

CNAM - Département Informatique

RCP216 Ingénierie de la fouille et de la visualisation de données massives  
M. Crucianu, R. Fournier, P. Cubaud

# **Visualisation d'information (5)**

## **Masses de {textes, images, livres}**

Pierre Cubaud <cubaud@cnam.fr>

juin 2015

**le cnam**

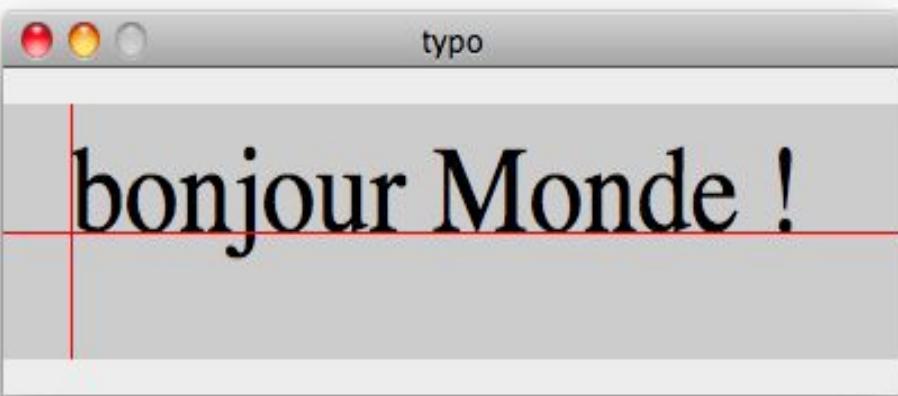
## **Plan de l'exposé**

1. Masses de textes : bibliométrie, scientométrie
2. Masses d'images : albums photos
3. Masses de livres : bibliothèques numériques

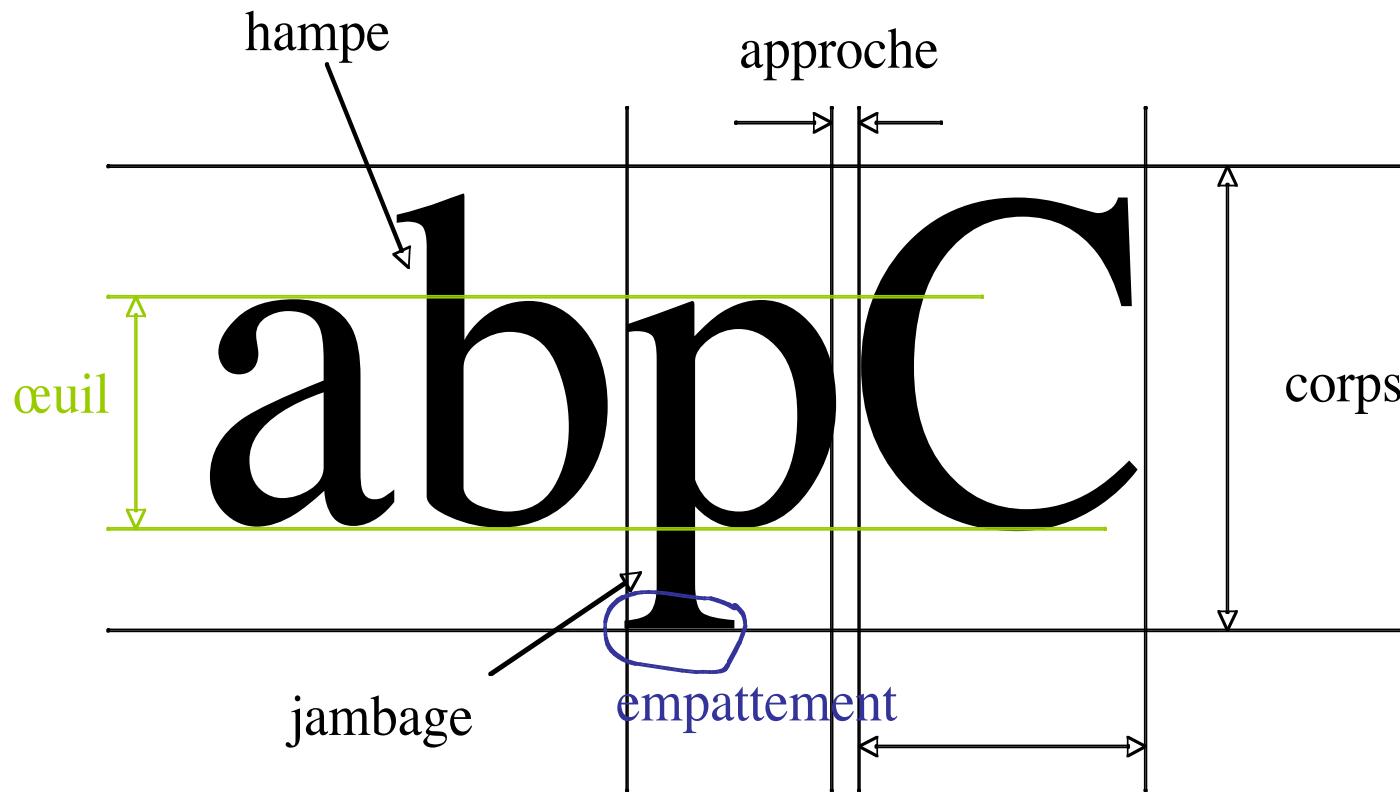
# 1. Masses de textes

## Afficher du texte (avec Processing)

```
typo
size(400,100);
PFont f = loadFont("Serif-48.vlw");
textFont(f);
fill(0);
text("bonjour Monde !", 30,50);
stroke(255,0,0);
line(30,0, 30,100);
line(0,50, 400,50);
```



Attention à  
l'origine du tracé  
 $\neq$  rect, ellipse, etc.



Unité de mesure : le point typographique

Point Didot = 0,3579mm

Point pica = 1/72 pouce = 0,35135mm

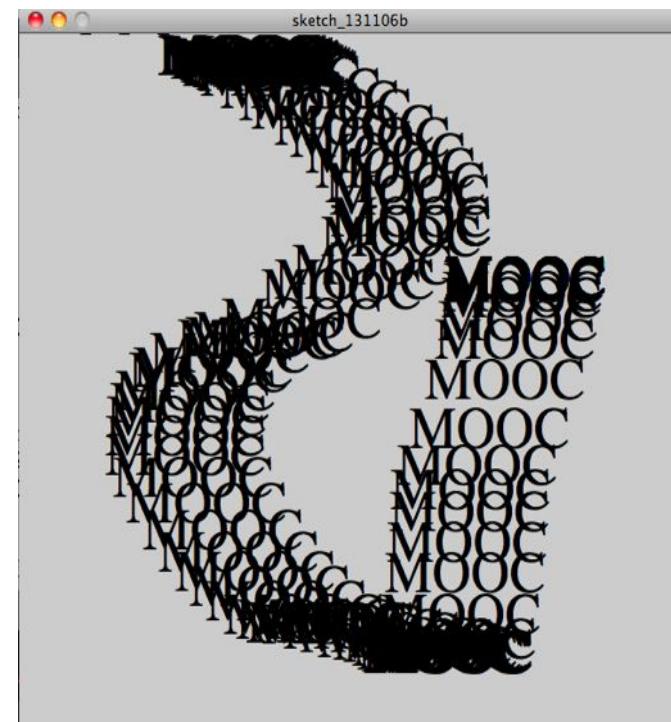
Point métrique = 0,4 mm

```
typomouse
```

```
void setup(){
  size(600,600);
  PFont f = loadFont("Serif-48.vlw");
  fill(0);
  textFont(f);
}
void draw(){
  text("MOOC", mouseX, mouseY);
}
```



étudier quelques variantes pour le placement du texte  
(aléa, etc)



## Tracé de taille variable

typozoom

```
void setup(){
    size(800,500);
    PFont f = loadFont("Serif-48.vlw");
    textAlign(f);
}

void draw(){
    background(200);
    float s=map(mouseX,0,width,0,500);
    fill(0);
    textSize(s);
    text("MOOC", 50, 400);
    fill(255);
    textSize(500-s);
    text("MOOC", 50, 400);
}
```



démonstration

# Réduction extrême

```
textered
```

```
String lines[] = loadStrings("methodeWRAP40.txt");
PGraphics pmp = createGraphics(600,lines.length*20,JAVA2D);

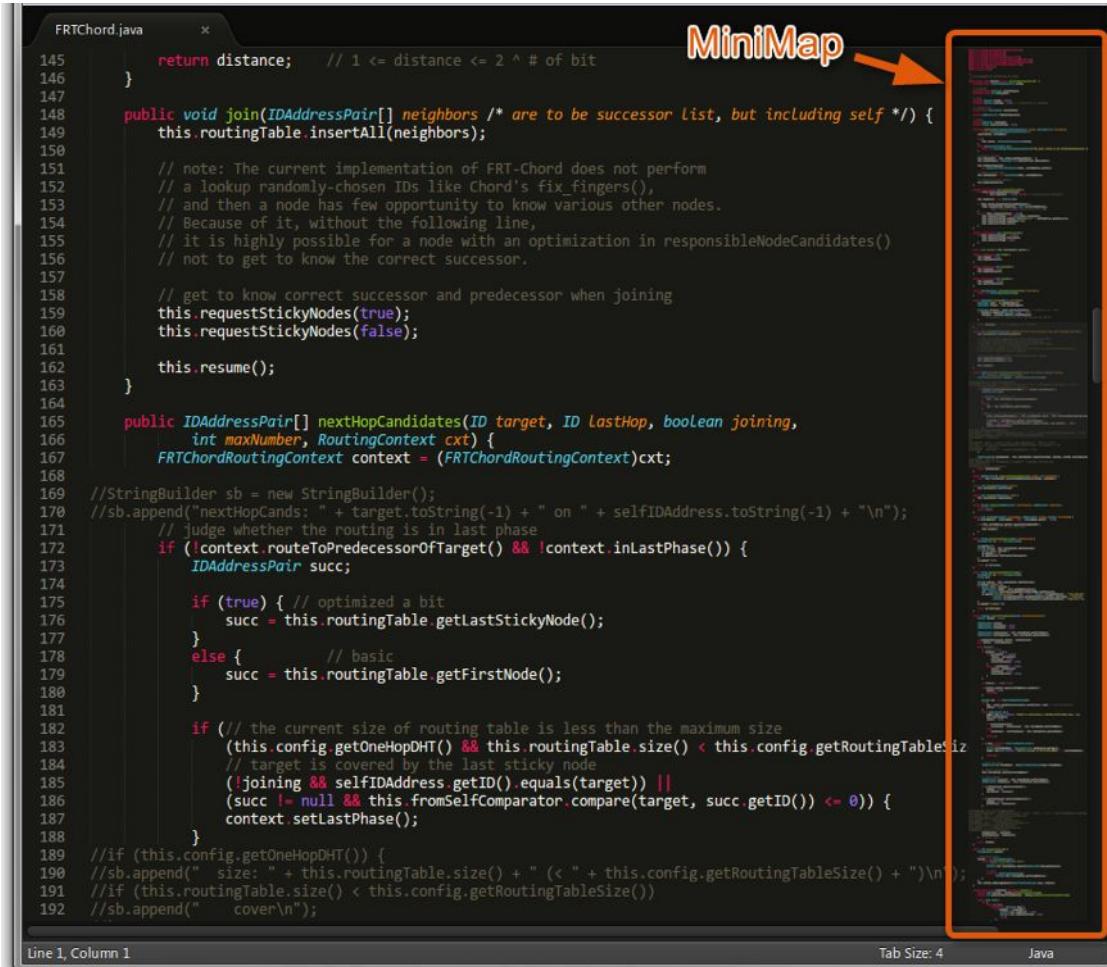
pmp.beginDraw();
pmp.textSize(20);
pmp.fill(10);
pmp.background(250);

int y=0;
for (int i=0;i<lines.length;i++){
  pmp.text(lines[i], 10,y);
  y += 20;
}

pmp.endDraw();
pmp.save("resu.png");
println("done");
```



# Exemple : éditeur Sublime Text

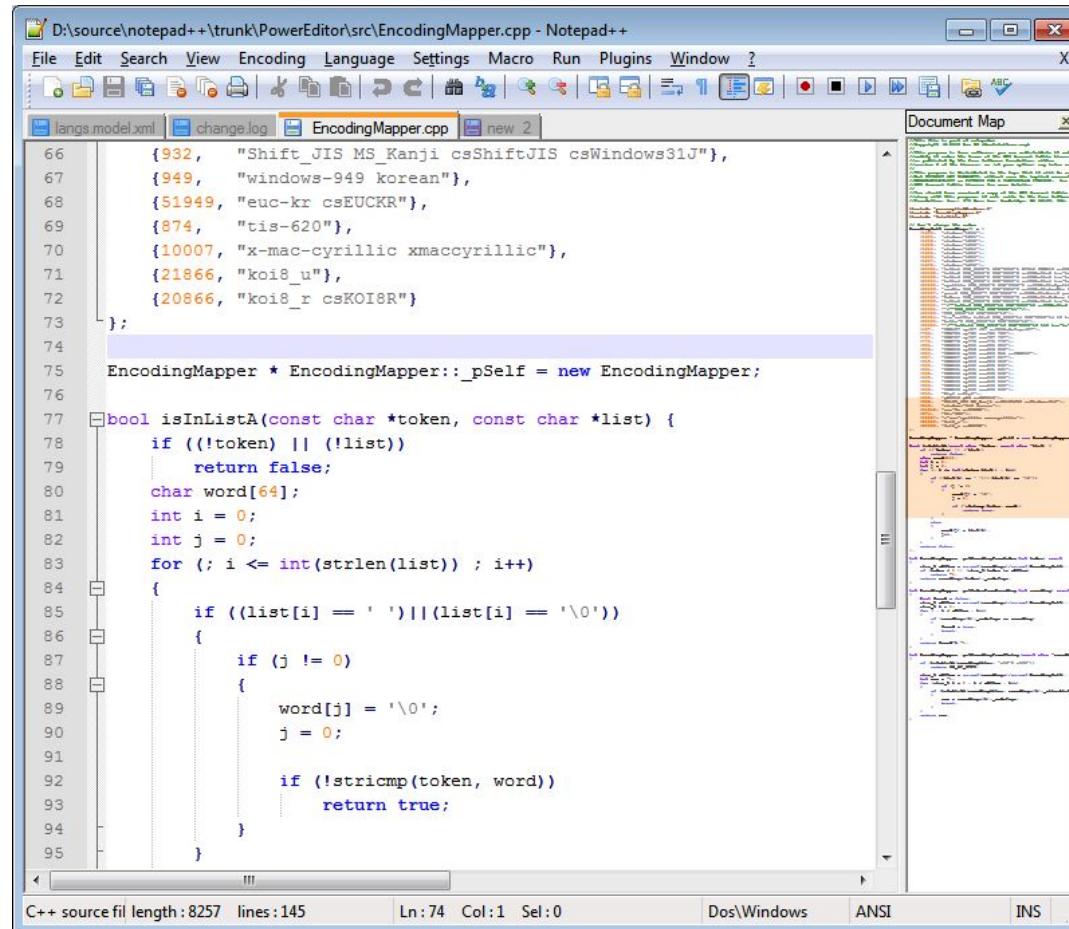


The screenshot shows a Sublime Text window with a Java file named `FRTChord.java` open. The code is a class implementation with various methods and comments. An orange arrow points from the text "MiniMap" to a vertical sidebar on the right side of the editor, which displays a compressed preview of the entire codebase. The status bar at the bottom indicates "Line 1, Column 1", "Tab Size: 4", and "Java".

```
145     return distance; // 1 <= distance <= 2 ^ # of bit
146 }
147
148 public void join(IDAddressPair[] neighbors /* are to be successor list, but including self */) {
149     this.routingTable.insertAll(neighbors);
150
151     // note: The current implementation of FRT-Chord does not perform
152     // a lookup randomly-chosen IDs like Chord's fix_fingers(),
153     // and then a node has few opportunity to know various other nodes.
154     // Because of it, without the following line,
155     // it is highly possible for a node with an optimization in responsibleNodeCandidates()
156     // not to get to know the correct successor.
157
158     // get to know correct successor and predecessor when joining
159     this.requestStickyNodes(true);
160     this.requestStickyNodes(false);
161
162     this.resume();
163 }
164
165 public IDAddressPair[] nextHopCandidates(ID target, ID lastHop, boolean joining,
166                                         int maxNumber, RoutingContext ctxt) {
167     FRTChordRoutingContext context = (FRTChordRoutingContext)ctxt;
168
169     //StringBuilder sb = new StringBuilder();
170     //sb.append("nextHopCands: " + target.toString(-1) + " on " + selfIDAddress.toString(-1) + "\n");
171     // judge whether the routing is in last phase
172     if (!context.routeToPredecessorOfTarget() && !context.inLastPhase()) {
173         IDAddressPair succ;
174
175         if (true) { // optimized a bit
176             succ = this.routingTable.getLastStickyNode();
177         } else { // basic
178             succ = this.routingTable.getFirstNode();
179         }
180
181         if (// the current size of routing table is less than the maximum size
182             (this.config.getOneHopDHT() && this.routingTable.size() < this.config.getRoutingTableSize()
183             // target is covered by the last sticky node
184             (&& target.equals(succ)) ||
185             (joining && selfIDAddress.getID().equals(target)) ||
186             (succ != null && this.fromSelfComparator.compare(target, succ.getID()) <= 0)) {
187             context.setLastPhase();
188         }
189     //if (this.config.getOneHopDHT()) {
190     //sb.append(" size: " + this.routingTable.size() + " (" + this.config.getRoutingTableSize() + ")\n");
191     //if (this.routingTable.size() < this.config.getRoutingTableSize())
192     //sb.append("    cover\n");
193 }
```

<http://did2memo.net/wp-content/uploads/2013/01/minimap-in-sublime-text-2.png>

## idem avec notepad++



The screenshot shows the Notepad++ interface with the following details:

