web", *Scientometrics*, vol.40, N°3, 1997, 493-506, 1997, et également in proceedings of the 6th International Conference on Scientometrics and Informetrics, Jerusalem, 131-142, 1997.
2. Grivel L., C. Francois, X. Polanco -"Analyse de l'information par cartographie neuromimétique et requêtes SQL sur le Web", - "4ème Conf. Intern. Hypertextes et Hypermedias : réalisation, outils méthodes", *Hypertextes et Hypermedias*, Editions Hermès, Vol.1, n°2 ,237-248, 1997.
3. Grivel L., X. Polanco, A. Kaplan -"Requêtes et navigation à partir de l'information structurée, le système HENOCH ", *Le Micro Bulletin*, N°70, 493-506, 1997.