

Expérimentations sémantiques

Autour de la Chanson de Roland

Jacques Ducloy, Thierry Daunois, Isabelle Turcan
<https://wicri-demo.istex.fr/wicri-chanson-roland.fr>

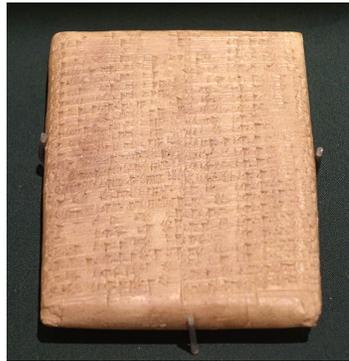
Bonjour, mon nom est Jacques Ducloy

Devenu humaniste numérique à 76 ans

- ▶ 1959 Je regarde par la fenêtre pendant la leçon sur la *Chanson de Roland*
- ▶ 1967 Calcul numérique (ENSEM) puis informatique en Lorraine
 - Crin=>Loria, IUCAL (=> CIRIL) / usagers : CRPG... et surtout le TLF
- ▶ 1980 Création GS Association Nationale du Logiciel (ADI, CNRS, CNET...)
 - Enquêtes / numérique Serveur Minitel Catalogues / Expositions (y compris USA)
- ▶ 1988 ... 2000, INIST (directeur informatique, production, R&D)
 - SGML pour la bibliothéconomie et veille stratégique => ennemi public numéro 1
- ▶ 1992 : Loria (Autoroutes de l'Information / Site Art Nouveau)
 - Ouverture internationale : ERCIM/Digital libraries ; DCMI (Dublin Core)
- ▶ 2007 : Contributions Wikipédia (Saul de Haendel... Philatélie)
- ▶ 2008 : Wicri (DRRT Lorraine ... ISTE... Paragraphe/Paris 8)
 - Avec Thierry Daunois, Frédérique Péguiron et Jean-Pierre Thomesse
- ▶ 2021 : Projet Chanson de Roland
 - avec une composition de Gilles Mathieu
et une problématique d'Isabelle Turcan

Changements de paradigmes pour les documents et bibliothèques

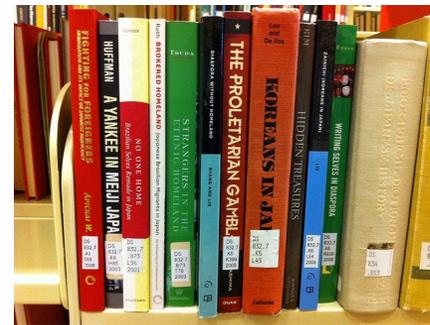
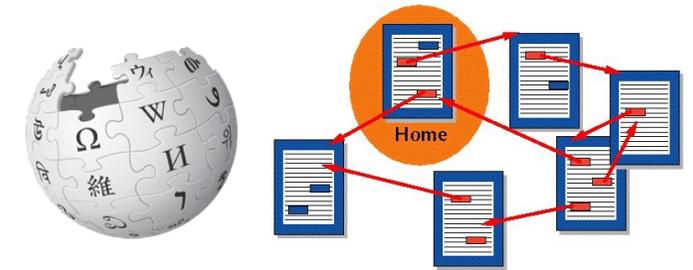
PAGE



CODEX



HYPERTEXTE
informationnel



HYPERTEXTE
Computational



Changement de paradigme

Pour les chercheurs et praticiens humanistes

- ▶ Avec les wikis programmables et les boîtes à outils XML
 - Un chercheur ou un praticien
 - peut devenir totalement indépendant pour développer un espace de travail numérique correspondant à ses besoins.
 - peut le partager avec ses collègues dans une stratégie collective.
 - L'informaticien développeur programmeur perd un peu sa place.
 - Mais un réseau de spécialistes demeure indispensable dans un protocole de type accompagnement.

Pour tous les humanistes : Apprendre à lire et à écrire dans d'immenses hypertextes



Plan

- ▶ Le contexte : Wicri, Dilib, ISTEEX
- ▶ La musique dans Wicri (et Wikipédia)
- ▶ La Chanson de Roland
 - Une thématique riche et complexe
 - Un gigantesque hypertexte
 - Actions sémantiques structurantes
- ▶ Bilan et perspectives

Galaxie WikiMedia Foundation

Un nouveau GAFA ???