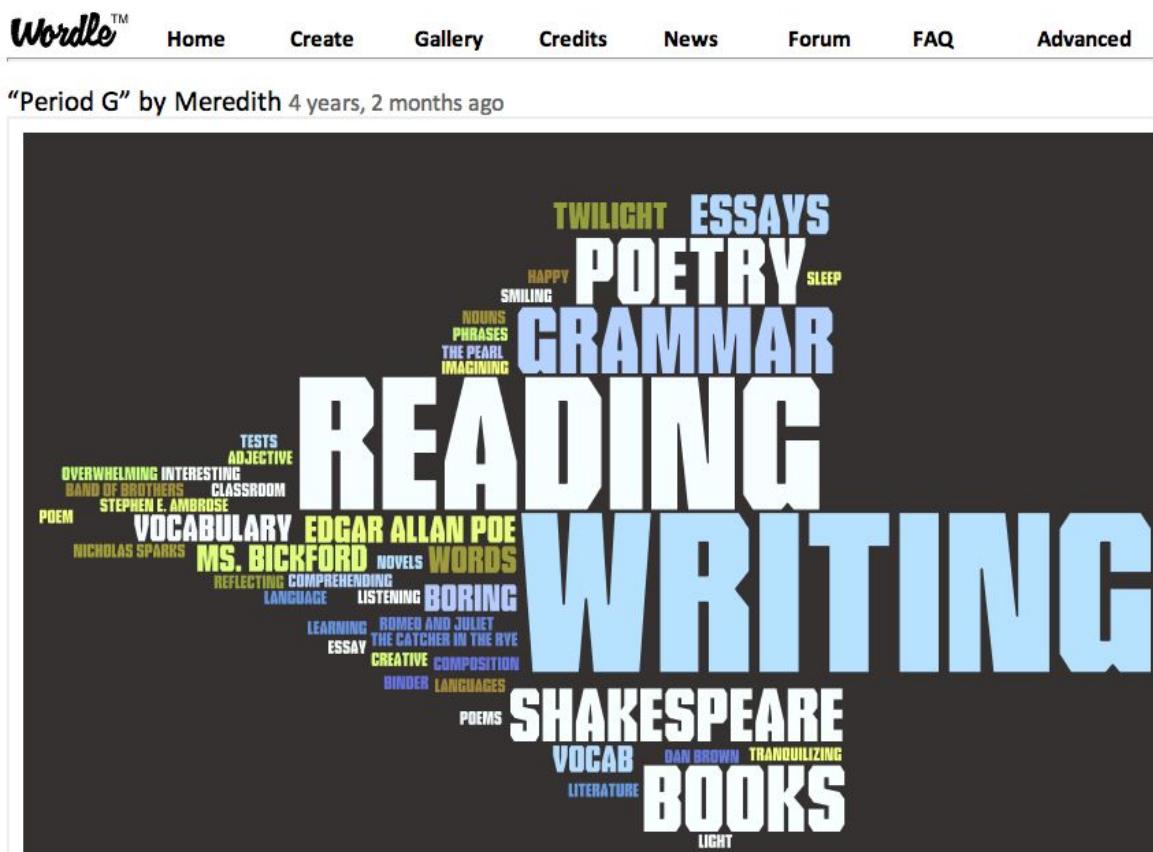
- Title Bar:** D:\source\notepad++\trunk\PowerEditor\src\EncodingMapper.cpp - Notepad++
- Menu Bar:** File, Edit, Search, View, Encoding, Language, Settings, Macro, Run, Plugins, Window, ?
- Toolbar:** Standard file operations like Open, Save, Print, Find, Replace, etc.
- Document Map:** A sidebar panel titled "Document Map" showing a hierarchical tree of the code structure.
- Code Editor:** The main window displays C++ code with syntax highlighting. The code includes declarations for character sets and a function for searching a list.

```
66     {932, "Shift_JIS MS_Kanji csShiftJIS csWindows31J"},  
67     {949, "windows-949 korean"},  
68     {51949, "euc-kr csEUCKR"},  
69     {874, "tis-620"},  
70     {10007, "x-mac-cyrillic xmaccyrillic"},  
71     {21866, "koi8_u"},  
72     {20866, "koi8_r csKOI8R"}  
73 };  
74  
75 EncodingMapper * EncodingMapper::_pSelf = new EncodingMapper;  
76  
77 bool isInListA(const char *token, const char *list) {  
78     if (!token) || (!list))  
79         return false;  
80     char word[64];  
81     int i = 0;  
82     int j = 0;  
83     for (; i <= int(strlen(list)) ; i++)  
84     {  
85         if ((list[i] == ' ')||(list[i] == '\0'))  
86         {  
87             if (j != 0)  
88             {  
89                 word[j] = '\0';  
90                 j = 0;  
91             }  
92             if (!strcmp(token, word))  
93                 return true;  
94         }  
95     }  
96 }
```

- Status Bar:** C++ source fil length : 8257 lines : 145 Ln : 74 Col : 1 Sel : 0 Dos\Windows ANSI INS

<https://notepad-plus-plus.org/assets/images/docMap.png>

# Les nuages de mots : à voir au TP#5



# Et encore

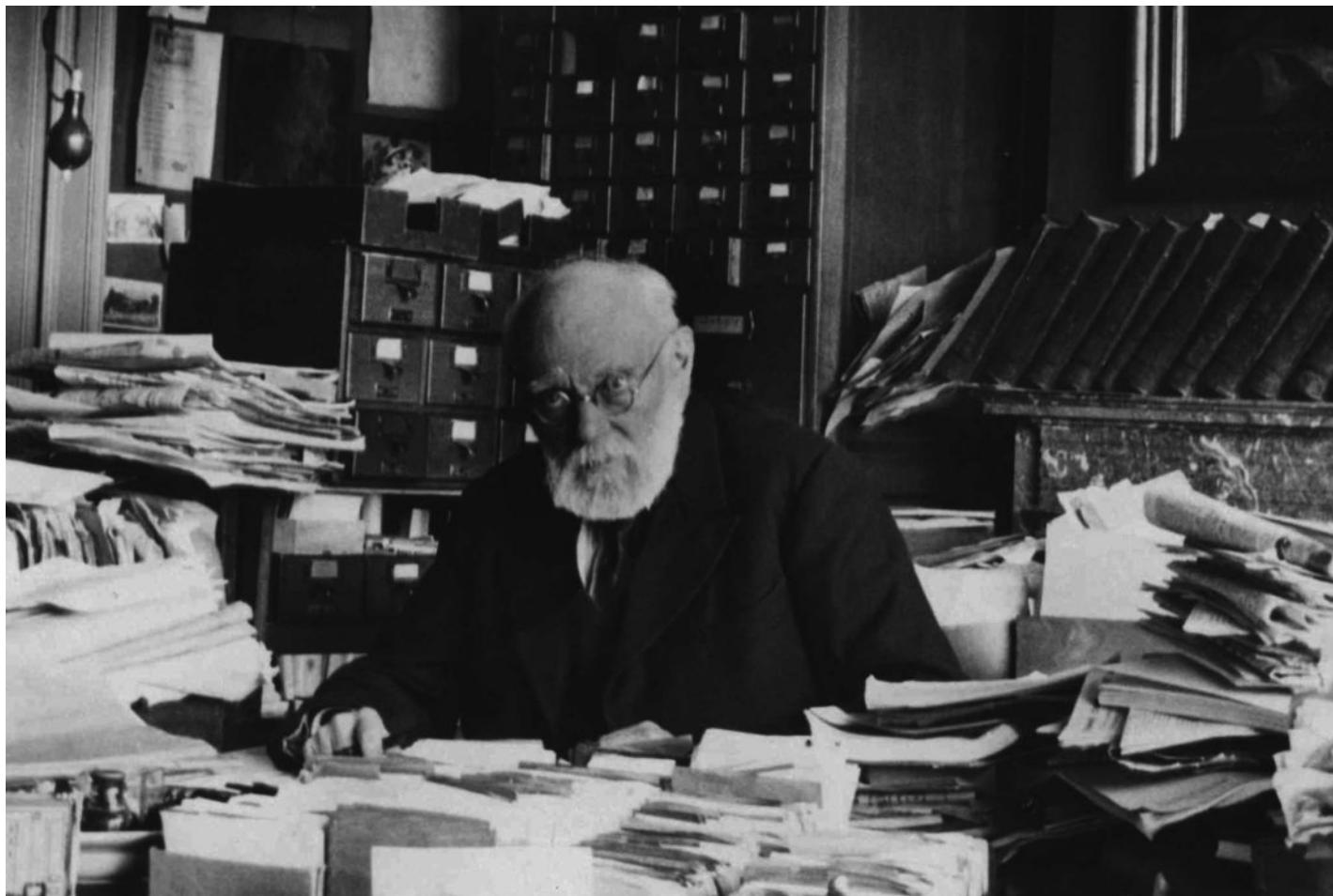


<http://www.davidsmall.com/talmud.html>

<http://acg.media.mit.edu/projects/thesis/DSThesis.pdf>

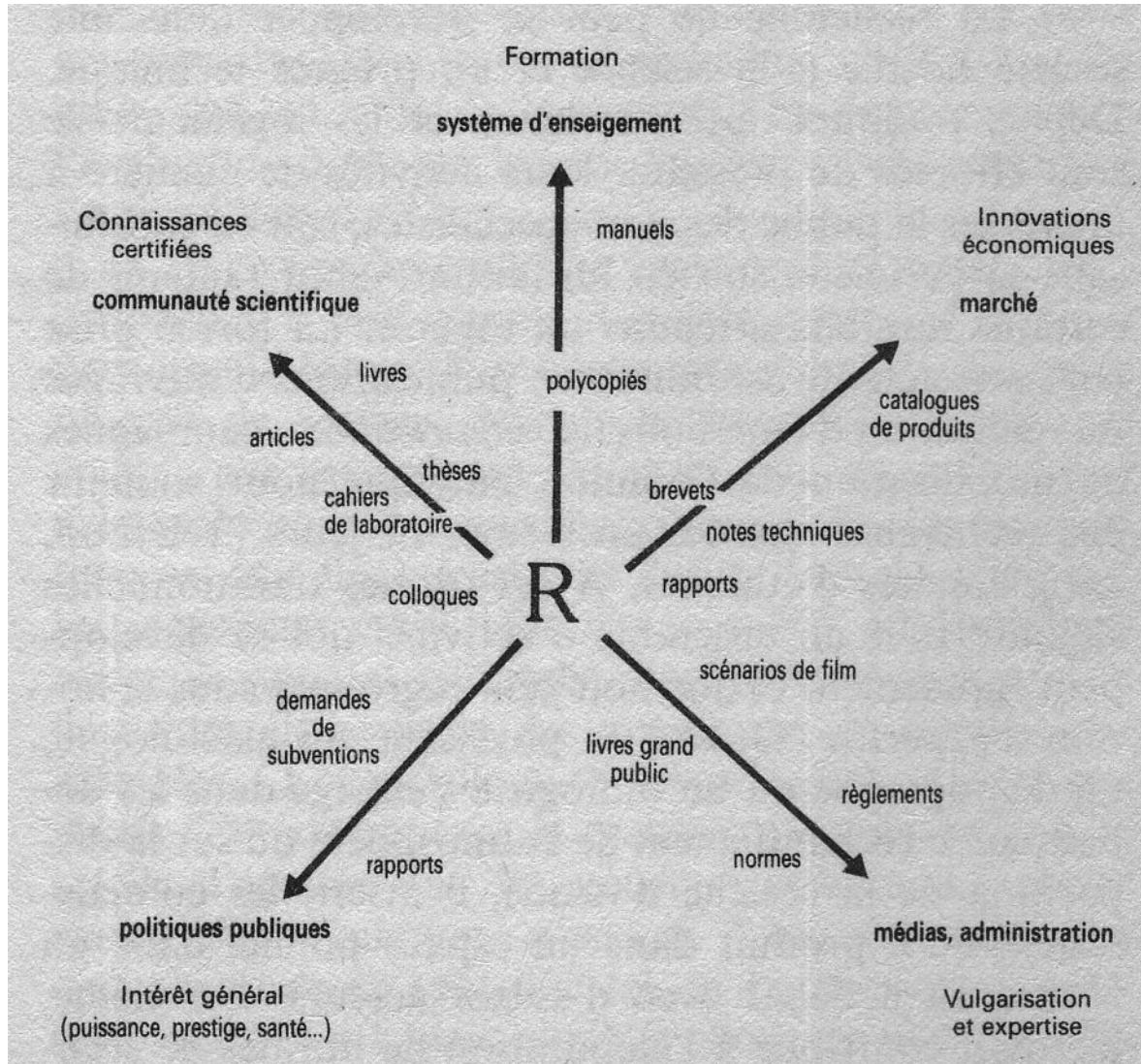
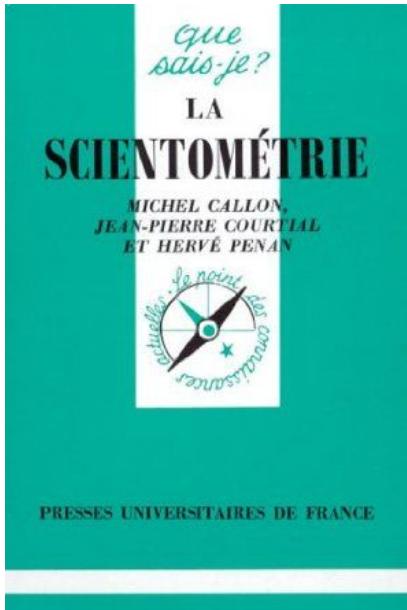
Article : <http://www.research.ibm.com/journal/sj/353/sectiond/small.pdf>

# bibliométrie / scientométrie



Paul Otlet : un internet de papier ?

# productions de la recherche



## les "connaissances certifiées"

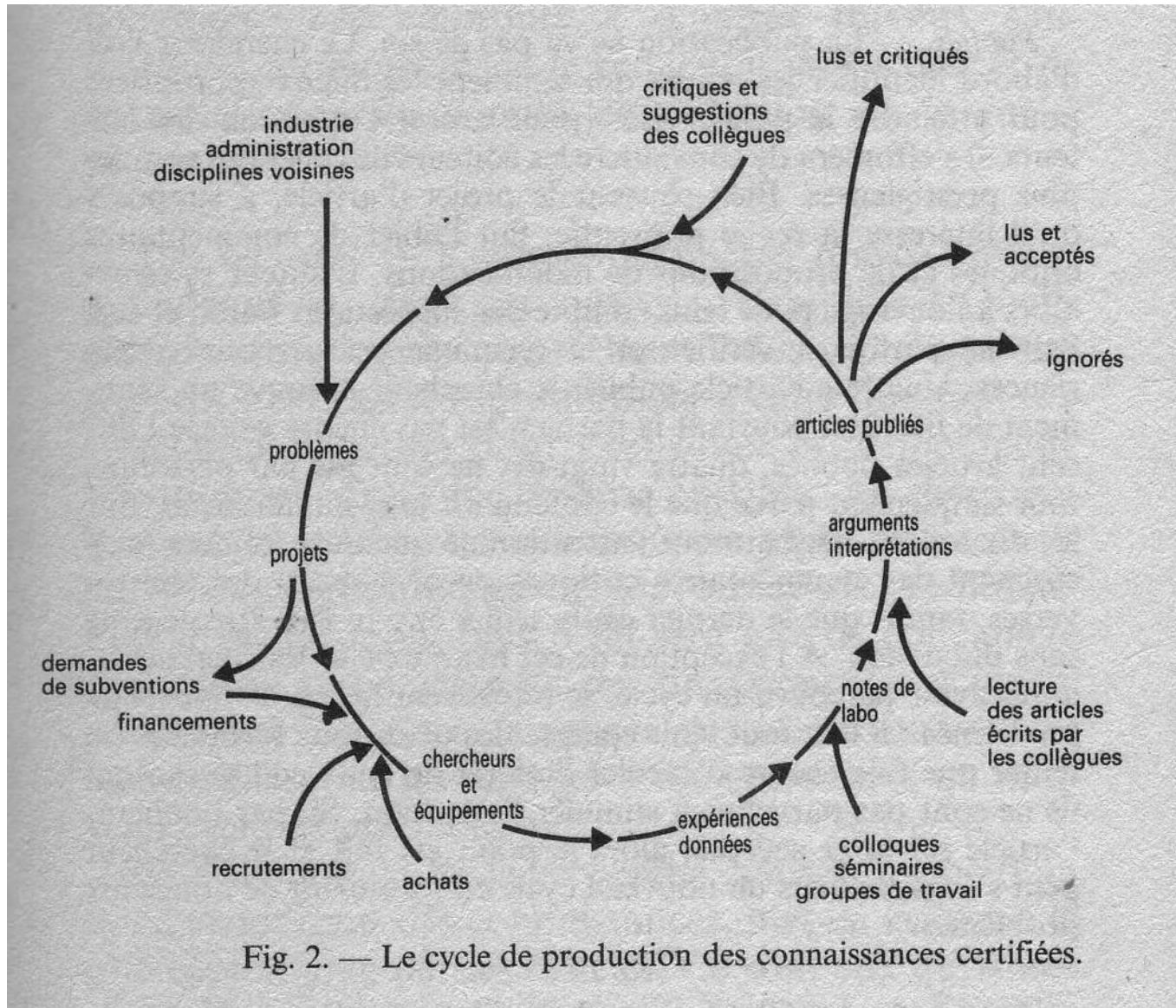
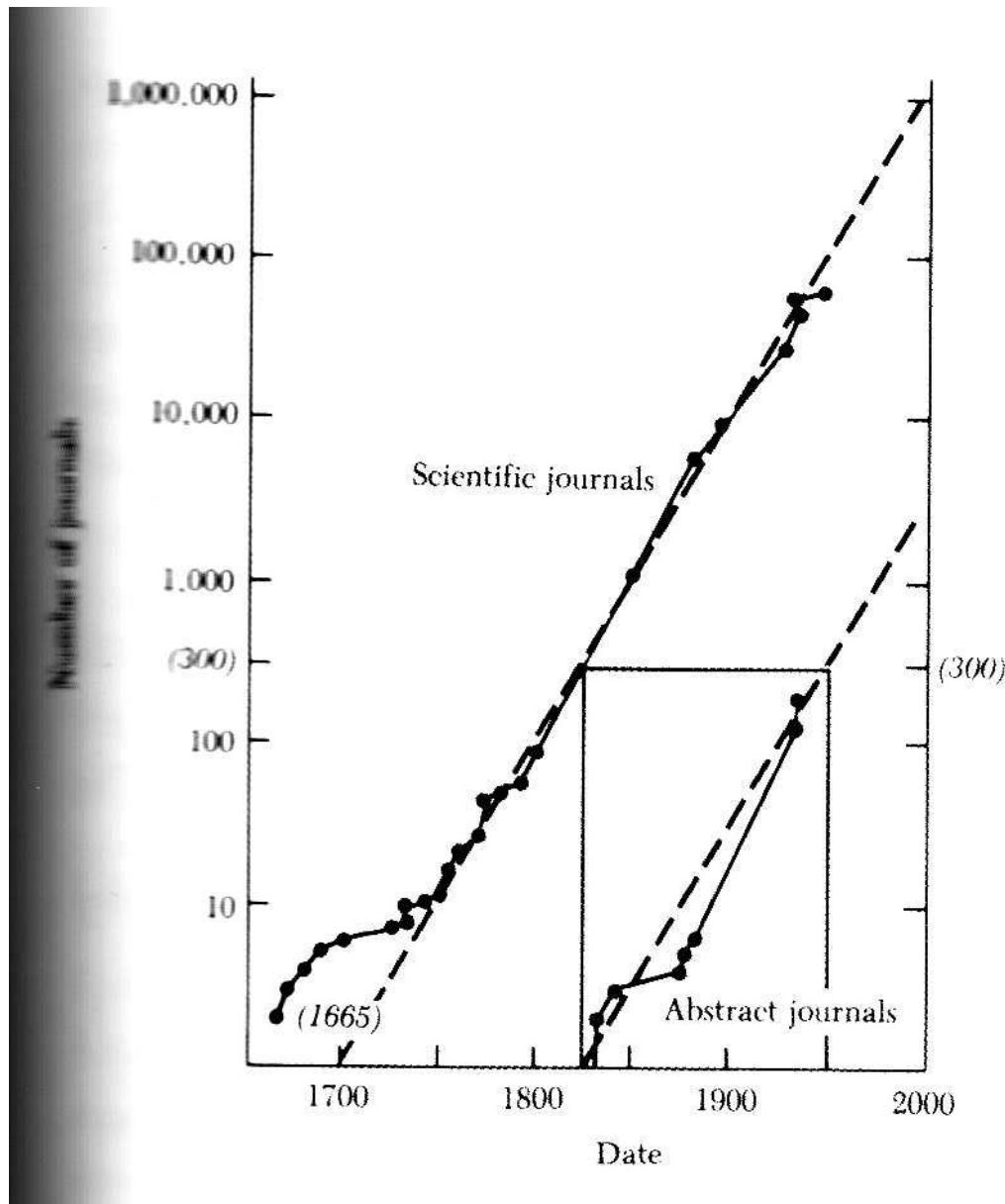


