

La problématique de la recherche de document

Étienne Vandeput

Journée de formation Form@HETICE

29 février 2008



Le document

- Une notion qui évolue
 - fichier numérique contenant un texte
 - fichier numérique
 - feuille de calcul
 - présentation
 - image
 - ...
 - ensemble d'informations affichables
 - pages Web

La ressource

- Une notion plus large encore
 - document
 - aux différents sens définis dans la diapo qui précède
- Une partie de document
 - un paragraphe
 - un bout de phrase
 - une image
 - un message de courriel
 - ...

Rechercher (retrouver) une ressource

- Un traitement formel qui nécessite
 - le choix de critères
 - la fourniture de valeurs
- Un critère
 - toute information susceptible d'identifier une classe de documents (ressources)

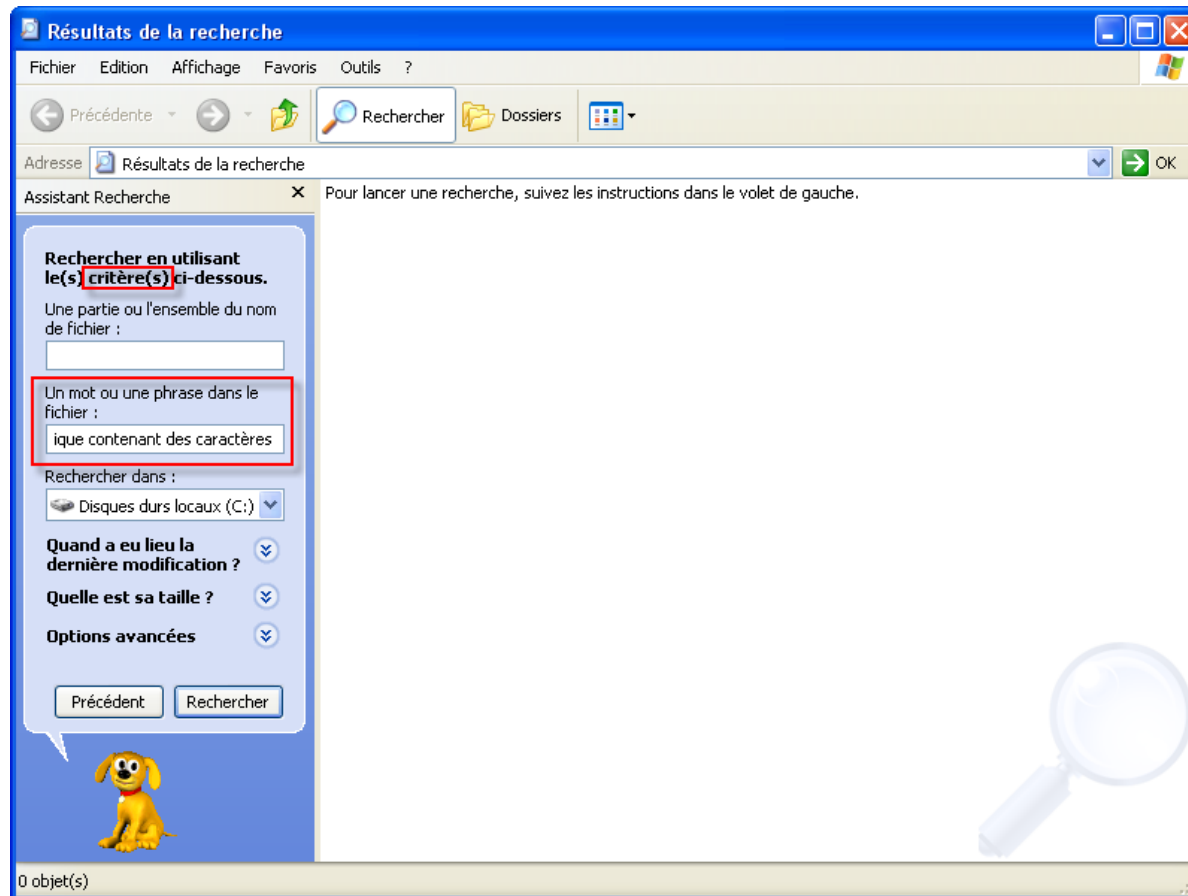
Rechercher (retrouver) une ressource

- Des valeurs pour les critères
 - informations potentielles (rechercher)
 - le contenu de la ressource
 - les métadonnées qui les concernent
 - informations ajoutées (retrouver)
 - les métadonnées issues du marquage ou de la caractérisation de la ressource