Ouvrages



FR: 2.388.429
 11.728.290
 EN : 6.436.960
 54.946.563

[Wikipédia](#)



Wiktionnaire
 Le dictionnaire libre

FR : 4 298 952
 EN : 6,912,797

[Wiktionnaire](#)

Entrepôts



80.001.464
 images
 multimédia

[Commons](#)

Bibliothèques

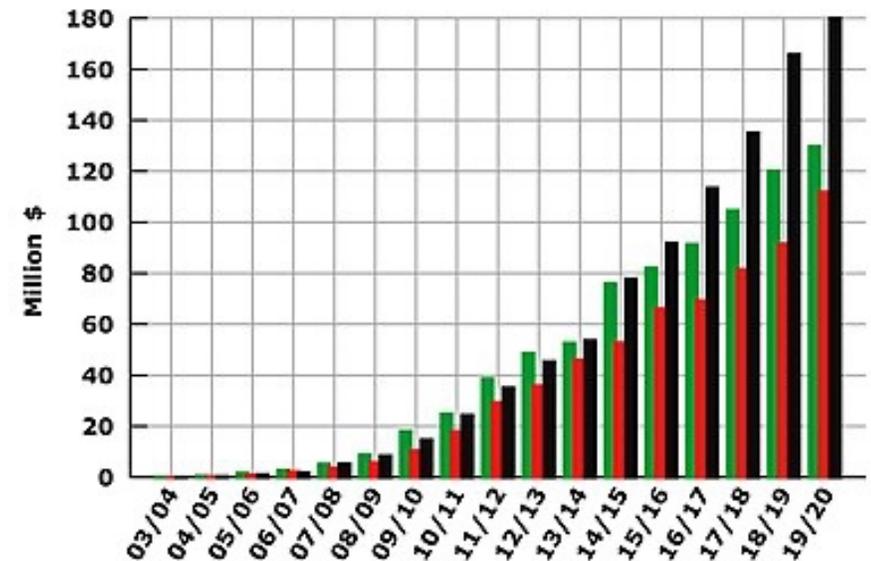


333.000
 Livres en français

[Wikisource](#)

Wikimedia Foundation Support and Revenue, Expenses and Net Assets at Year End

(Net Assets do not include funds in Wikimedia Endowment and Tides Advocacy. Expenses 2015–2020 include \$25 million paid to Wikimedia Endowment.)



MediaWiki :



Un moteur, un peu complexe à maintenir, mais très puissant

- ▶ Un moteur programmable par le contributeur
- ▶ Le moteur de Wikipédia => Compatibilité (modèles, extensions)
- ▶ Volumétrie conséquente :
 - 48.000.000 de pages (anglais) : 6.000.000 articles, 42.000.000 de métadonnées
- ▶ Wikipédia
 - Obligation de « *sourcer* » (pas de contributions innovantes)
 - Organisation par type de données (encyclopédie, ouvrages)

Aspects spécifiques de Wicri

- ▶ Contributions signées (pas d'anonymat)
- ▶ Espace de publication (textes originaux)
- ▶ Réseau de wikis en organisation thématique
 - Différents types d'informations dans un wiki
 - Pour des besoins complémentaires



Sites MediaWiki

► Dizaines de milliers



Ouvrages

Entrepôts

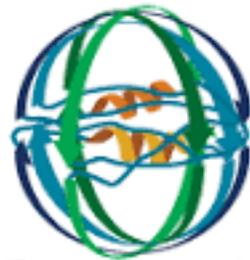
Bibliothèques



Wiktionnaire
Le dictionnaire libre



Free Choral Music



PROTEOPEDIA
— LIFE IN 3D —



Réseau Wicri
50 familles multilingues
Contributeurs identifiés



Observatoire
encyclopédique



Encyclopédie
+ serveurs



Encyclopédie +
+ publications



Encyclopédie + ...
+ bibliothèque



Hypertexte +++

Boîte à outil XML



- ▶ Ilib 1990 (INIST) – Dilib 1992 (Loria)
- ▶ Premières expérience en lex (préfiguration XSLT...)

```
<doc><tit>Tintin au Congo</tit><aut>Hergé</aut><date>1948</date><doc>  
<doc><tit>Tintin en Amérique</tit><aut>Hergé</aut><date>1949</date><doc>
```

```
%%  
"<doc>"      ;  
"</doc>"     printf("\n\n");  
"<tit>"      printf("<b>");  
"</tit>"     printf("</b>\n");  
"<aut>"      printf("::");  
"</aut>"     ;  
"<date>"     printf (" (");  
"</date>"    printf (")");  
%%
```

```
<b>Tintin au Congo</b>  
::Hergé (1948)  
  
<b>Tintin en Amérique</b>  
::Hergé (1949)
```