Fig. 2. — Le cycle de production des connaissances certifiées.

## **croissance des publications et pb de l'evaluation**



Google publish or perish

Web Images Maps Shopping Plus Outils de recherche

Environ 1 410 000 résultats (0,40 secondes)

Les cookies assurent le bon fonctionnement de nos services. En utilisant ces derniers, vous acceptez l'utilisation des cookies.

OK En savoir plus

### Articles universitaires correspondant aux termes **publish or perish**

- [\*\*Publish or perish\*\* - Parchomovsky](#) - Cité 125 fois
- [\*\*Publish or perish\*\* - Clapham](#) - Cité 30 fois
- [\*\*Publish or perish\*\* - Hackathorn](#) - Cité 12 fois

une spirale infernale ??? => automatisation

# bases bibliographiques

le grand ancêtre : MEDLINE (1964)

```
^F File ^E Edit ^A Search ^L Limit ^V View ^T Tools ^O Options ^Y Help
1 primordial dwarfism.mp. [mp=ti, sh, ab, it, hw, tn, mf, rw] 90
2 remove duplicates from 1 51
3 (children and nutrition).mp. [mp=ti, sh, ab, it, hw, tn, mf, 7819
4 (children and nutrition).ft. 154
5 remove duplicates from 4 152

Ovid - CINAHL & EMBASE & MEDLINE

[To select option hold Ctrl and letter indicated. Press ^Y for Help.] Enter subject, then press <Enter>
:

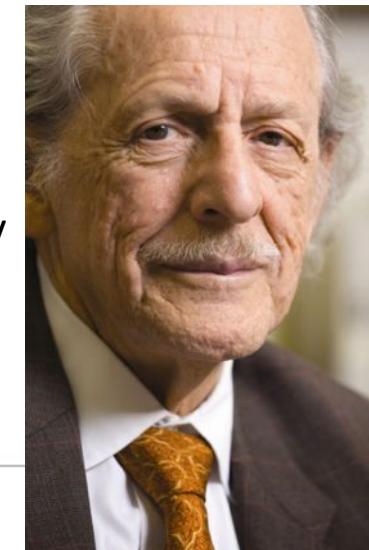
^U Author ^J Journal ^G Limit Set ^M Combine ^B Save
^R Textword ^D Database ^K View Set ^P Print Set ^X Exit
```

1,6 G requêtes en 2011 sur pubMED

The screenshot shows the PubMed.gov homepage. At the top, there is a navigation bar with links for "PubMed", "Advanced", "Search", and "Help". Below the navigation bar, a blue banner displays a message: "PubMed is open, however it is being maintained with minimal staffing due to the lapse in government funding. Information will be updated to the extent possible, and the agency will attempt to respond to urgent operational inquiries. For updates regarding government operating status see USA.gov." The main content area of the page is currently empty.

# index de citations

<http://www.garfield.library.upenn.edu/>



## Eugene Garfield

From Wikipedia, the free encyclopedia

*This article is about the scientist. For the lawyer and railroad executive, see [Eugene K. Garfield](#).*

**Eugene Garfield** (born September 16, 1925) is an American scientist, one of the founders of bibliometrics and scientometrics.<sup>[1]</sup> He received a PhD in Structural Linguistics from the University of Pennsylvania in 1961. Dr. Garfield was the founder of the Institute for Scientific Information (ISI), which was located in Philadelphia, Pennsylvania. ISI now forms a major part of the science division of Thomson Reuters company. Garfield is responsible for many innovative bibliographic products, including *Current Contents*, the *Science Citation Index* (SCI), and other citation databases, the *Journal Citation Reports*, and *Index Chemicus*. He is the founding editor and publisher of *The Scientist*, a news magazine for life scientists. In 2007, he launched *HistCite*, a bibliometric analysis and visualization software package.

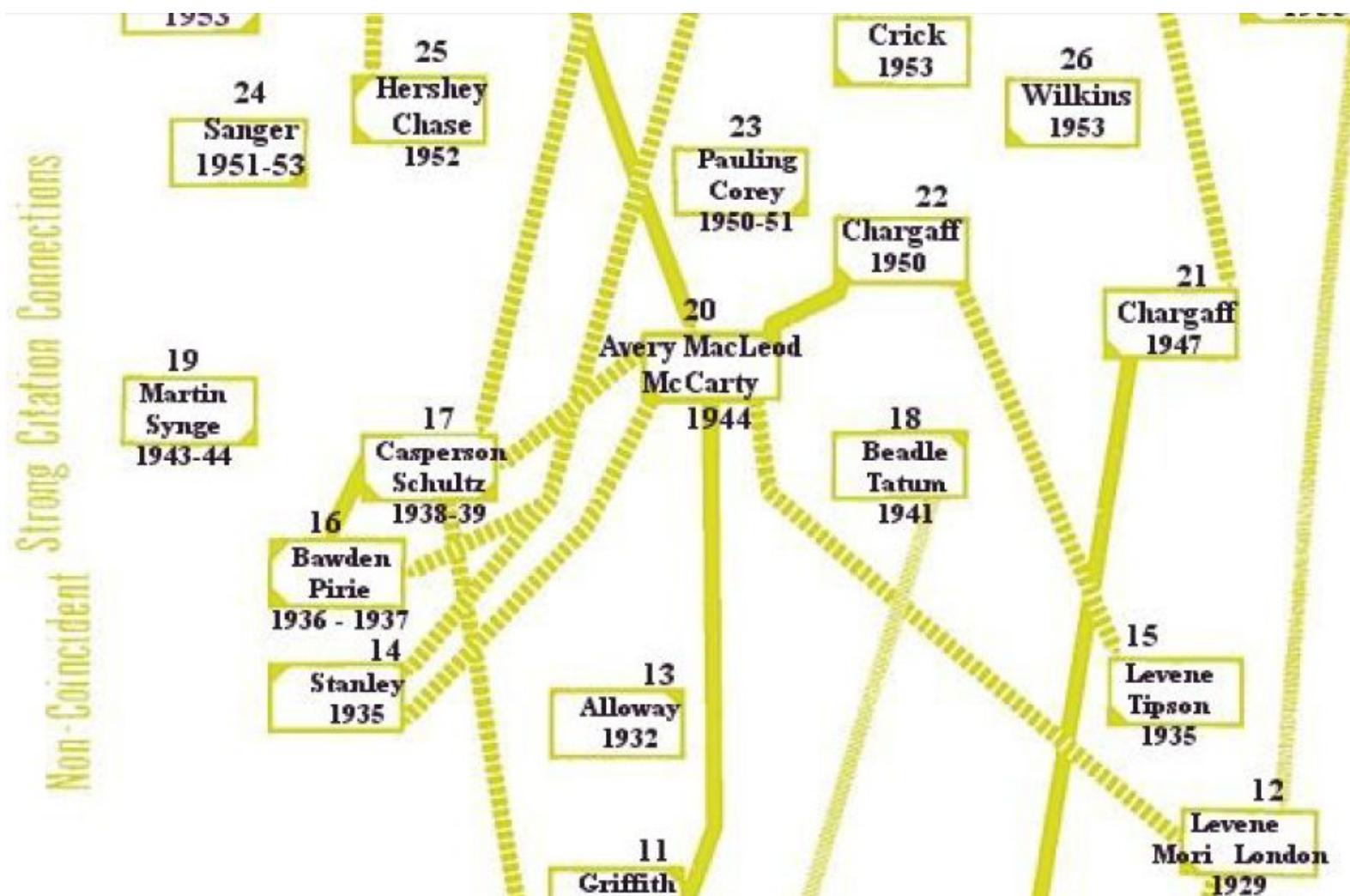
Following ideas inspired by Vannevar Bush's famous 1945 article "As We May Think", Garfield undertook the development of a comprehensive citation index showing the propagation of scientific thinking; he started the Institute for Scientific Information in 1955. The creation of the *Science Citation Index* made it possible to calculate impact factor,<sup>[2]</sup> which measures the importance of scientific journals. It led to the unexpected discovery that a few journals like *Nature* and *Science* were core for all of hard science. The same pattern does not happen with the humanities or the social sciences. [citation needed]

Garfield's work led to the development of several Information Retrieval algorithms, like HITS and Pagerank. Both use the structured citation between websites through hyperlinks.

<b>Eugene Garfield</b>	
<b>Born</b>	September 16, 1925 (age 88) New York City, New York, United States
<b>Education</b>	Ph. D., University of Pennsylvania (1961)
<b>Occupation</b>	Scientist
<b>Known for</b>	One of the founders of bibliometrics and scientometrics <i>Science Citation Index</i> Institute for Scientific Information
<b>Website</b>	
<a href="http://www.garfield.library.upenn.edu/">http://www.garfield.library.upenn.edu/</a>	

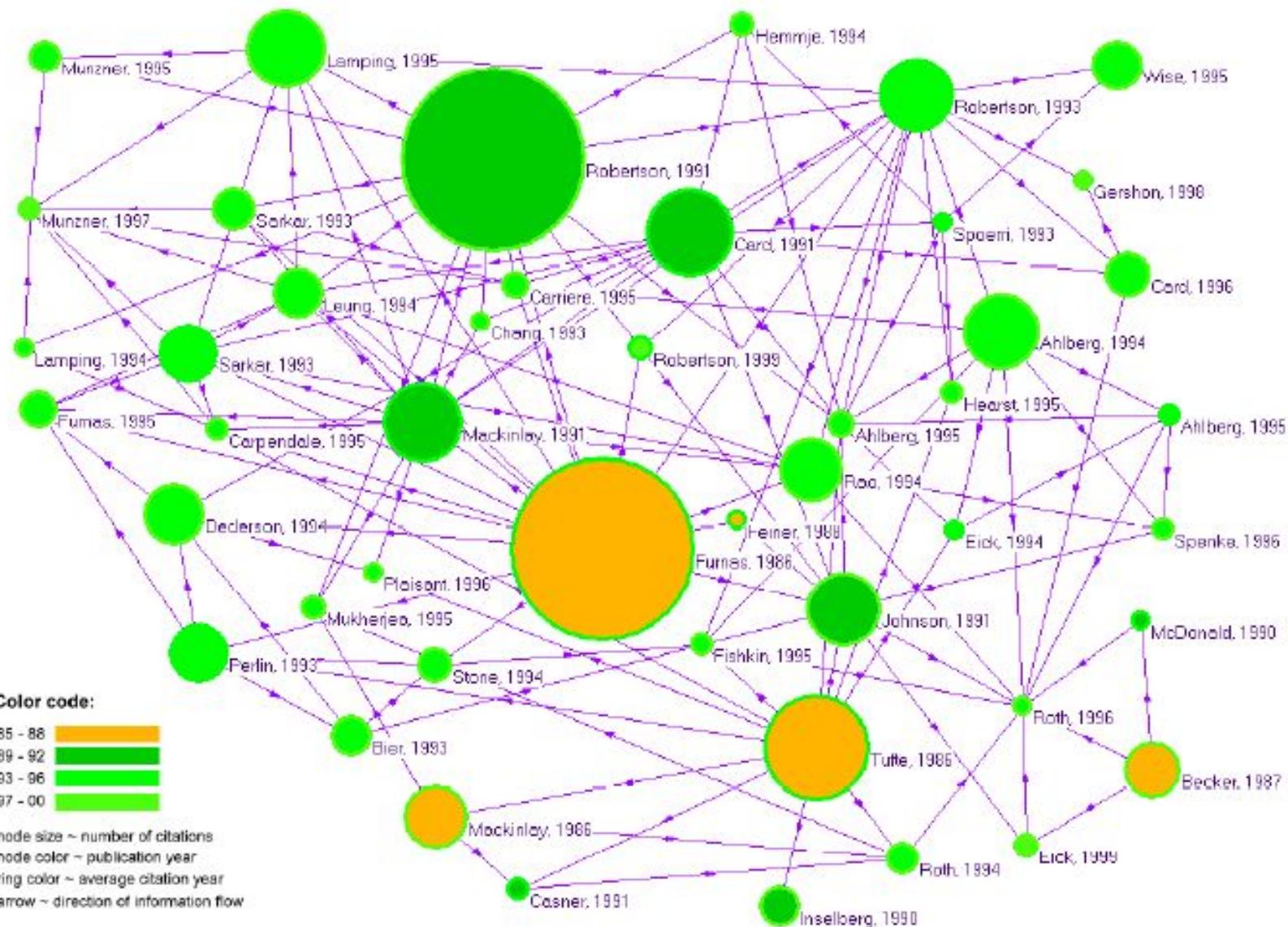
[Contents](#) [hide]

# graphes de citation



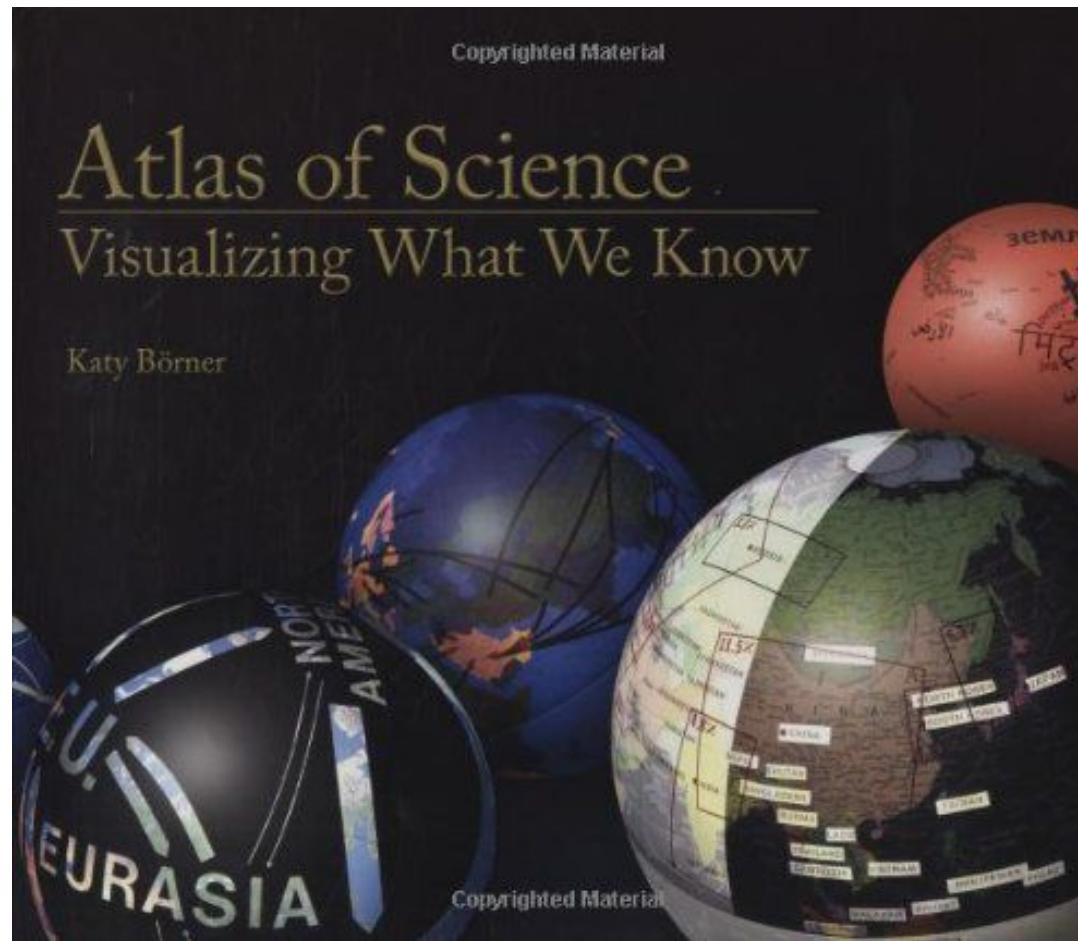
Garfield et al. 1964 (avec Asimov !)