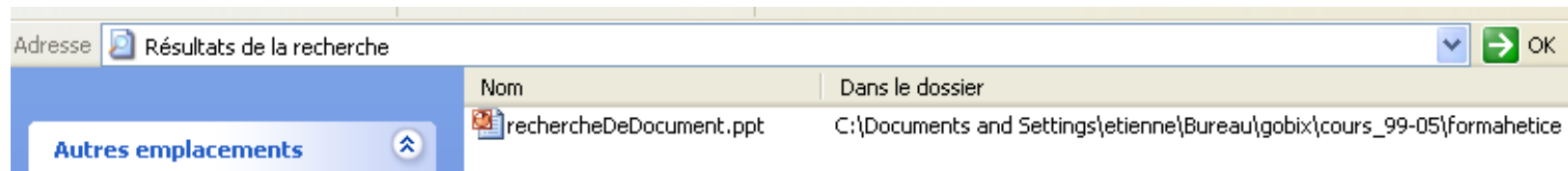
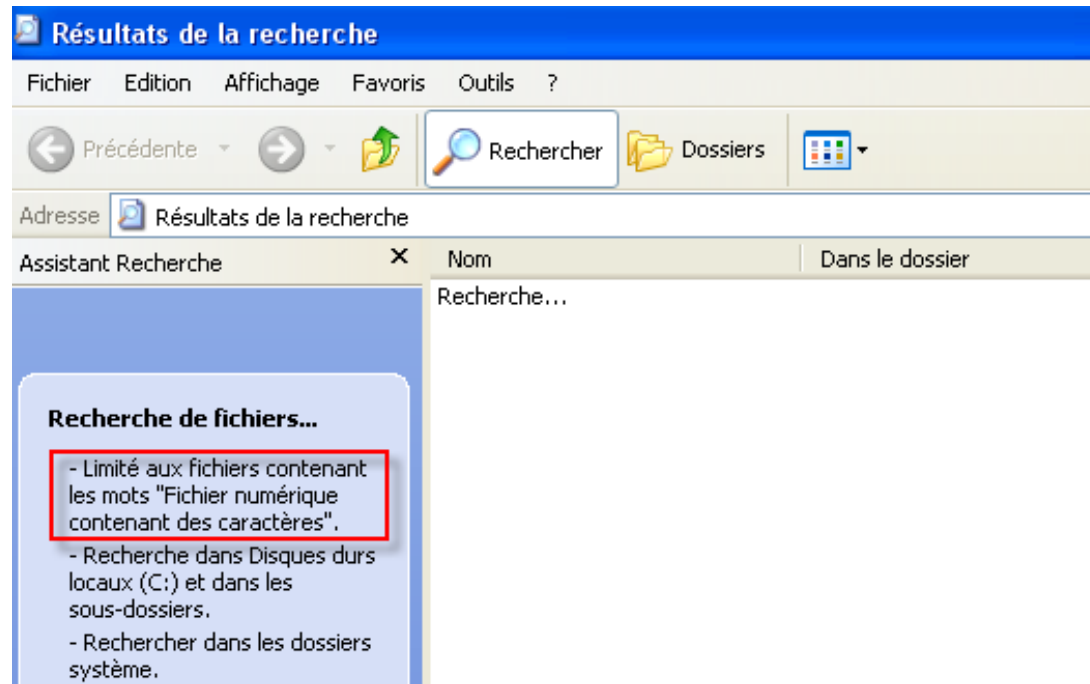
Le contenu de la ressource

- Fichier numérique contenant des caractères
 - toutes les chaînes de caractères qui le composent
 - mots
 - groupes de mots
 - parties de mots
 - mots-clés

Exemple SE



Exemple SE



Exemple Web

Web [Images](#) [Maps](#) [News](#) [Shopping](#) [Gmail](#) [more](#) ▼

Google™

classification bayésienne

Search

[Advanced Search](#)
[Preferences](#)

Web

Results 1 - 10 of about 8,630

[PDF\] DEA IARFA - Techniques du traitement d'images Segmentation par ...](#)