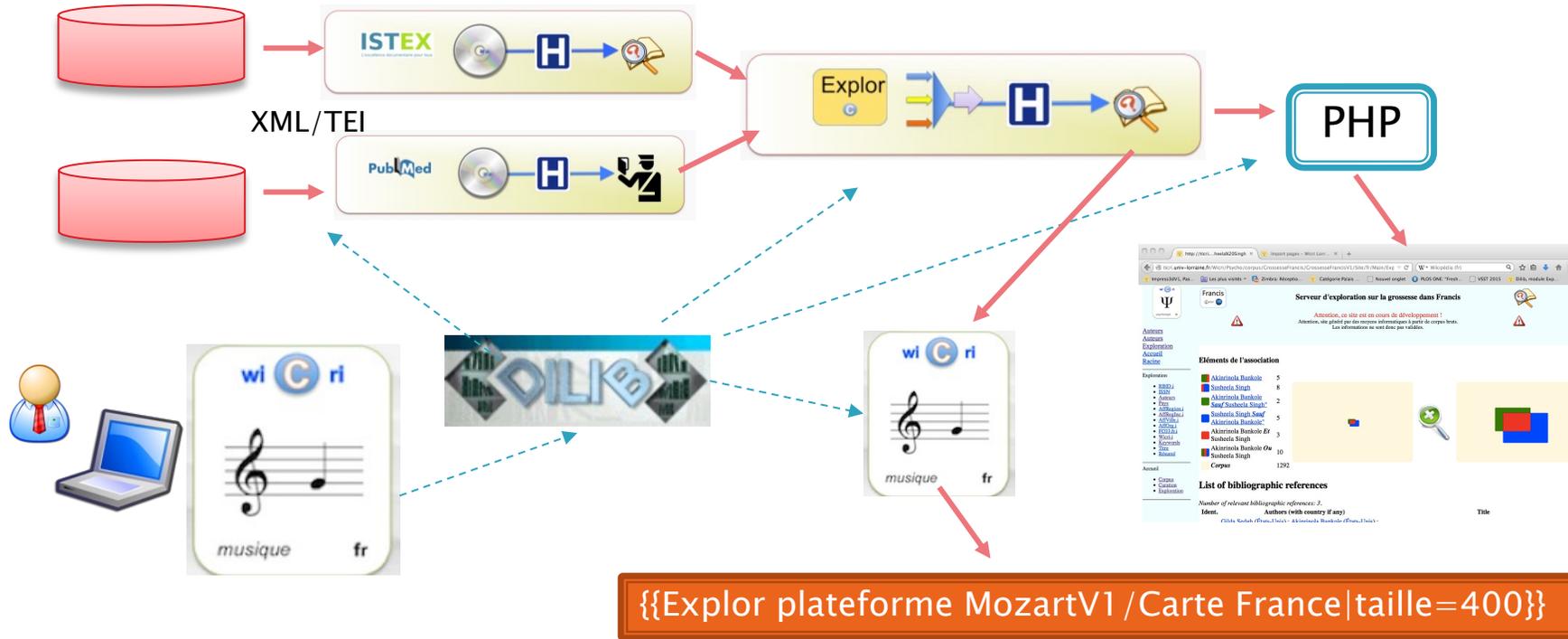
Tintin au Congo

Hergé (1948)

Tintin en Amérique

Hergé (1949)

ISTEX – Serveur d’exploration – génération



Pays	Région	Villes
1. France (67) ↗	1. Californie (11) ↗	1. Paris (9) ↗
2. États-Unis (31) ↗	2. Île-de-France (9) ↗	2. Marseille (5) ↗
3. Royaume-Uni (14) ↗	3. Occitanie (région administrative) (7) ↗	3. Montpellier (4) ↗
4. Allemagne (14) ↗	4. Massachusetts (4) ↗	4. Londres (4) ↗
5. Canada (11) ↗	5. Anglet (6) ↗	5. Grenoble (4) ↗
6. Italie (10) ↗	6. État de New York (5) ↗	6. Berlin (4) ↗
7. Espagne (8) ↗	7. Marlyland (5) ↗	7. Toulouse (3) ↗
8. Suisse (6) ↗	8. Caroline du Nord (5) ↗	8. Prague (3) ↗
9. Australie (6) ↗	9. Arizona (5) ↗	9. Montréal (3) ↗
10. Pays-Bas (5) ↗	10. Washington (Etat) (4) ↗	10. Zurich (2) ↗
Mots-clés anglais	Mots des titres	ISSN/revue
1. Astrophysics (3) ↗	1. data (10) ↗	1. SPIE proceedings series (6) ↗
2. State of the art (2) ↗	2. analysis (7) ↗	2. 1932-6203 (5) ↗
3. Software package (2) ↗	3. software (6) ↗	3. Lecture Notes in Computer Science (4) ↗
4. Real time (2) ↗	4. microbial (6) ↗	4. Eos Trans. AGU (3) ↗
5. Quebec (2) ↗	5. marine (5) ↗	5. 2024-6250 (3) ↗
6. Perspective (2) ↗	6. genome (5) ↗	6. 1091-6490 (3) ↗
7. Open source software (2) ↗	7. distributed (5) ↗	7. 0096-3941 (3) ↗
8. Measurement sensor (2) ↗	8. genomic (4) ↗	8. 0027-8424 (3) ↗
9. Library network (2) ↗	9. control (4) ↗	9. 2047-217X (2) ↗
10. Information policy (2) ↗	10. web (3) ↗	10. 1545-7885 (2) ↗



```

<idx>
  <k>Salieri</k>
  <f>2</f>
  <l>
    <e>001230</e>
    <e>002345</e>
  </l>
</idx>