<http://www.garfield.library.upenn.edu/papers/useofcitdatawritinghistofsci.pdf>



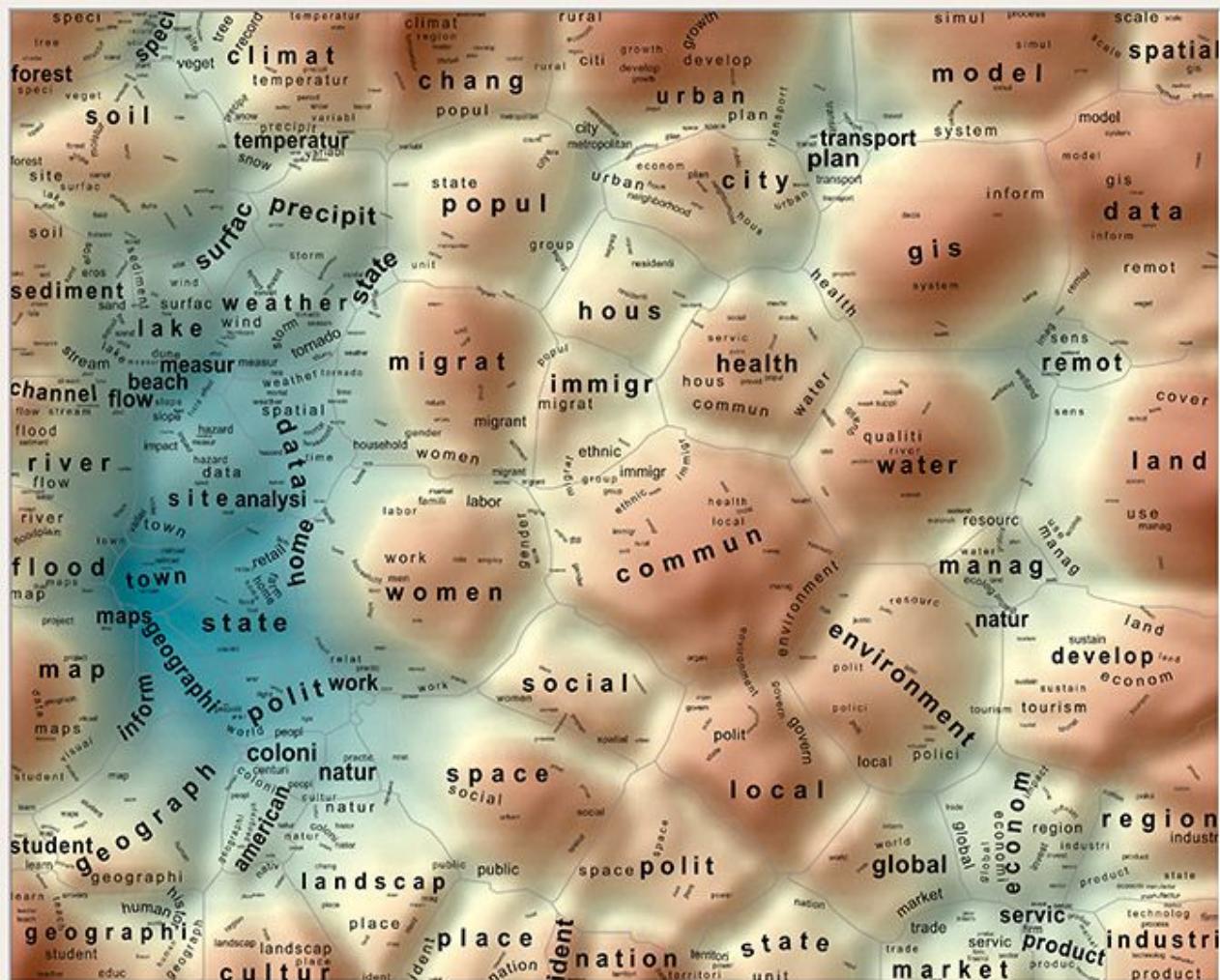
Major Information Visualization Authors, Papers and Topics in the ACM Library  
 Weimao Ke, Katy Borner and Lalitha Viswanat

# cartographie de la connaissance



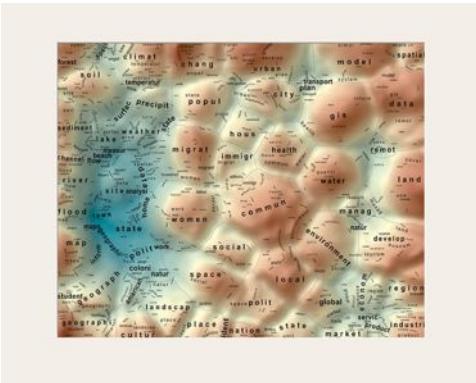
livre magnifique (gd format) et pas trop cher (MIT Press, 30€)

[www.scimaps.org/atlas](http://www.scimaps.org/atlas)



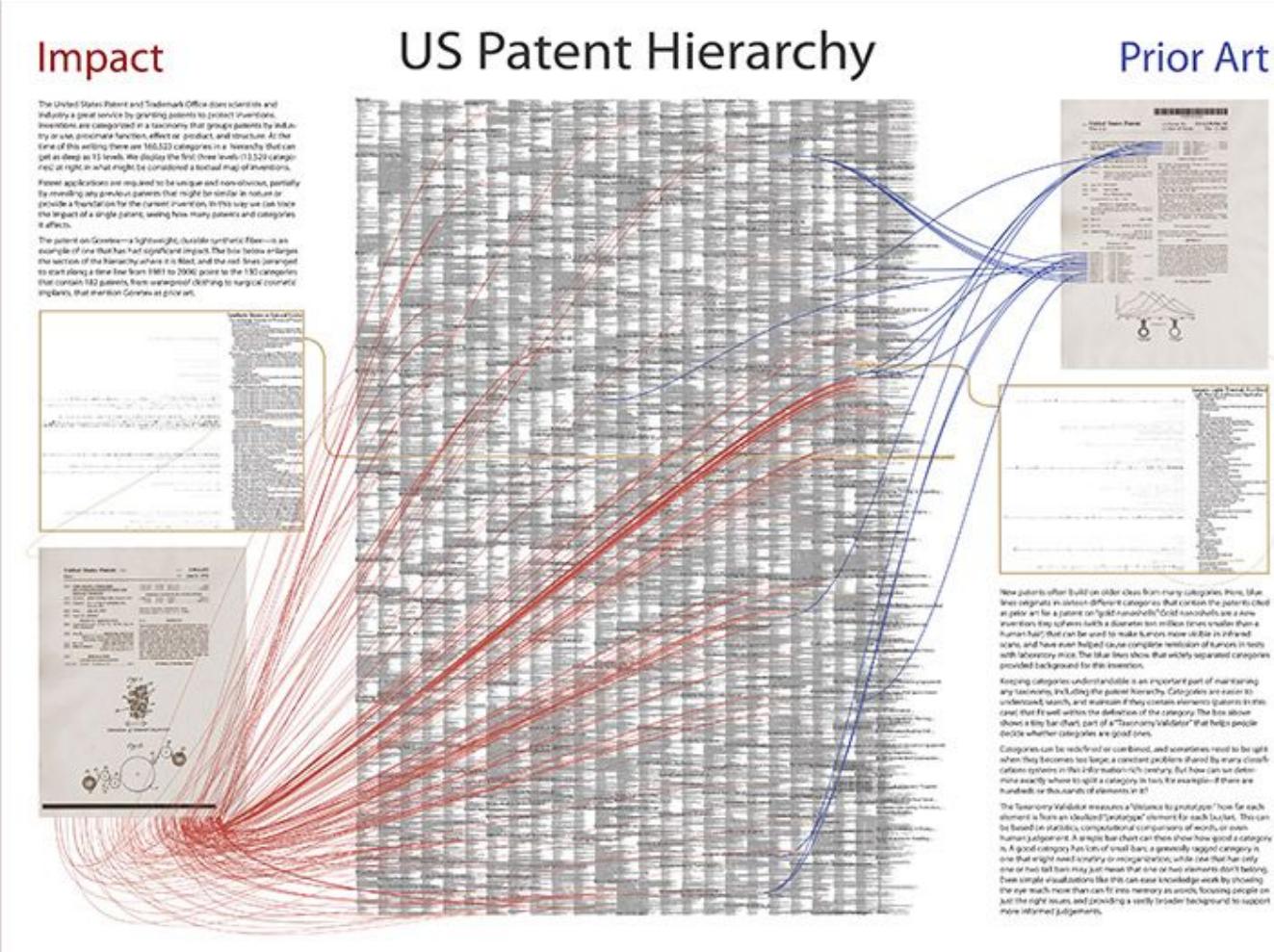
A. Skupin - In terms of geography (2005)

+22000 abstracts des articles  
de la conf. de l'AAG (1993-2002)



carte de Kohonen (SOM) :  
-entrée : vecteur n-dim occurrences  
des mots de l'abstract  
-sortie : vecteur dim 2

altitude liée à la définition des sujets

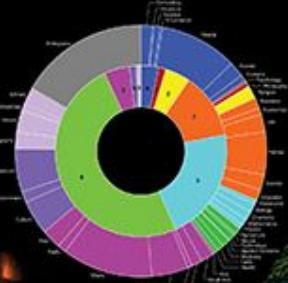


K. Börner et al. Taxonomy visualisation of Patent Data, 2006  
3,17M brevets + 160K catégories

# DESIGN VS. EMERGENCE: VISUALIZATION OF KNOWLEDGE ORDERS

## WIKIPEDIA'S CATEGORY STRUCTURE

The Wikipedia category structure was extracted from the January 2008 dump by the Wikimedia Foundation. Wikipedia has approximately 1.5 million categories, and roughly 10% of all Wikipedia pages are categorized. By contrast, the UDC category structure, as forced by the strict hierarchical nature of the classification, is much more rigid and organized, while the Wikipedia "Category Merge tool" documents "at least 100,000 articles that do not agree with the standard classification scheme." Categories are also used to group pages into the 12 main categories of the MediaWiki namespace, each with the exact article having 40 children nodes. The refutor visualization is based on the first 10 levels of the hierarchy, consisting of 85,708 nodes and 80,714 edges. Node size is based on its total degree, while edge width is based on the total weight it represents along the path between the rightmost node in depth.

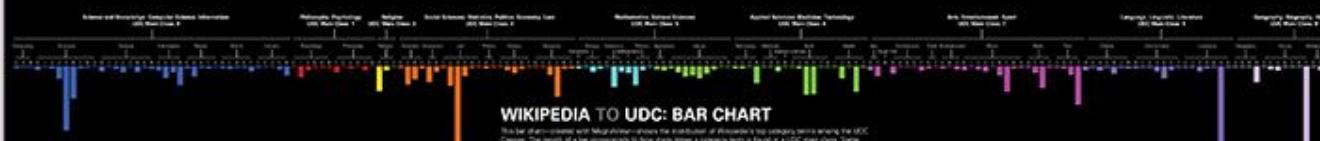
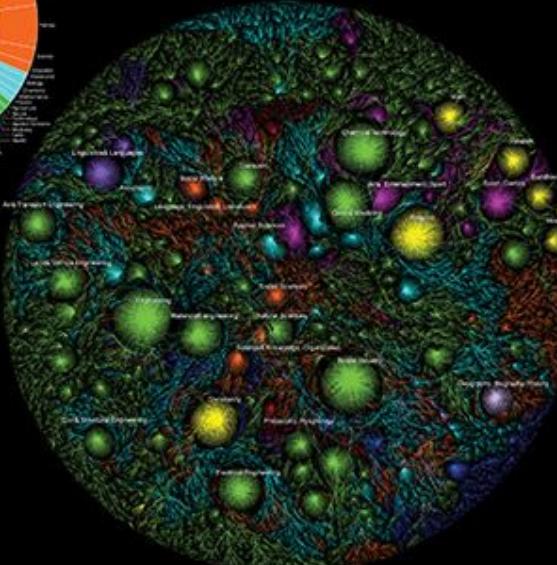


## CATEGORY DISTRIBUTION OF WIKIPEDIA & UDC

This sunburst chart shows the distribution of the top 100 UDC categories and the top 100 Wikipedia categories cluster. Wikipedia categories cluster around the following UDC classes, ordered according to their size: Q27 (Human Sciences) and 510 (Applied Sciences), followed by Q16 (Natural Sciences) and 510 (Applied Sciences). Wikipedia's 100 largest categories, however, are distributed across many categories, including Q16 (Natural Sciences) and 510 (Applied Sciences), Q27 (Human Sciences), and 510 (Applied Sciences). A large part (17%) of Wikipedia categories are located in the category "Unknown".

## UNIVERSAL DECIMAL CLASSIFICATION

The UDC was created in 1908 by Paul Otlet and Henri La Fontaine as an attempt at unification of existing knowledge. The UDC is a hierarchical classification system that attempts to categorize all the world's knowledge. One data extract from a 2004 edition of the UDC has a total of 101 levels, organized in a tree with branching not to the depth of nine, organized under nine main classes. This visualization is based on the first 10 levels of the UDC, which contains a total of 96,404 nodes and 113,033 edges. Node size is based on its total degree, while edge width is based on the total weight it represents along the path between the rightmost node in depth.



## WIKIPEDIA TO UDC: BAR CHART

This bar chart visualizes the relationship between the distribution of Wikipedia to the categories defined among the UDC. This chart is color-coded with 10 major categories. The distribution of Wikipedia to the categories defined among the UDC can be observed. Some categories have strong connection with a single UDC class. For example, "Mathematics" appears only in class 5, whereas "Geography" appears in almost every class. The legend is based on the following categories: 1. Mathematics, 2. Natural Sciences, 3. Applied Sciences, 4. Human Sciences, 5. Technology, 6. Social Sciences, 7. Economics, 8. History, 9. Geography, 10. Unknown. In the 43,000 categories of Wikipedia, the majority of them belong to the first two categories. This visualization is based on the first 10 levels of the UDC. We can see that most of the categories in the first level of the UDC are highly connected to the first two categories. To this end, we have located 100 people's 100 categories within the UDC's exploration table and verified that they match correctly. Each one of the 100 categories of Wikipedia is mapped to over 1000 nodes long, and on the rightmost section,

VISUALIZING TRENDS AND DYNAMICS

# 30 YEARS OF SCIENTIFIC DEVELOPMENT



Social  
sciences

Health &  
medical  
sciences

Biomedical  
sciences

Life & earth  
sciences

Chemistry

Physics &  
engineering

Mathematics  
& computer  
sciences

## A lire sur le sujet

# Visualization of Text Streams: A Survey

Artur Šilić

Department of Electronics, Microelectronics, Computer and Intelligent Systems  
University of Zagreb, Faculty of Electrical Engineering and Computing  
Unska 3, 10000 Zagreb, Croatia

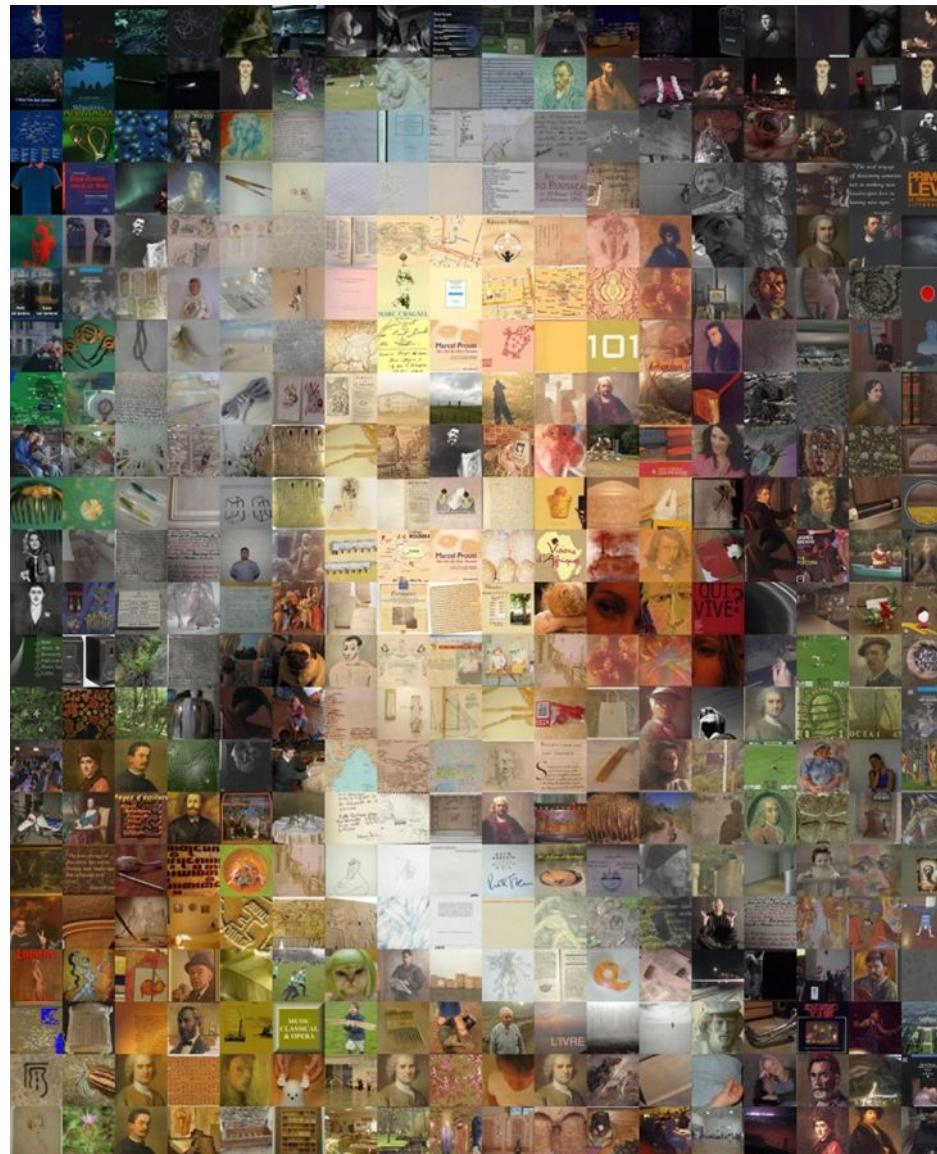
**Abstract**— This work presents related areas of research, types data collections that are visualized, technical aspects of creating visualizations, and evaluation methodologies. Existing methods are structured and explained from the aspect of visualization process. Successful applications are noted and some future trends in the field are anticipated.