File Format: PDF/Adobe Acrobat - [View as HTML](#)

Classification bayésienne. Méthode supervisée. • Probabilité $P(y, s, \{\omega, i\})$ apprentissage sur un ensemble de données classées ...

www.tsi.enst.fr/~tupin/cours/dea_iarfa/cours_seg.pdf - [Similar pages](#)

[agglomération : Classification Bayésienne](#) - [[Translate this page](#)]

Détection d'agglomérations dans les images SPOT avec utilisation de la morphologie mathématique et de la **classification Bayésienne** [Rapport d'activité ...

www-projet.cnes.fr:8220/agglomeration/cl_bayésienne.html - 4k - [Cached](#) - [Similar pages](#)

[John Libbey Eurotext : Éditions médicales et scientifiques France ...](#)

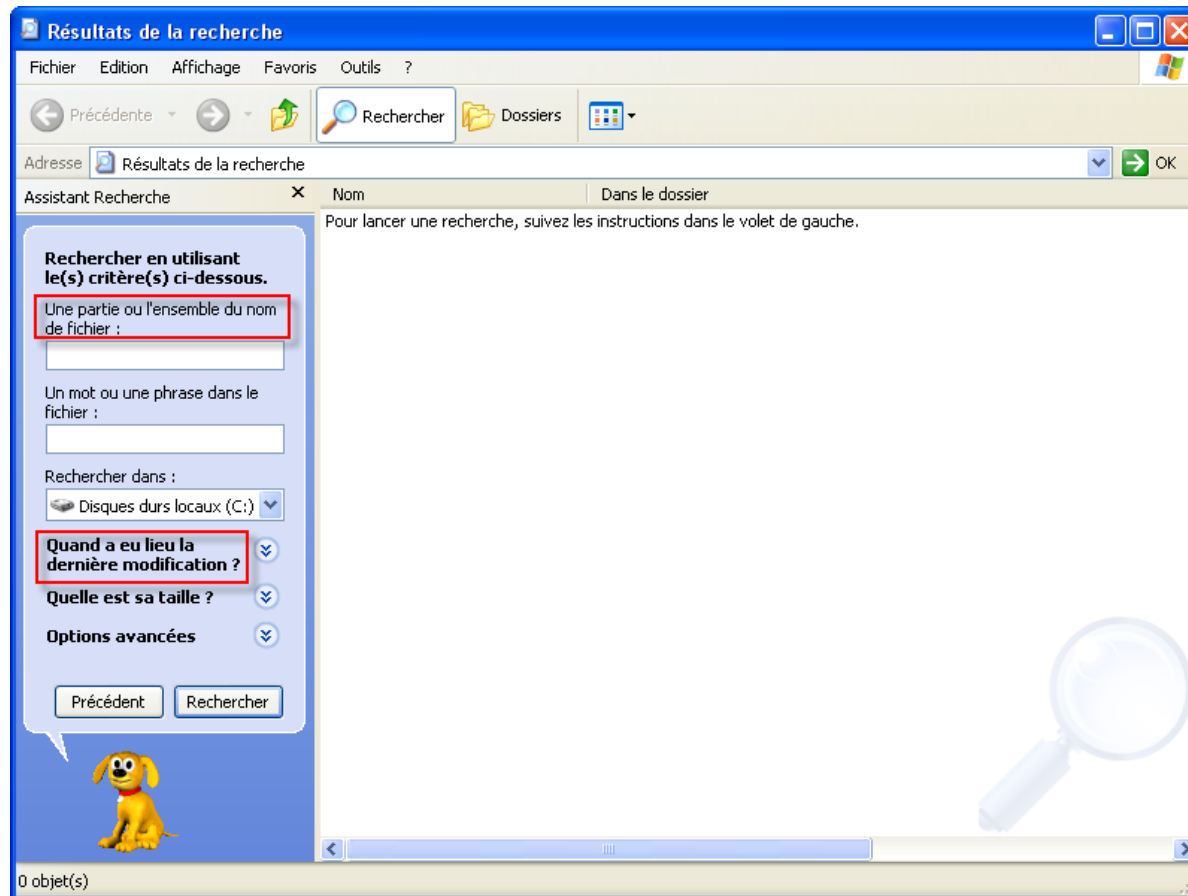
B : de l'ensemble de la région de Ouargla (image entière) après **classification bayésienne** par maximum de vraisemblance (CBMV). ...