**Keywords**— Information Visualization, Visual Analytics, Topic Detection and Tracking, Text Mining, Trend Discovery, Visualization Evaluation, Dimensionality Reduction, Text Representation, Information Extraction, User Interaction.

Another field of study that relates to text visualization is Topic Detection and Tracking (TDT) [7], a DARPA-sponsored initiative to investigate the state-of-the-art in finding and tracking new events in a stream of broadcast news [8]. Pilot phase started in 1998 and final phase of the project ended in 2004. The developed methods for detection and tracking stories are elaborate and data-specific. They can be used as a background tool for text stream visualization in which the foreground drawing methods need not be complex but the visualizations can still be highly informative. An example

[https://www.fer.unizg.hr/\\_download/repository/KDI\\_-\\_Artur\\_Silic\\_-\\_Visualization\\_of\\_Text\\_Streams.pdf](https://www.fer.unizg.hr/_download/repository/KDI_-_Artur_Silic_-_Visualization_of_Text_Streams.pdf)

## 2. Masses d'images



**Robert Silvers, Photomosaic, 1995**

# Serendipité ?



# Mobile Interface Design for Browsing Large Collections of Images

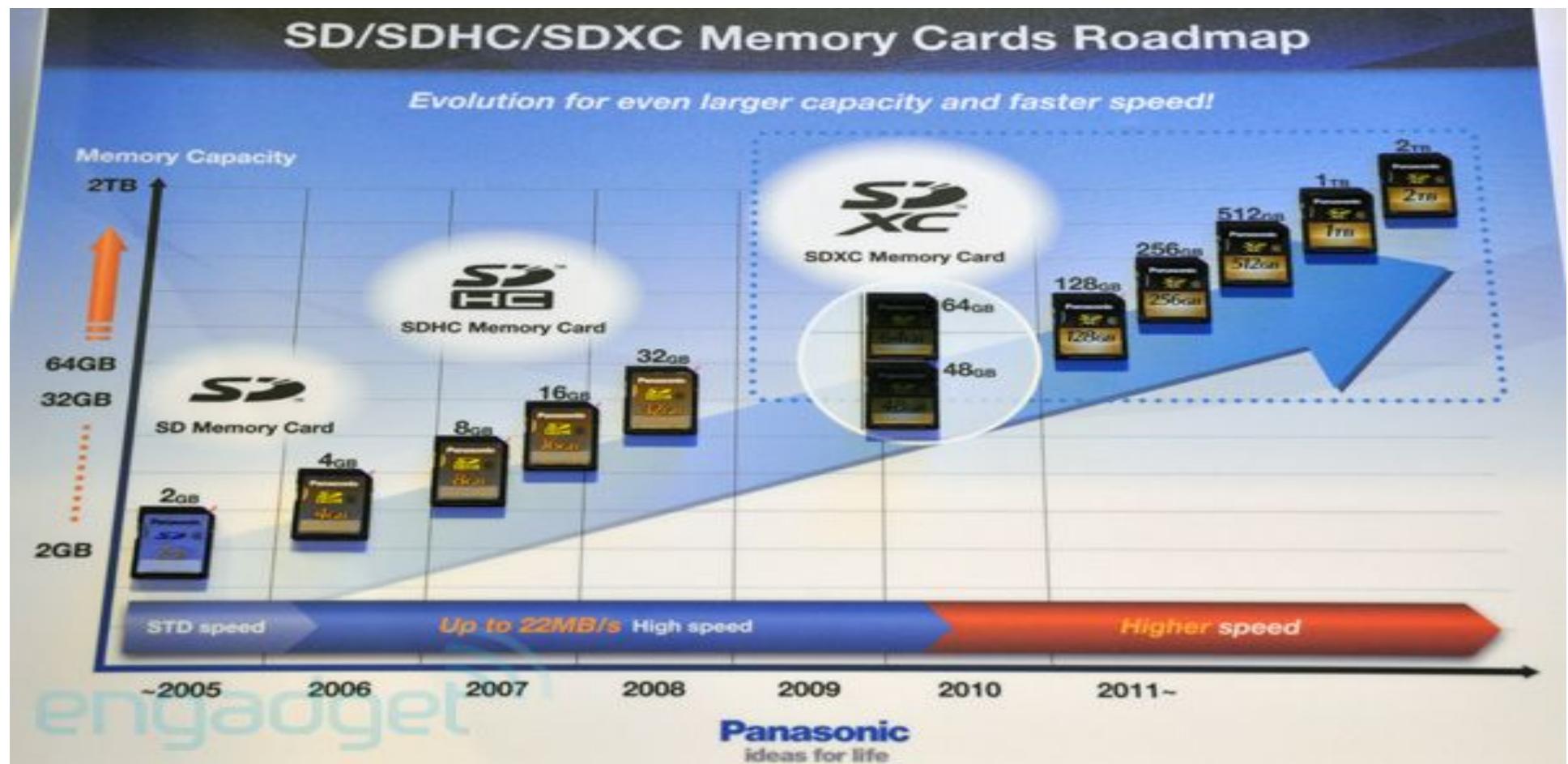
Shuo-Hsiu HSU

12.01.2011

# (Context of subject)

Digital photos in the context of mobility

- Availability of the image-capturing devices (low price, variation in products )
- The growth of the storage device



# (Photo Categorization)

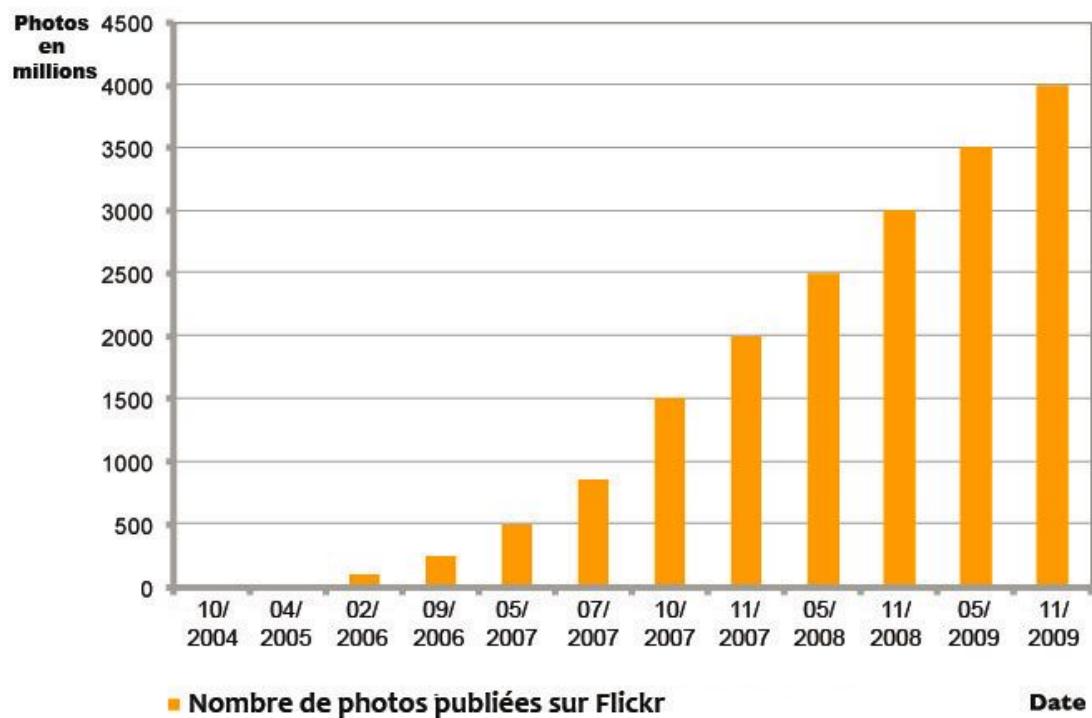


*Great photos are more than pictures;  
they are “**stories**”*

**Canon**

# (Context of subject)

- Online service facilitates the storage and encourage the sharing of photos
  - Google Picassa
  - Yahoo Flickr
  - Orange
  - Facebook\*  
*(the biggest one)*
  - Amazon web service
  - Apple iCloud



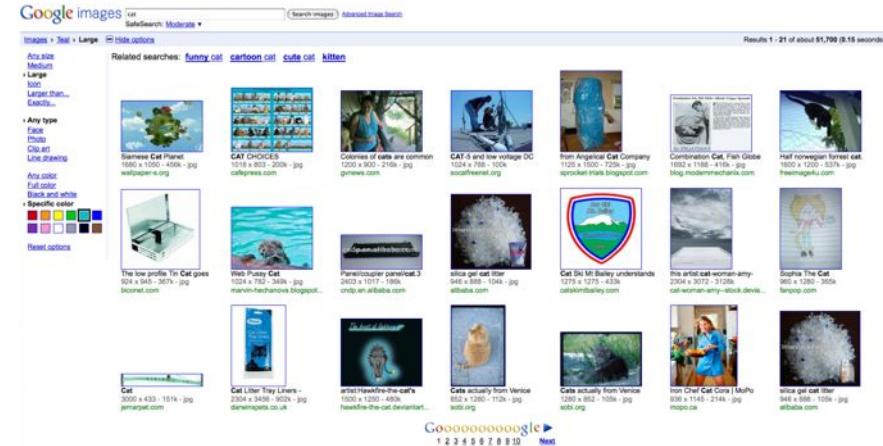
# (Context of subject)

Definition of the keywords in this subject :

- Photo albums = personal photo albums
- InfoVis = Casual Infovis
- Small screen
- Mobile device
- Unconventional interaction model (instead of mouse, joystick, keyboard)

# (Problematic)

- Conventional image browsers implemented with clustering algorithms are often **query-based**.



- **Single** usage to deal with **various** types of photos within conventional system tools, rarely integrated.

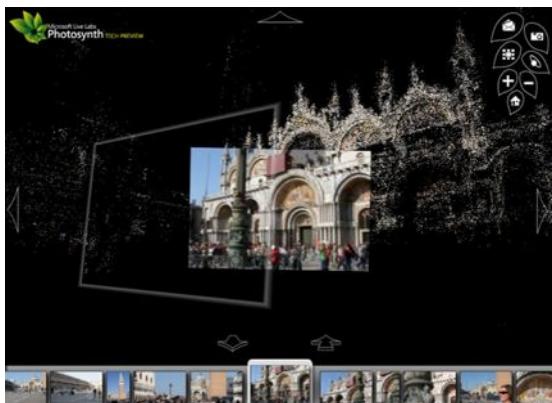
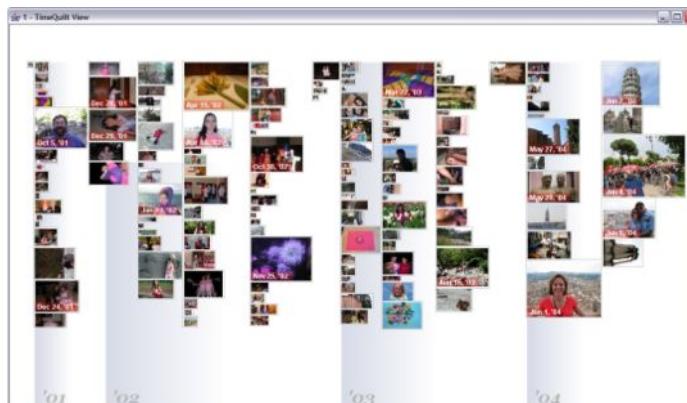


Photo tourism (Snavely et al., 2005)



Time Quilt (Huynh et al., 2005)

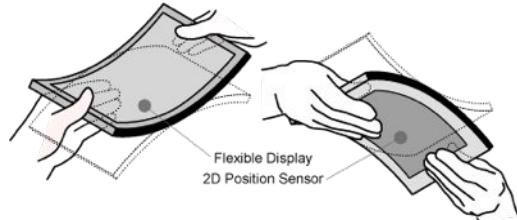


Face Bubble (Xiao et al., 2008)

# (Interaction Techniques)

What kind of natural interaction techniques for mobiles?

Haptic manipulation

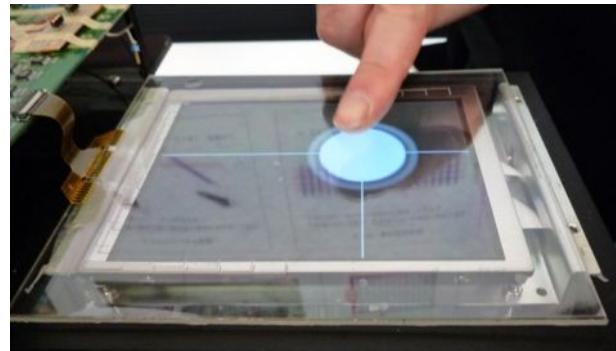


Gummi [Schwesig et al., 2003]



TWEND  
[Herkenrath et al., 2008]

Remote gesture



Mitsubishi 3D control panel, 2009



Toshiba Qosmio 55, 2009

Touch-based  
gesture



Apple iPad, 2010

# (Interaction Techniques)

What kind of natural interaction techniques for mobiles?

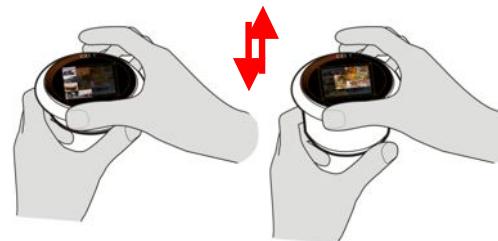
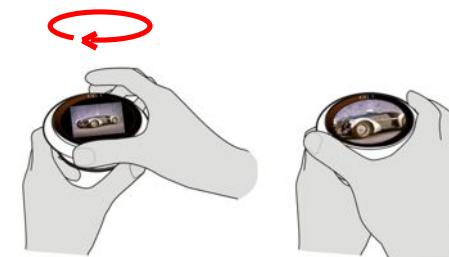
Haptic manipulation – upcoming news



Nokia kinetic future: flexible screen and twisted interface  
[Nokia, 2011]

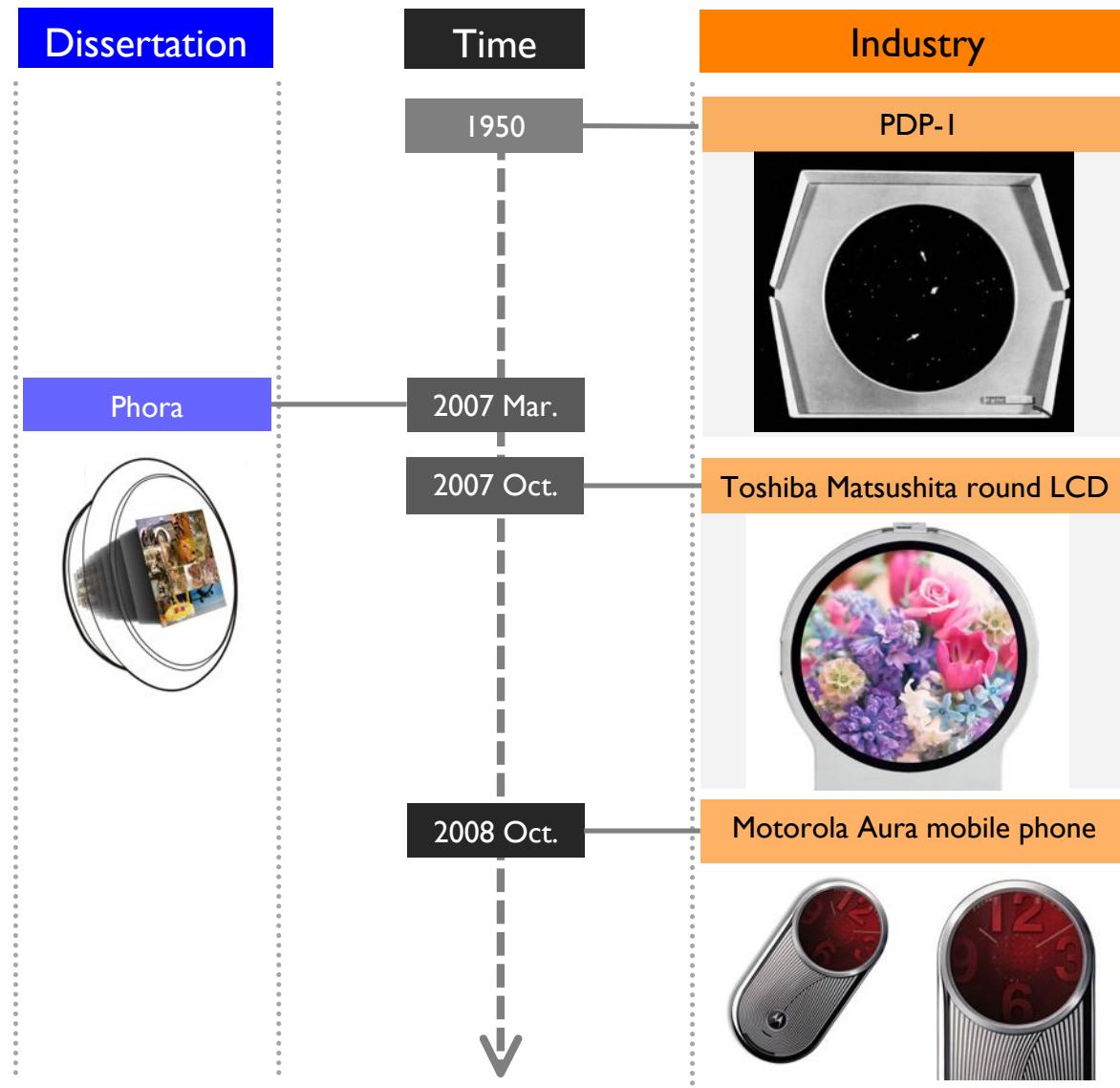
# (Device Design - Phora)

Design & 3 gestures



\* Phora is inspired from french noun « forage», which means digging.

# (Round screen)



# (Device Design - Phora)

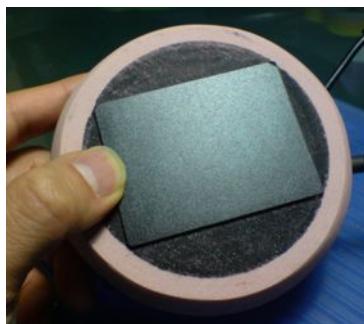
## Prototype



[demo]



SpaceNavigator



Substitute of  
round touch screen:  
TouchPad



SpaceNavigator in  
Phora

## Small scale user study



- 6 participants
- Evaluation of haptic manipulation
- Tasks simulating the usage in everyday scenario
- Involving different size of albums (300/1000P)

# (Device Design – Phora II)

## Two potential tasks of casual usage of photos on mobile devices

- Simple photo retrieval on mobiles (the album size is not large)
- Shuffling photos for reminiscing (recalling memories)

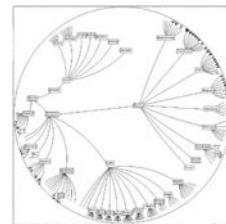
### Trends of mobile browsers

- Usage on mobile image browser is casual, users only need basic functions.
- As function requirements is low, the visual presentation dominates the interaction.

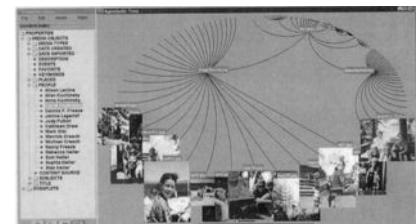
### Overview of circular visualization / interface



Fisheye  
(Furnas, 1986)



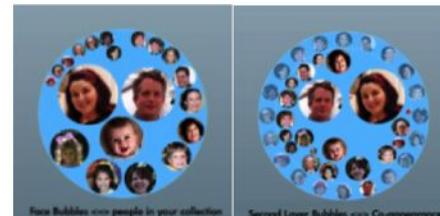
Hyperbolic tree  
(Lamping, 1995)



Fotofile  
(Kuchinsky et al., 1999)



Circular interface  
(Shen et al., 2003)



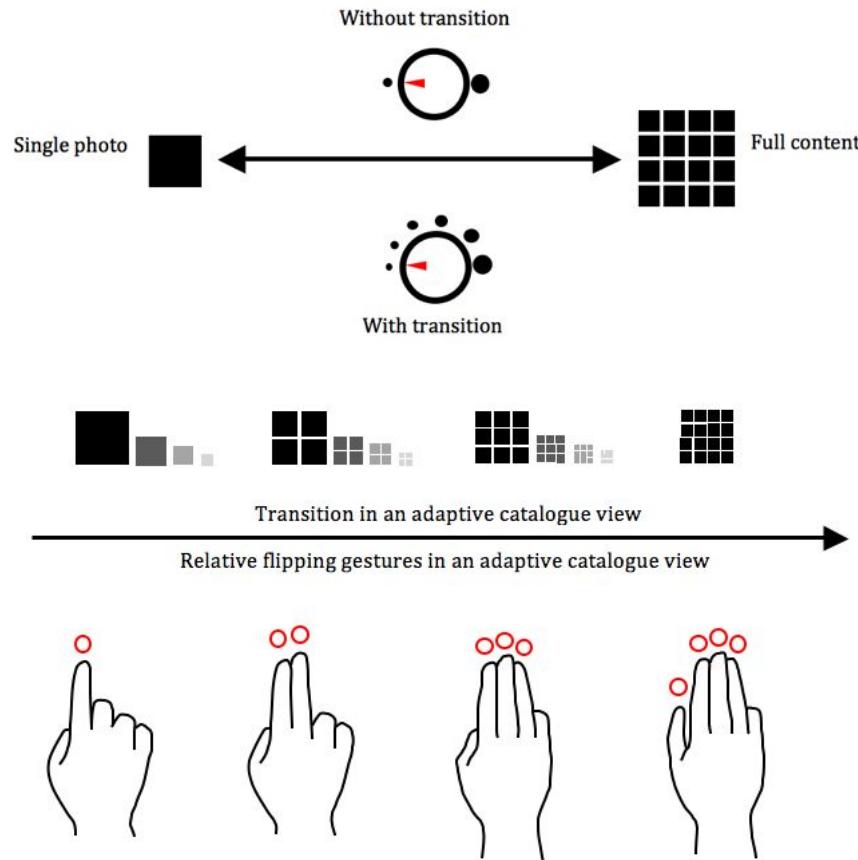
Face Bubble  
(Xiao & Zhang, 2008)



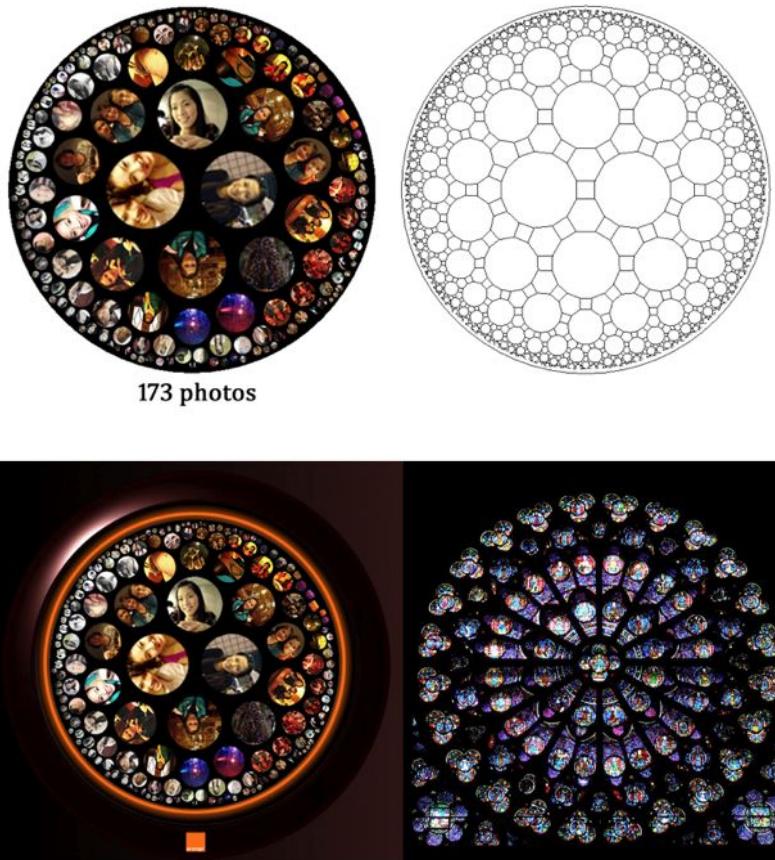
MyPhotoEgg  
(2008)

# (Device Design – Phora II)

## Potential mechanism in image browser

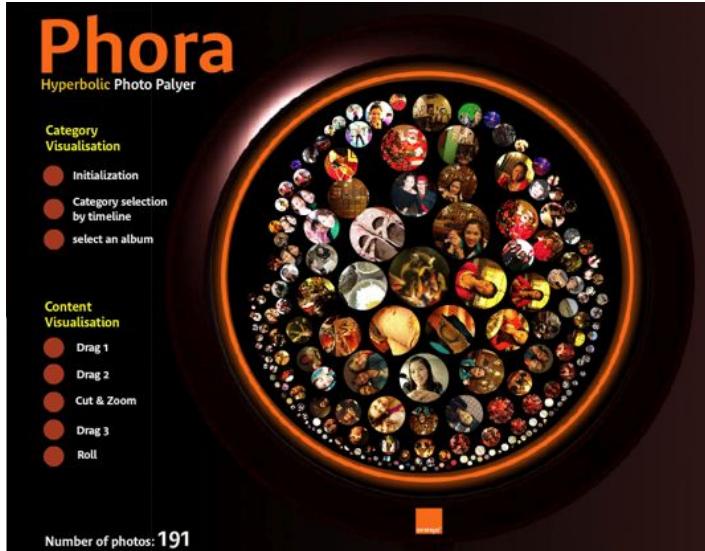


## Phora based on hyperbolic tree



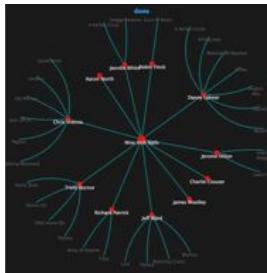
# (Device Design – Phora II)

## Design

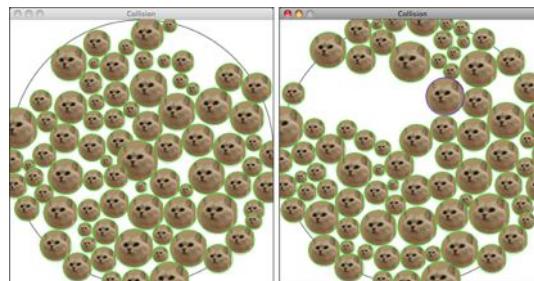


demo

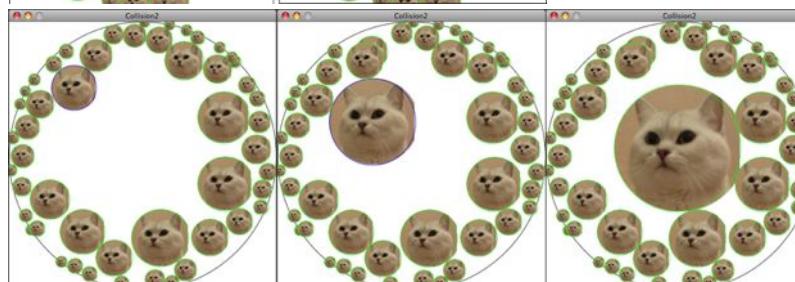
## Prototyping problem



Hypertree by  
Javascript  
InfoVis  
toolkits



Progress in  
computing  
collision in  
Processing.



Progress in  
zooming in  
the center.

### 3. Masses de livres (les bibliothèques numériques)

# machines à lire

AUXILIAIRES

LIVRE



SUBSTITUTS



Microfilm



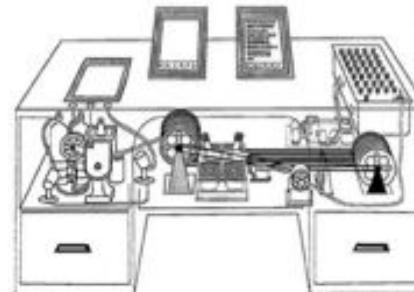
P. Otlet, livre téléphoté, 1930



JCR Licklider, Library of the future, 1965



A. Korn, FAX, 1907



V. Bush, MEMEX, 1945

# Vanevar BUSH : MEMEX

@ The Atlantic | July 1945 | As We May Think | ...

Précédente Suivante Arrêter Actualiser Démarrage Remplissage automatique Imprimer Courrier

Adresse <http://www.theatlantic.com/unbound/flashbks/computer/bushf.htm>

Mac OS X Google webmail Abebook deptinfo StatHome

**Atlantic** ACCELERATED INTELLIGENCE starts here! Click here to SUBSCRIBE- Atlantic THE NEW EDITION No Risk!

Search [Tips] Go

**THE Atlantic online**

E-MAIL ARTICLE PRINTER FORMAT SUBSCRIBE TO THE ATLANTIC

The Atlantic Monthly | July 1945

**AS WE MAY THINK**  
BY VANNEVAR BUSH

As Director of the Office of Scientific Research and Development, Dr. Vannevar Bush has coordinated the activities of some six thousand leading American scientists in the application of science to warfare. In this significant article he holds up an incentive for scientists when the fighting has ceased. He urges that men of science should then turn to the massive task of making more accessible our bewildering store of knowledge. For years inventions have extended man's physical powers rather than the powers of his mind. Trip hammers that multiply the fists, microscopes that sharpen the eye, and engines of destruction and detection are new results, but not the end results, of modern science. Now, says Dr. Bush, instruments are at hand which, if properly developed, will give man access to and command over the inherited knowledge of the ages. The perfection of these pacific instruments should be the first objective of our scientists as they emerge from their war work. Like Emerson's famous address of 1837 on "The American Scholar," this paper by Dr. Bush calls for a new relationship between thinking man and the sum of our knowledge. —THE EDITOR

T his has not been a scientist's war; it has been a war in which all have had a part. The scientists, burying their old professional competition in the demand of a common cause, have shared greatly and learned much. It has been exhilarating to work in effective partnership. Now, for many, this appears to be approaching an end. What are the scientists to do next?

For the biologists, and particularly for the medical scientists, there can be little indecision; for their war has hardly required them to leave the old paths. Many indeed have been able to carry on their war research in their familiar peacetime laboratories. Their objectives remain much the same.

It is the physicists who have been thrown most violently off stride, who have left academic pursuits for the making of strange destructive gadgets, who have had to devise new methods for their unanticipated assignments. They have done their part on the devices that made it possible to turn back the enemy, have worked in combined effort with the physicists of our allies. They have felt within themselves the stir of achievement. They

Return to "Flashback: Prophets of the Computer Age"

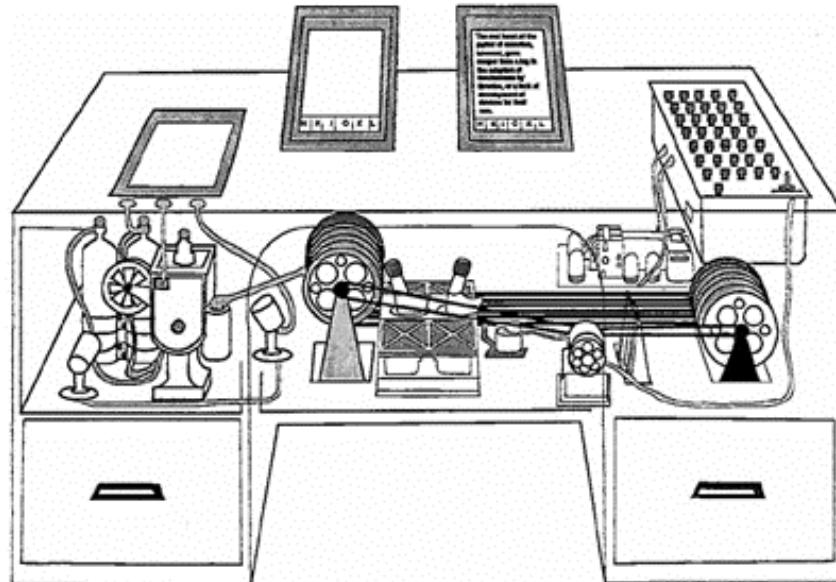
More on technology from *The Atlantic Monthly*.

Be the first to sport *Atlantic* merchandise

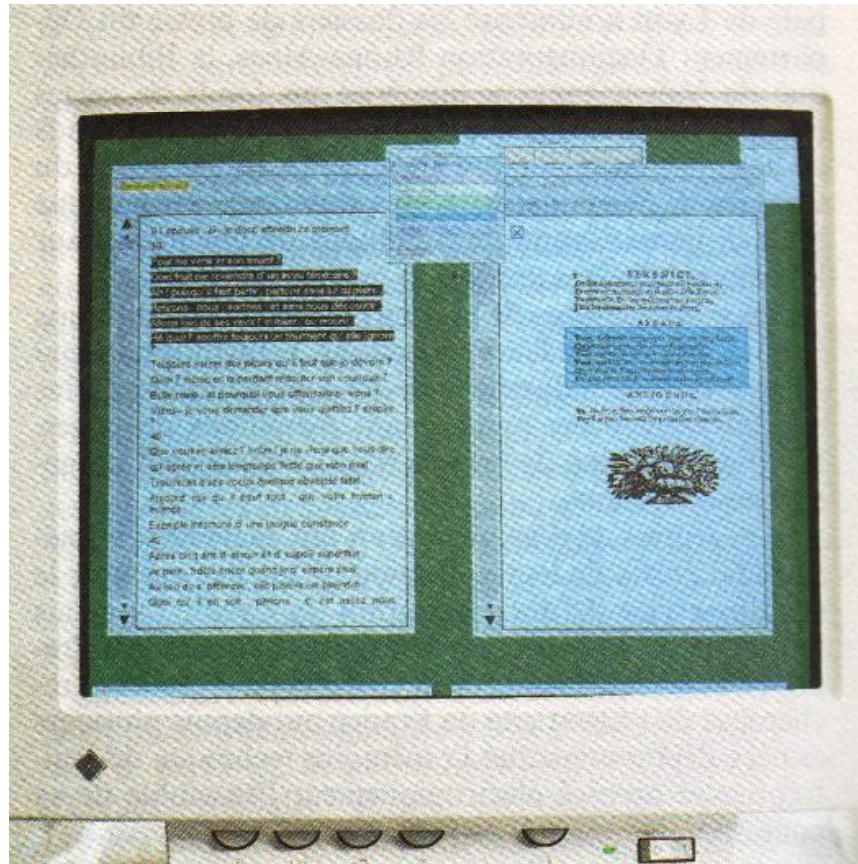
Perfect your French (click here)  
Maintain Spanish (click here)  
Improve German (click here)  
Enrich your Italian (click here)

SUBSCRIPTION CHAMPION AUDIOMAGAZINE

click here



# MEMORIA : Le PLAO (1988 - 1997)

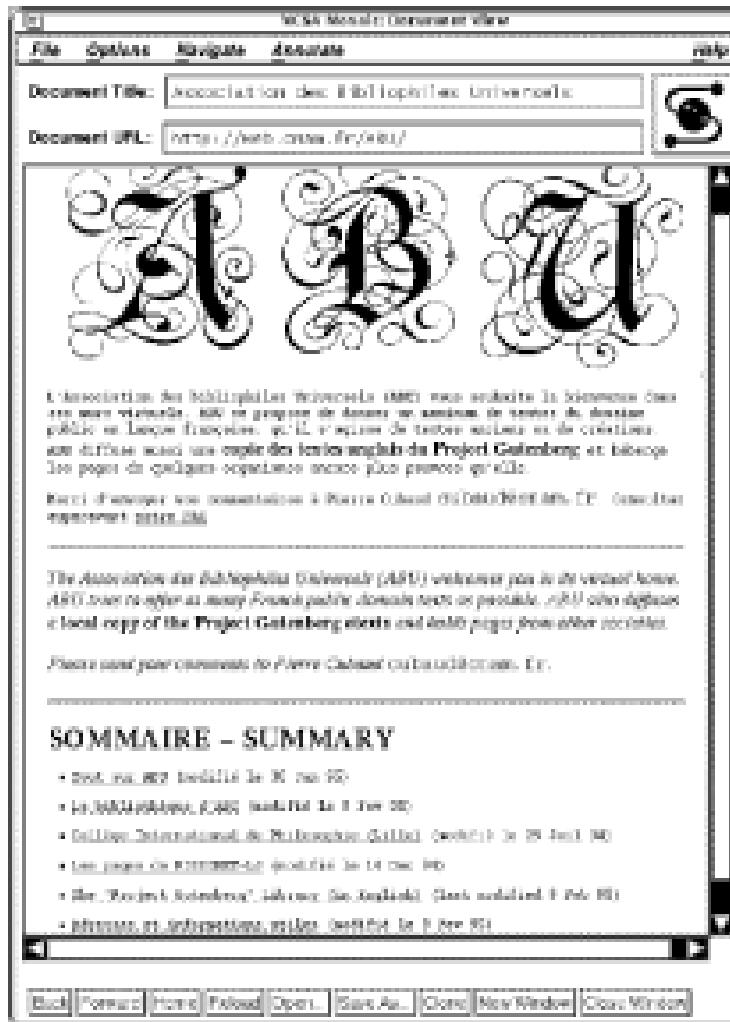


J. VIRBEL. Reading and managing texts on the Bibliothèque de France station. in G.P. Landow, P. Delany (eds.) *The digital word. Text-based computing in the humanities*. MIT Press, 1993

Y. MAIGNIEN. Chronique partielle d'une bibliothèque virtuelle. *Doc. numérique* vol. 2 (3-4), pp. 281-290

A. PELFRENE. Les postes de lecture assistée par ordinateur (PLAO) in C. PAGANELLI (dir.) *Interaction homme-machine et recherche d'information*. Hermès, 2002.