www.jle.com/en/revues/agro_biotech/sec/e-docs/00/03/82/B5/article.md?fichier=images.htm -

Métadonnées (gestion du document)

- Fichier numérique géré par le SE
 - nom
 - partie du nom
 - date de création
 - date de dernière modification

Exemple SE

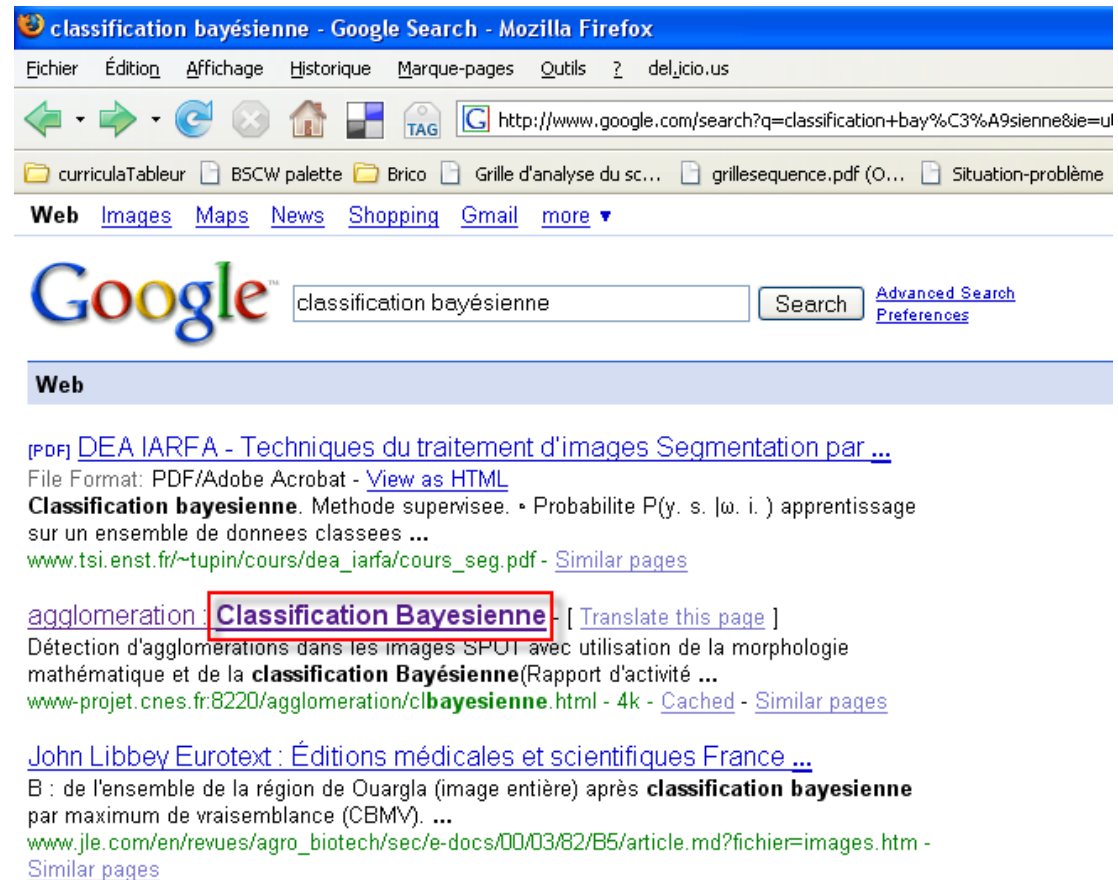


Métadonnées

(caractérisation du document)

- Pages Web
 - titre
 - données « méta »
 - auteur
 - description
 - mots-clés
 - ...

Exemple Web



The screenshot shows a Mozilla Firefox browser window with the following elements:

- Browser Title Bar:** "classification bayésienne - Google Search - Mozilla Firefox"
- Menu Bar:** "Fichier", "Édition", "Affichage", "Historique", "Marque-pages", "Outils", "?", "del_jcio.us"
- Address Bar:** "http://www.google.com/search?q=classification+bay%C3%A9sienne&ie=ul"
- Bookmarks Bar:** "curriculaTableur", "BSCW palette", "Brico", "Grille d'analyse du sc...", "grillesequence.pdf (O...", "Situation-problème"
- Navigation Bar:** "Web", "Images", "Maps", "News", "Shopping", "Gmail", "more ▾"
- Search Interface:** The Google logo is on the left. The search input field contains "classification bayésienne". To the right of the input field is a "Search" button. Further right are links for "Advanced Search" and "Preferences".
- Search Results:** A section titled "Web" contains three search results:
 - Result 1:** A PDF document titled "DEA IARFA - Techniques du traitement d'images Segmentation par ...". Below the title, it says "File Format: PDF/Adobe Acrobat - View as HTML". The main text reads: "Classification bayésienne. Methode supervisee. • Probabilite P(y. s. |ω. i.) apprentissage sur un ensemble de donnees classees ...". The URL is "www.tsi.enst.fr/~tupin/cours/dea_iarfa/cours_seg.pdf" with a "Similar pages" link.
 - Result 2:** A page titled "agglomeration : Classification Bayésienne" (the title is highlighted with a red box). It includes a "[Translate this page]" link. The text describes "Détection d'agglomerations dans les images SPUT avec utilisation de la morphologie mathématique et de la classification Bayésienne(Rapport d'activité ...". The URL is "www-projet.cnes.fr:8220/agglomeration/clbayesienne.html" with "4k - Cached - Similar pages" information.
 - Result 3:** A page titled "John Libbey Eurotext : Éditions médicales et scientifiques France ...". The text reads: "B : de l'ensemble de la région de Ouargla (image entière) après classification bayésienne par maximum de vraisemblance (CBMV). ...". The URL is "www.jle.com/en/revues/agro_biotech/sec/e-docs/00/03/82/B5/article.md?fichier=images.htm" with a "Similar pages" link.

Exemple Web

NOUS AVONS LU sur le Net

Évaluation des compétences en TIC

Par **Étienne Vandeput**, dans Bernard André, Georges-Louis, Baron Éric Bruillard (Dir.), *Premières journées francophones de didactique des progiciels* (10 et 11 juillet 2003).

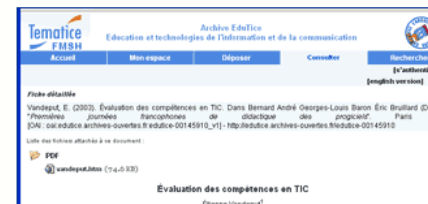
L'auteur analyse quelques épreuves (PCIE, B2i...) censées vérifier les compétences TIC des apprenants. Les évaluations qu'elles proposent sont-elles pertinentes et efficaces ? « Comme toujours, lorsque divers outils donnent l'impression de ne pas convenir totalement à la tâche envisagée, il convient d'en proposer d'autres sur mesure. » Il nous propose donc une démarche effectuée dans le contexte particulier d'un cours pour l'obtention du Diplôme d'Études Spécialisées en Technologie de l'Éducation et de la Formation organisé conjointement par les universités de Liège et de Namur.

Extrait concernant le B2i (niveau 1) français

Rappelons qu'un des
d'apprentissages à des
sans distinction au sein
cas du deuxième module

Il est question de « savo

« ...l'élève doit être cap
d'utiliser à bon escient le



```
<HTML>
```

```
<HEAD>
```

```
<META HTTP-EQUIV="Content-Type" CONTENT="text/html; charset=iso-8859-1">
```

```
<META HTTP-EQUIV="Content-Language" CONTENT="fr">
```

```
<META NAME="Author" CONTENT="EPI">
```

```
<META NAME="Copyright" CONTENT="Association EPI">
```

```
<META NAME="Classification" CONTENT="ENSEIGNEMENT, FORMATION, INFORMATIQUE, TICE">
```

```
<TITLE>Lu sur le Net : Évaluation des compétences en TIC</TITLE>
```

```
<META NAME="Ref-epi" CONTENT="10706t">
```

```
</HEAD>
```

```
<BODY TEXT="#003090" BGCOLOR="#FFFFFF" LINK="#0000EE" VLINK="#0000DD" ALINK="#FF0000">
```

```
<!AK 14/06/2007>
```

Métadonnées (processus de marquage)

- Mots-clés associés à la ressource
 - tagging
 - libre: folksonomies
 - imposé: ontologies