# 1993 : first web site for digital library (French classics)



<http://abu.cnam.fr> (still online)



Conservatoire numérique des Arts et Métiers

Bibliothèque numérique en histoire des sciences et des techniques  
Au 27/10/2014, le Cnum contient 1144 titres, 2255 volumes, 692 098 pages,  
2223 unités téléchargeables

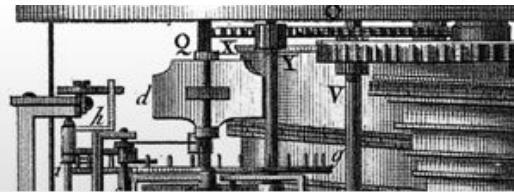
À PROPOS

CATALOGUE GÉNÉRAL

RECHERCHER DANS LE :



RECHERCHE AVANCÉE



Le site du Cnum fait peau neuve ! Merci d'adresser vos remarques et questions à cnum@cnam.fr

### Thématiques



Catalogues de constructeurs  
217 titres



Expositions universelles  
426 titres



Technologies de l'information et de la communication  
217 titres



Construction  
117 titres



Histoire du Cnam  
18 titres



Transports  
158 titres



Énergie  
102 titres



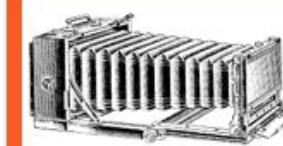
Machines et instrumentation scientifique  
236 titres



Généralités scientifiques et vulgarisation  
13 titres

### Nouveautés

#### LES DÉBUTS DE LA PHOTOGRAPHIE



Voir les dernières mises en ligne

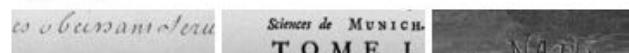
### Expositions

#### PROCHAINEMENT



Palais, pavillons et galeries : les bâtiments des expositions universelles en représentation (1798-1900)

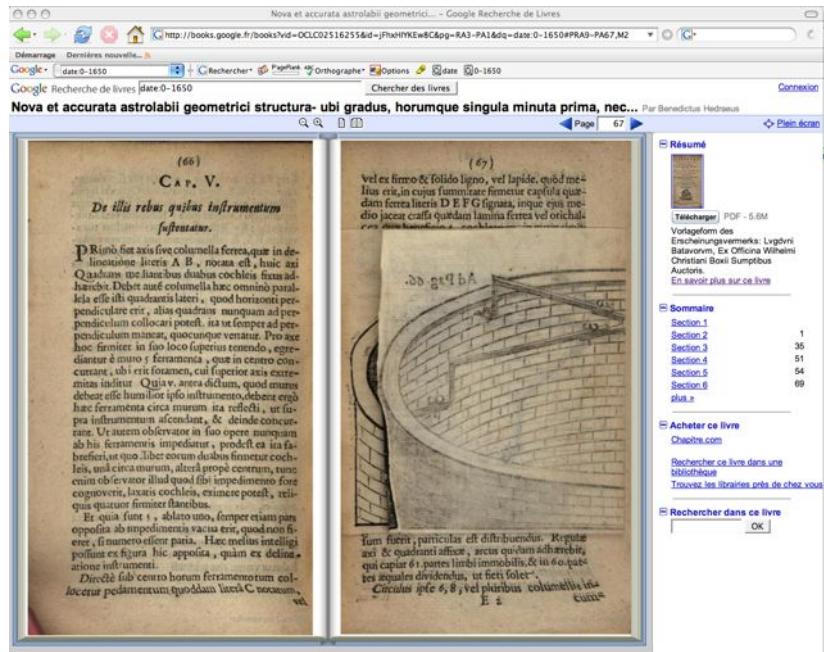
### Typologies



≥ 1998 : <http://cnum.cnam.fr>

± 1Mpages

# Research challenges ?

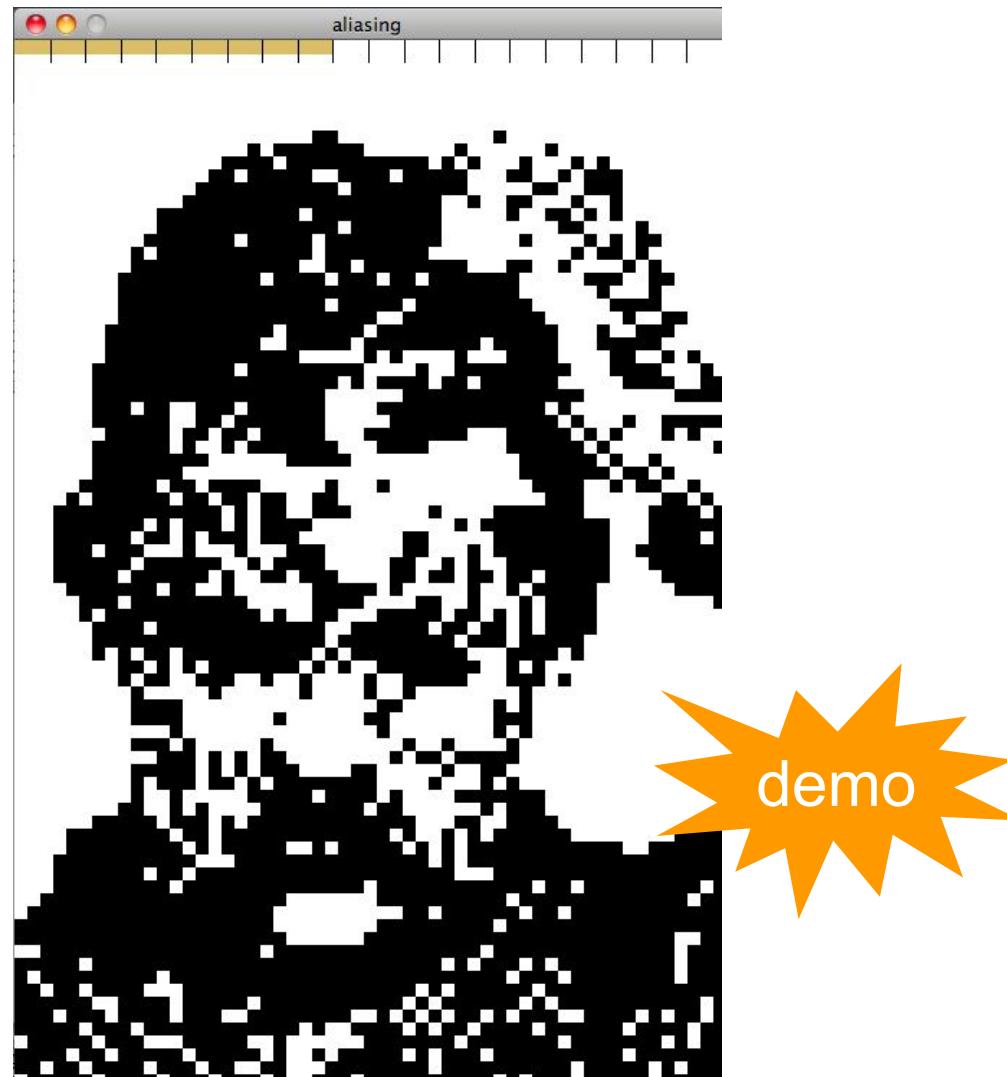


2005 : google books



2010 : Apple ipad

# Digitization ?



# strolling ?

The screenshot shows the Google Book Search homepage. At the top is the Google logo with "Book Search" below it and "BETA" to the right. Below the logo is a search bar with a "Search Books" button to its right. To the right of the search bar are links for "Advanced Book Search" and "Google Book Search Help". Underneath the search bar is a radio button group labeled "Search:  All books  Full view books". A large green banner below the search bar reads "Search the full text of books (and discover new ones)". At the bottom of the page are links for "About Google Book Search", "Information for Publishers", and "Google Home".



<http://www.rauzier-hyperphoto.com/voyages-extraordinaires/>

# Au Xerox Parc

Stuart K. Card, George G. Robertson, and William York

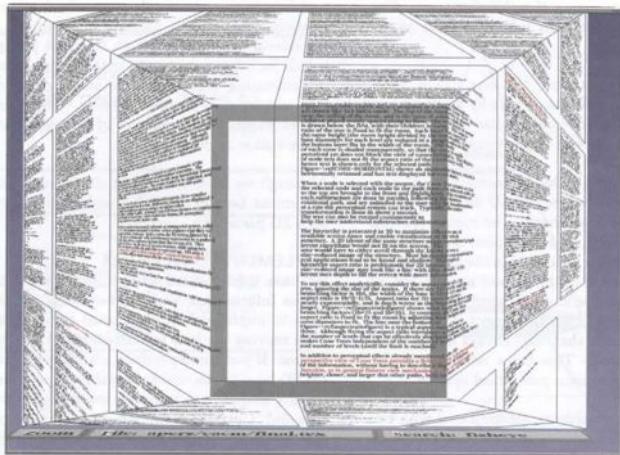
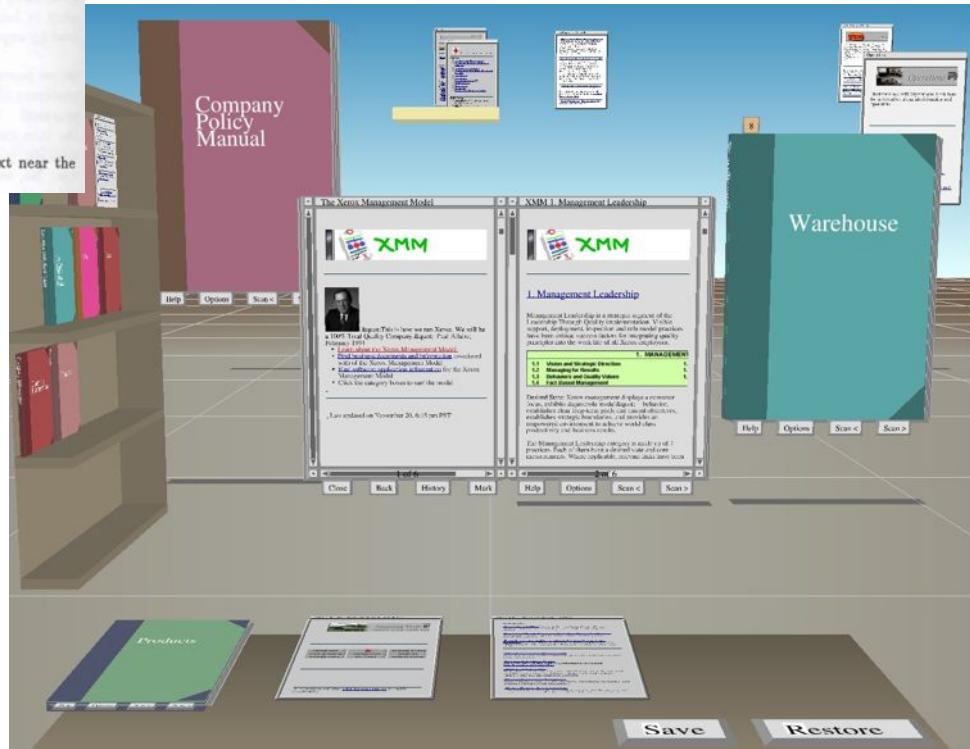


Figure 3: Document Lens with lens pulled toward the user. The resulting truncated pyramid makes text near the lens' edges readable.



The WebBook and the Web Forager: An Information Workspace for the World-Wide Web

ACM CHI'1996



# Et Microsoft

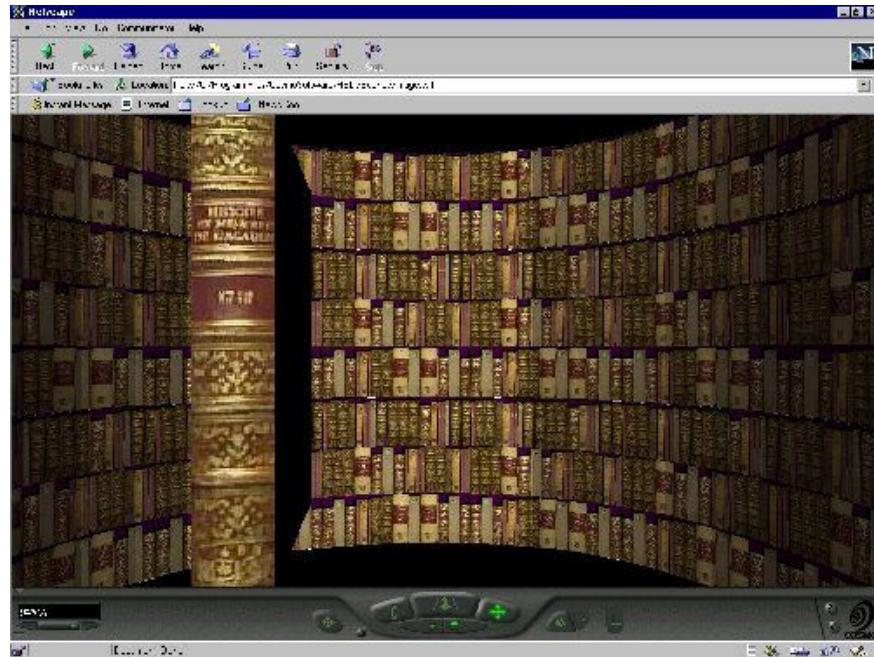


Data mountain  
Robertson et al UIST'1998

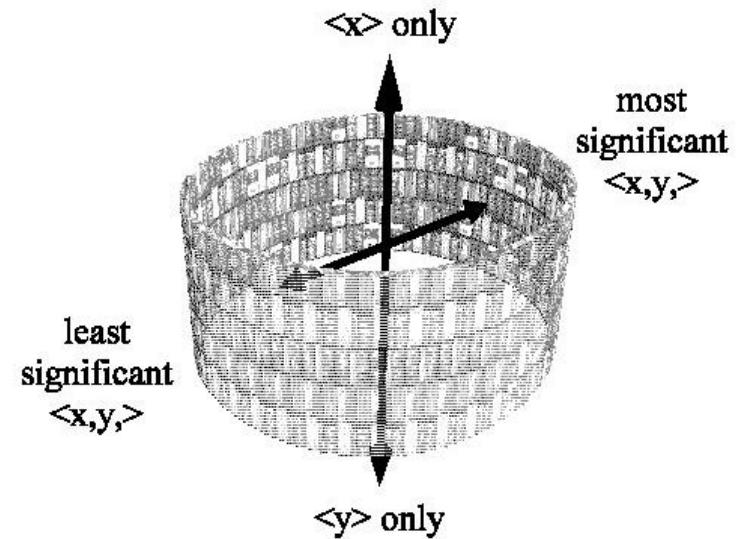
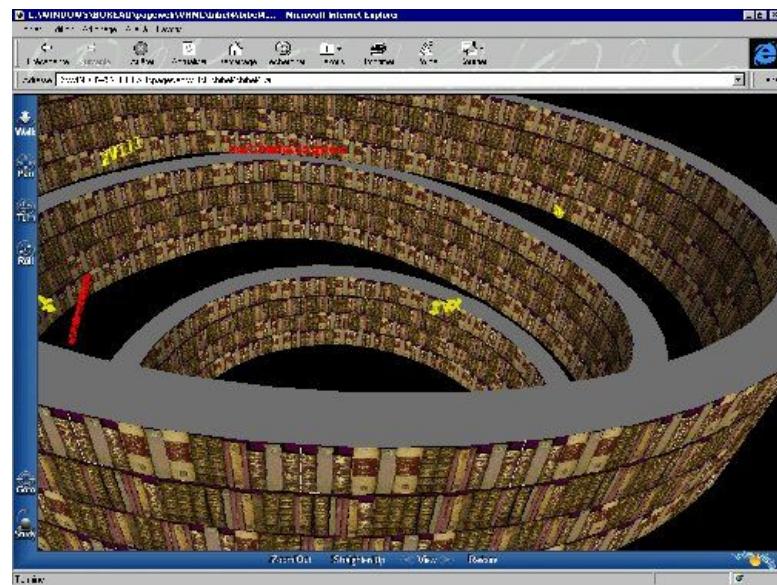