Exemple SweetWiki

The screenshot displays the SweetWiki interface for the 'crifa' page. The page is structured as follows:

- Header:** Sweet Wiki logo and navigation links: Welcome Eva | Home Page | Disconnect.
- Search:** Keyword search and Tag search fields.
- Navigation:** Webs (Main, Documentation, Prive, edit), Pages (Sand Box), and Administration (Manage Ontologies, Reload Coresé).
- Page Content:**
 - crifa** (highlighted in yellow)
 - crifa is a subclass of :
 - ste, ulg, universite, liege, (highlighted in red)
 - in the same category :
 - stecrifa, palette, tools, services, copeit, recherche, recherches, expertise, tool, sweetwiki, cop, cops, elogbook, ergonomie, scapin, bastien, meda, amaya1, thesaurus, (highlighted in red)
 - Pages tagged with crifa :
 - DocumentationHome
 - FirstSteps
 - HowtoUseSW
 - RelatedWorkAndPositioning
 - ReportSweetwiki
 - ReturnOnExperience
 - SweetWiki
 - SweetWikiDesignGoals
 - SweetWikiDiscussion
 - SweetWikiIntroduction
 - UsabilityKupu
 - UsabilitySW
 - WhyAndHowToTag
 - WikiNewsTICFA
 - AppleHumanInterface
- Page tags:** crifa is a subclass of :
 - in the same category :
 - Pages tagged with crifa :
 - Images tagged with crifa (click on the images to go to the corresponding page)
 - Pages that contains videos tagged with crifa (click on the link to go to the corresponding page)
 - Files tagged with crifa

Exemple SweetWiki

Welcome Eva | Home Page | Disconnect

Search

- ◆ **Keyword search**
- ◆ **Tag search**
- ◆ Recent changes
- ◆ Advanced search
- ◆ All users

Webs

- ◆ Main
- ◆ Documentation
- ◆ Prive
- ◆ edit

Pages

- ◆ Sand Box

Administration

- ◆ Manage Ontologies
- ◆ Reload Corese

elogbook

elogbook is a subclass of :

[services](#), [palette](#), [crifa](#), [ste](#), [ulg](#), [universite](#), [liege](#),



in the same category :

[sweetwiki](#),

Pages tagged with elogbook :

- ◆ [ELogBook](#)
- ◆ [PaletteToolsAndServices](#)


Images tagged with elogbook (click on the images to go to the corresponding page)

- ◆ 
- ◆ 

Pages that contains videos tagged with elogbook (click on the link to go to the corresponding page)

Files tagged with elogbook

Exemple SweetWiki



Welcome Eva | Home Page

Search

- ◆ **Keyword search**
- ◆ **Tag search**

Tags' informations

- ◆ **Category** : [ergonomie](#), [expertise](#),
- ◆ **Related tags** : [scapin](#), [meda](#), [bastien](#),

Exemple del.icio.us

The screenshot shows the del.icio.us interface for user 'evandepu'. At the top, there's a navigation bar with links for 'popular', 'recent', 'your bookmarks', 'your network', 'subscriptions', 'links for you', and 'post'. The user is logged in as 'evandepu' with links for 'settings', 'logout', and 'help'. Below the navigation is a search bar and a dropdown menu set to 'del.icio.us'. The main content area displays a list of 14 items, with 'page 1 of 2' shown. Each item includes a title, a link to edit/delete, and a 'saved by' count. A right-hand sidebar contains a 'tags' section with a list of 20 tags and their counts, and a 'tag options' section.

del.icio.us / evandepu /

popular | recent

logged in as **evandepu** | settings | logout | help

All your items (14) del.icio.us

« earlier | later » page 1 of 2

L'enseignement en Communauté française de Belgique [edit / delete](#)
to [accueil enseignement communaute francaise ...](#) [saved by 14 other people](#) ... 8 mins ago

Educnet : nouvelles technologies de l'information et de la communication pour l'éducation - TICE - TIC Enseignement [edit / delete](#)
to [educnet accueil education ...](#) [saved by 182 other people](#) ... 10 mins ago

Main Page - Wikipedia, the free encyclopedia [edit / delete](#)
to [wikipedia accueil ...](#) [saved by 304 other people](#) ... 17 mins ago