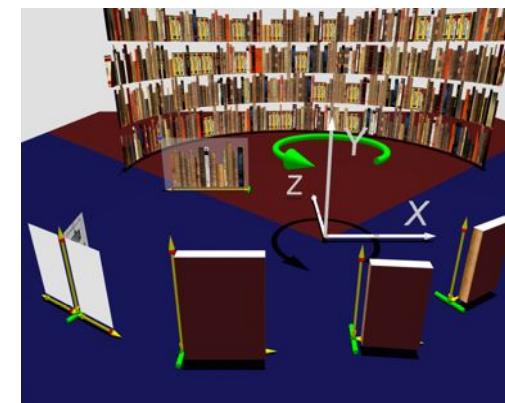
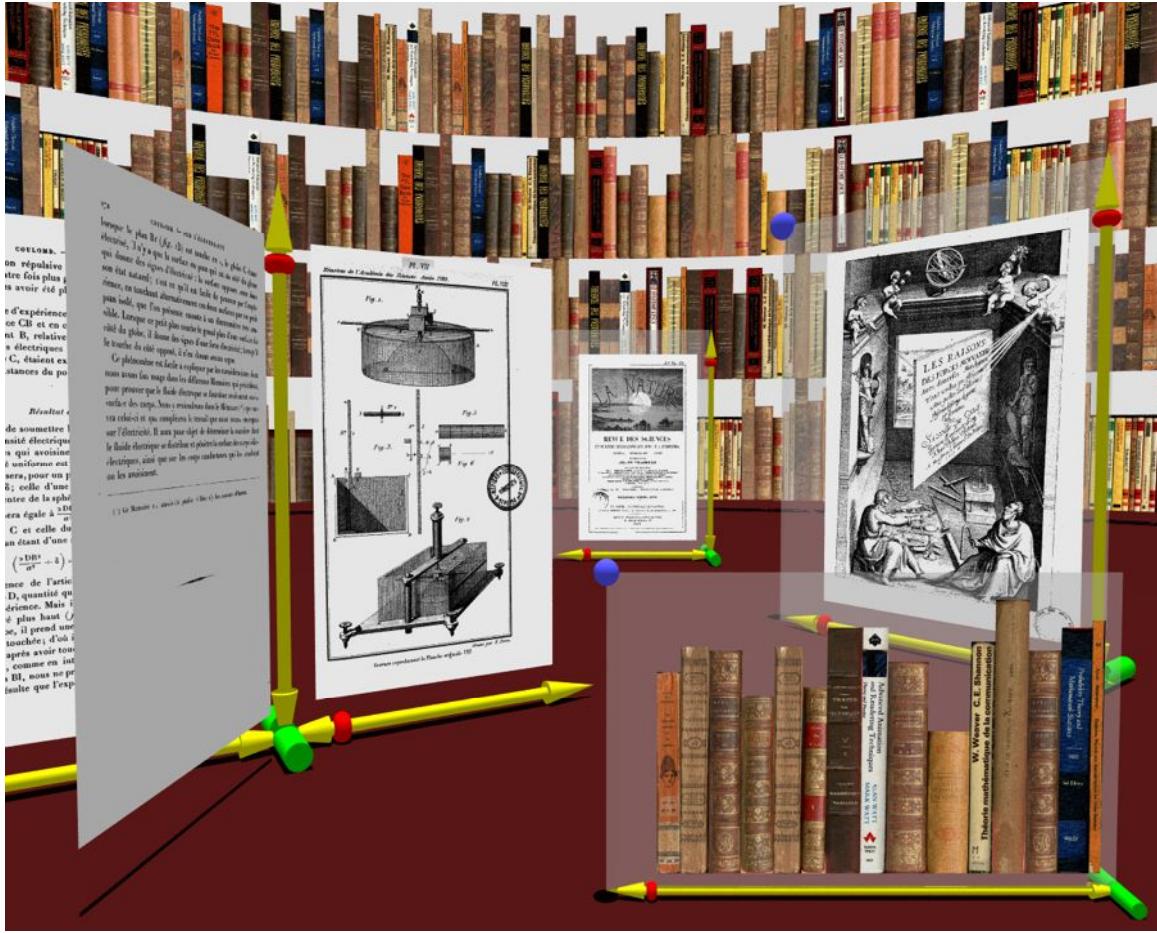
Task Gallery  
Robertson et al. CHI'2000

# Au CNAM



Cubaud, Thiria, Topol - ACM  
DL'1998





Cubaud et al. ACM-IEEE joint conf. digital libraries, 2002

# Skimming ?

Google Recherche de livres

[À propos de ce livre](#)

**Mémoir**



De Antoine La  
Lavoisier

Publié 1920  
[Gauthiers-Villa](#)

67 pages

Copie de l'exe

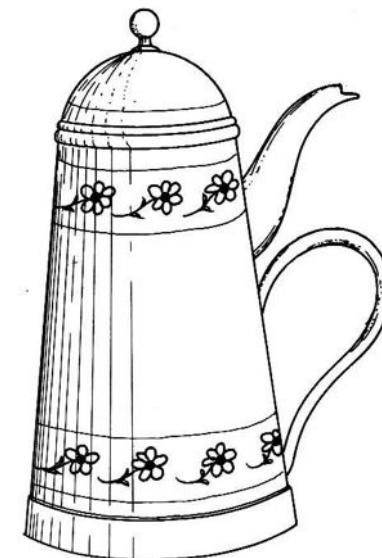
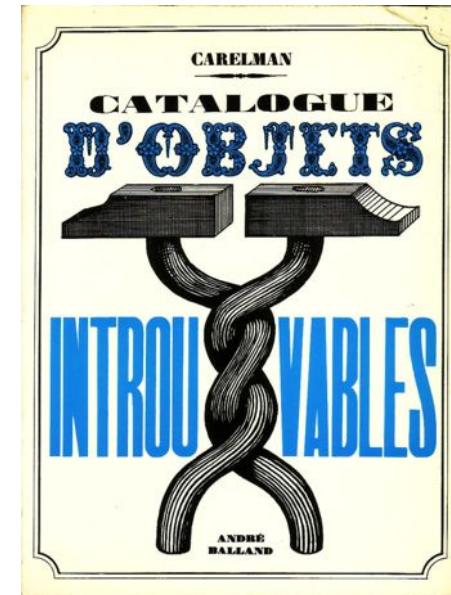
Numérisé le 8

[Ajouter à ma b](#)

[Rédiger un co](#)

**Rechercher dans ce livre**

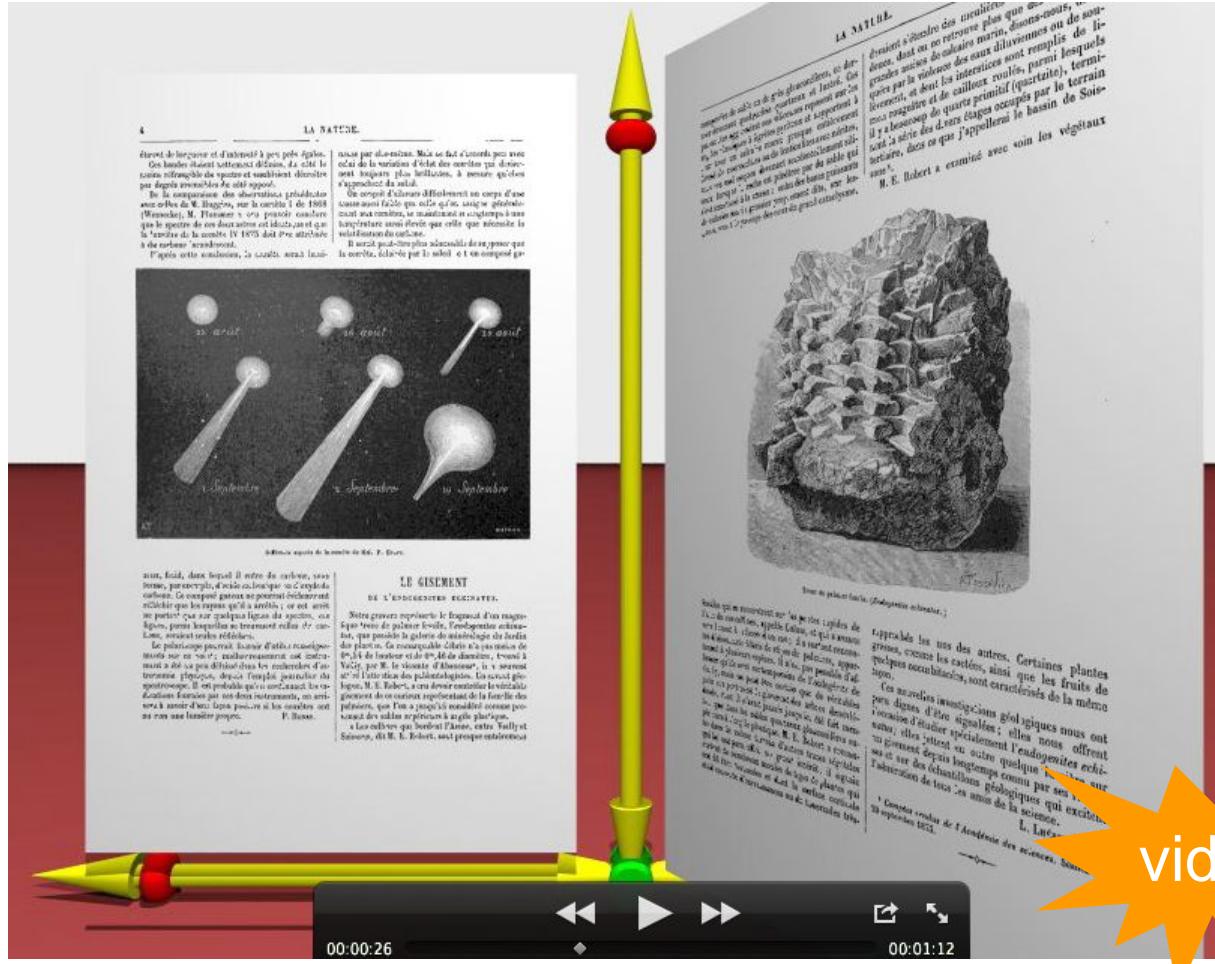
# affordance



E19 — Cafetière pour masochiste.  
Nous pensons que le dessin est suffisamment explicite pour ne pas s'appesantir sur des détails qui pourraient s'avérer pénibles.

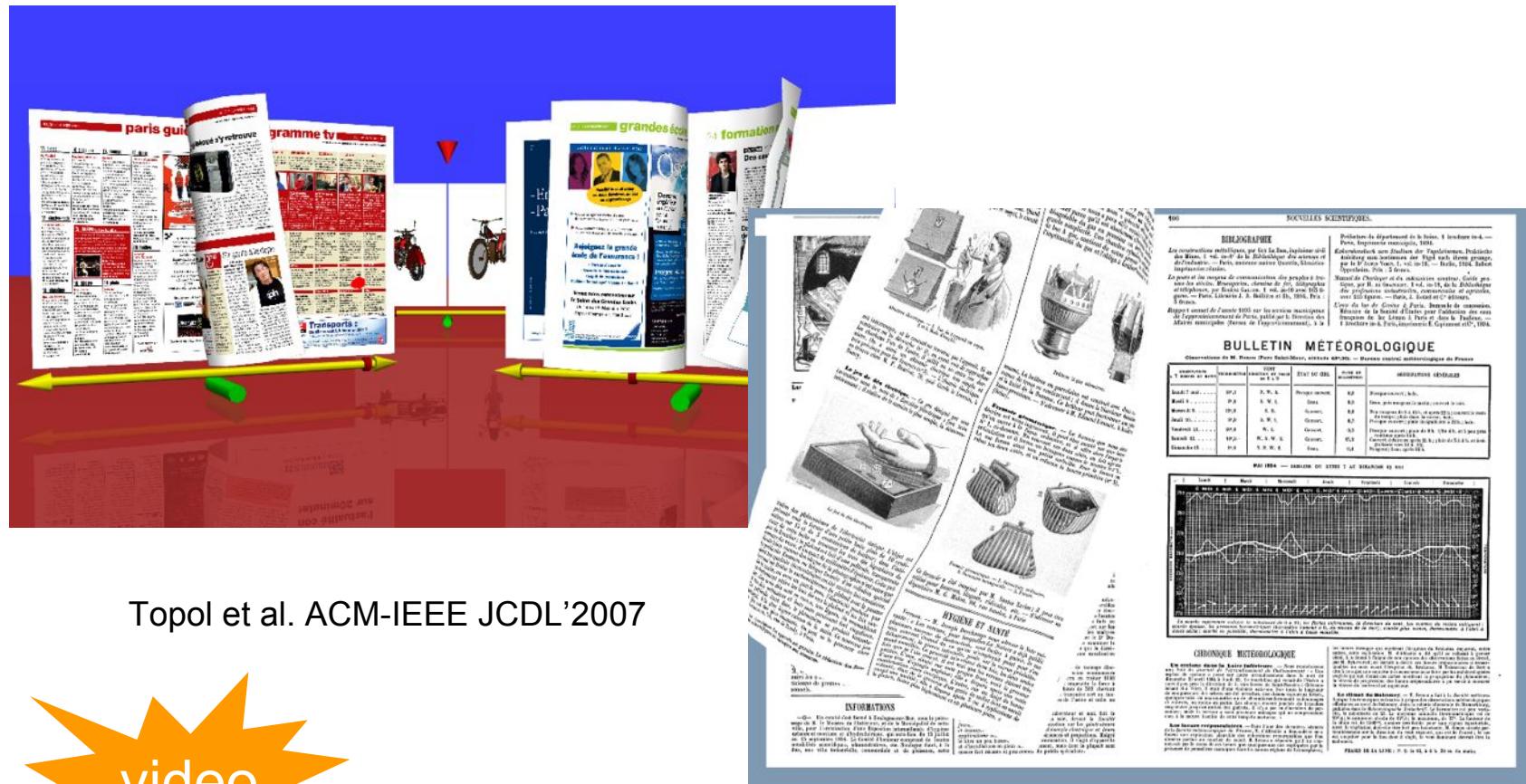


« Turning the page » CARPENTER L., SHAW S., PRESCOTT A. (eds)  
Towards the digital library : The british library initiatives for access programme.  
London, British lib., 1998 (p. 61 et suiv.)



# Cubaud, Topol ACM Web3D'2001

# moteur physique



Topol et al. ACM-IEEE JCDL'2007



Topol et al. CIDE'2012

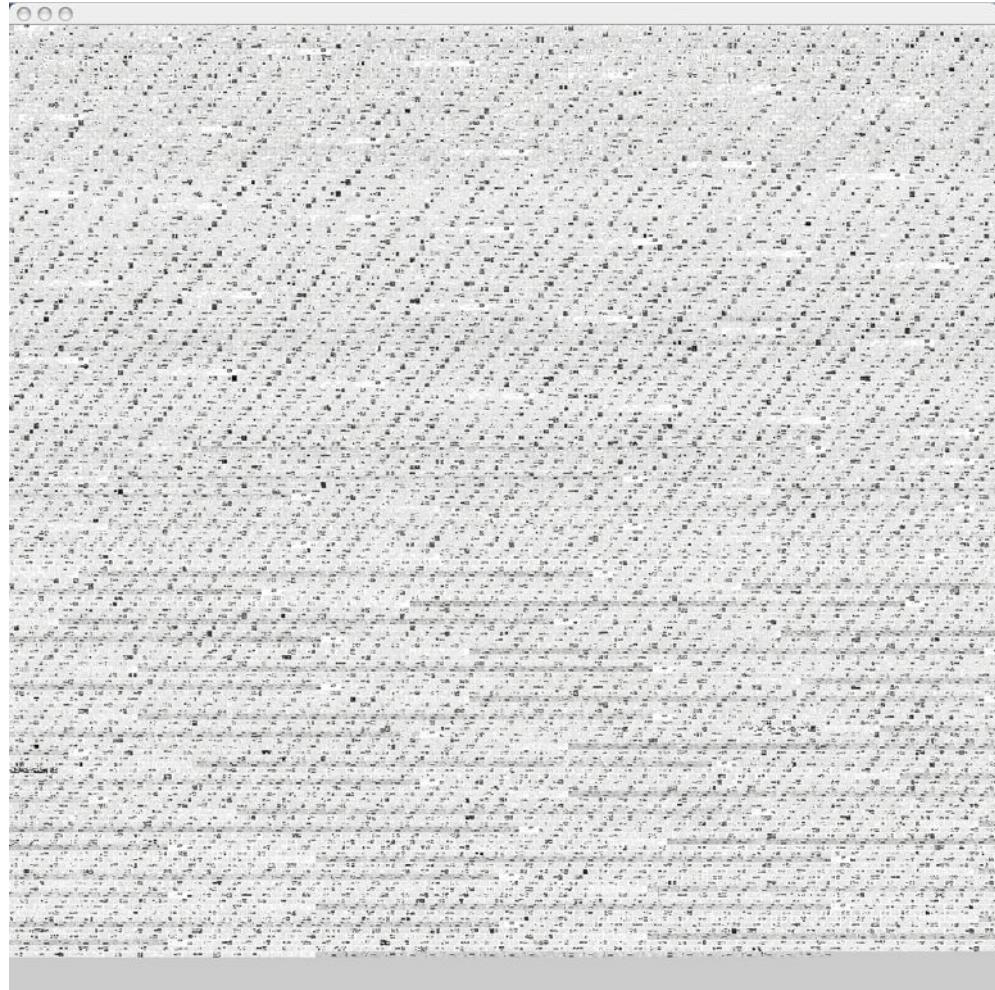
feuilleter en survol = 500 images/seconde !!!



T. Nakashima, Y. Watanabe, T. Komuro, M. Ishikawa  
U. Tokyo, mars 2010

Jer-  
ue; il  
naire  
ursik  
evant  
e pile  
isente  
votre  
, La

le vos  
abon-  
ze ans  
e, me  
le, je  
it l'é-  
rairie  
rquai  
iméro  
Je ne  
parer  
nfin à  
naga-  
ndai,  
le nu-  
al qui  
tant.  
it de  
nicile  
n un  
com-  
frap-  
porte



Thirty first years of "la Nature" = 32500 pages