Secrétariat de l'Enseignement Catholique en Communautés Française et Germanophone de Belgique (SeGEC) [edit / delete](#)
to [segec accueil ...](#) 2 hours ago

DET - FUNDP [edit / delete](#)
to [det fundp accueil ...](#) 2 hours ago

Grille d'analyse du scénario pédagogique/enseignement stratégique [edit / delete](#)
to [scenarior pedagogique sp ...](#) 3 hours ago

SweetWiki : Prive/MaitriseDesTIC [edit / delete](#)
to [hetice ...](#) 3 hours ago

SweetWiki : Prive/RessourceS [edit / delete](#)
to [crifa competences tic ...](#) 3 hours ago

Université de Liège - - - Accueil [edit / delete](#)
to [accueil ulg ...](#) 3 hours ago

FUNDP - FUNDP [edit / delete](#)
to [fundp accueil ...](#) [saved by 8 other people](#) ... 3 hours ago

▼ **tags**

- 9 [accueil](#)
- 1 [animation](#)
- 1 [carnets](#)
- 1 [cefis](#)
- 1 [communaute](#)
- 1 [competences](#)
- 2 [crifa](#)
- 1 [det](#)
- 1 [education](#)
- 1 [educnet](#)
- 1 [enseignement](#)
- 1 [francaise](#)
- 3 [fundp](#)
- 1 [hetice](#)
- 1 [logiciel](#)
- 1 [media](#)
- 1 [multimedia](#)
- 1 [pedagogique](#)
- 1 [perso](#)
- 1 [promethee](#)
- 1 [scenario](#)
- 1 [segec](#)
- 1 [sp](#)
- 1 [sw](#)
- 1 [tic](#)
- 1 [ulg](#)
- 1 [wikipedia](#)

▼ **tag options**

Exemple del.icio.us

 [del.icio.us](#) / tag /
[your bookmarks](#) | [your network](#) | [subscriptions](#) | [links for you](#) | [post](#)

[popular](#) | [recent](#)
logged in as **evandepu** | [settings](#) | [logout](#) | [help](#)

Popular tags on del.icio.us del.icio.us

This is a **tag cloud** - a list of tags where size reflects popularity.
sort: [alphabetically](#) | [by size](#)

design blog art music software web2.0 tools webdesign video programming linux
photography reference education blogs howto shopping css news web travel
inspiration tutorial politics food recipes free books flash imported technology development
opensource games business toread science research search tips internet javascript history google humor sga
fic windows resources health mac photoshop funny ruby python wiki social media java movies tutorials online rails
culture security php writing fashion teaching download portfolio mobile community home slash fun audio apple article computer
diy graphics illustration math freeware architecture iphone marketing tv photo au ubuntu photos images cooking seo bandslash
wordpress osx jobs tech library environment hardware youtube ajax mckay/sheppard typography radio email film english lifehacks collaboration
wishlist finance learning recipe videos clothing 2008 mp3 flickr reading blogging comics firefox productivity language school entertainment podcast
magazine ideas shop animation green electronics literature furniture

(red tags are tags you share with everyone else)

Métadonnées

(processus de caractérisation)

- Définition de facettes
 - facettes et valeurs
 - définies classiquement (structure d'une BD)
 - issues d'une ontologie

Exemple BayFac

FormaHetic - Document recherche

Recherche Langue Connexion

Mots clés

Matière

Auteur

Type de contenu

Date
From
To

Type de support

Rechercher

266 résultats sur 266 dans la base de connaissances

- ➔ [1 annexe publieWeb EV cefis](#)
- ➔ [1 annexe util netmeeting](#)
- ➔ [1 Article ress ppt](#)
- ➔ [1 ecolEte intro](#)
- ➔ [1 Invitation plen 23mai2006 vf](#)
- ➔ [1 Questionnaire groupes th matiques](#)
- ➔ [1compte rendu 18 12 01](#)
- ➔ [1compterendu 100501](#)
- ➔ [1reu101001](#)
- ➔ [2 annexe netmeeting](#)

1 - 2 - 3 - 4 - 5 - 6 - 7 - 8 - 9 - 10 - 11 - 12 - 13 - 14 - 15 - 16 - 17 - 18 - 19 - 20 - 21 - 22 - 23 - 24 - 25 - 26 - 27 [Suivant »](#)