```

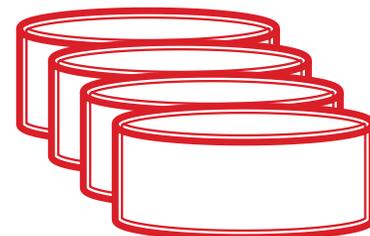
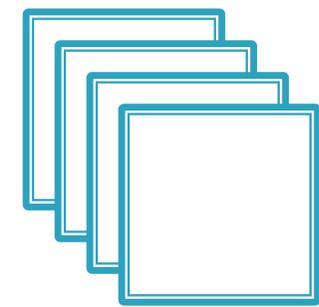
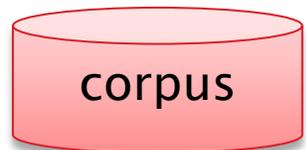
Santé : Serveurs à génération rapide (5')



NlmPubMedExplorCorpus -q influenza -s 1000



Part	Hyper	View
1. France (17) p.	1. Colombie (17) p.	1. Pays (17) p.
2. États-Unis (17) p.	2. Nouvelle-Zélande (17) p.	2. Nouvelle-Zélande (17) p.
3. Royaume-Uni (17) p.	3. Royaume-Uni (17) p.	3. Royaume-Uni (17) p.
4. Allemagne (17) p.	4. Royaume-Uni (17) p.	4. Royaume-Uni (17) p.
5. Canada (17) p.	5. Royaume-Uni (17) p.	5. Royaume-Uni (17) p.
6. Italie (17) p.	6. Royaume-Uni (17) p.	6. Royaume-Uni (17) p.
7. Espagne (17) p.	7. Royaume-Uni (17) p.	7. Royaume-Uni (17) p.
8. Suisse (17) p.	8. Royaume-Uni (17) p.	8. Royaume-Uni (17) p.
9. Australie (17) p.	9. Royaume-Uni (17) p.	9. Royaume-Uni (17) p.
10. Portugal (17) p.	10. Royaume-Uni (17) p.	10. Royaume-Uni (17) p.



Pages paramètres

Base Xml

Ecrire la musique



<score>

```
{  
  \time 2/4  
  \clef bass  
  c4 c g g a a g2  
}
```

Les commandes sont précédées d'un \

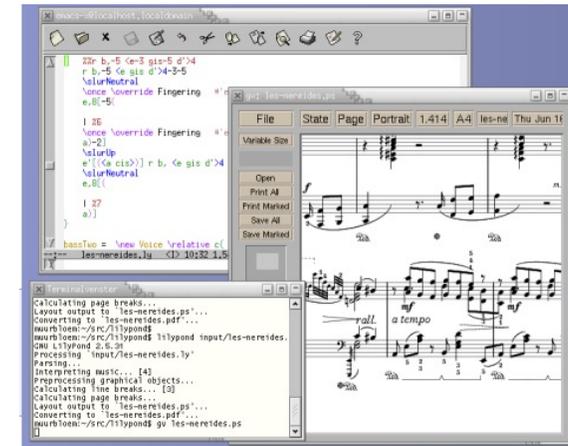
Les lettres représentent les notes

Les chiffres annoncent la durée

</score>



a = la
b = si
c = do
d = ré
...



Études pour choristes amateurs curieux..

Lassus, musica ficta, contrafacta

Impression pour musicien professionnel

Altérations implicites

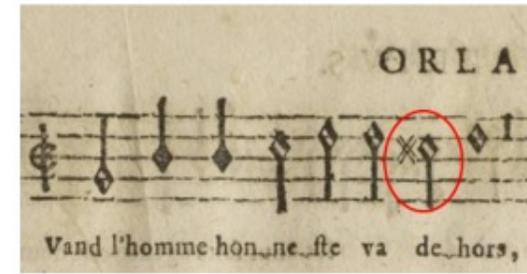


S V P E R I V S.

Vand mon mary vient de dehors, Ma



0:00 / 0:00



ORLA

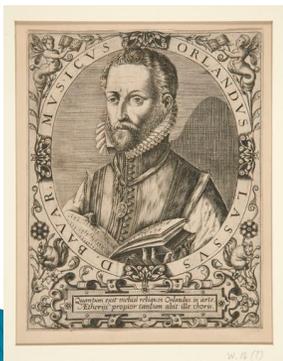
Vand l'homme hon.ne.ste va de.hors,



0:00 / 0:00

Pour amateurs

Altérations explicites



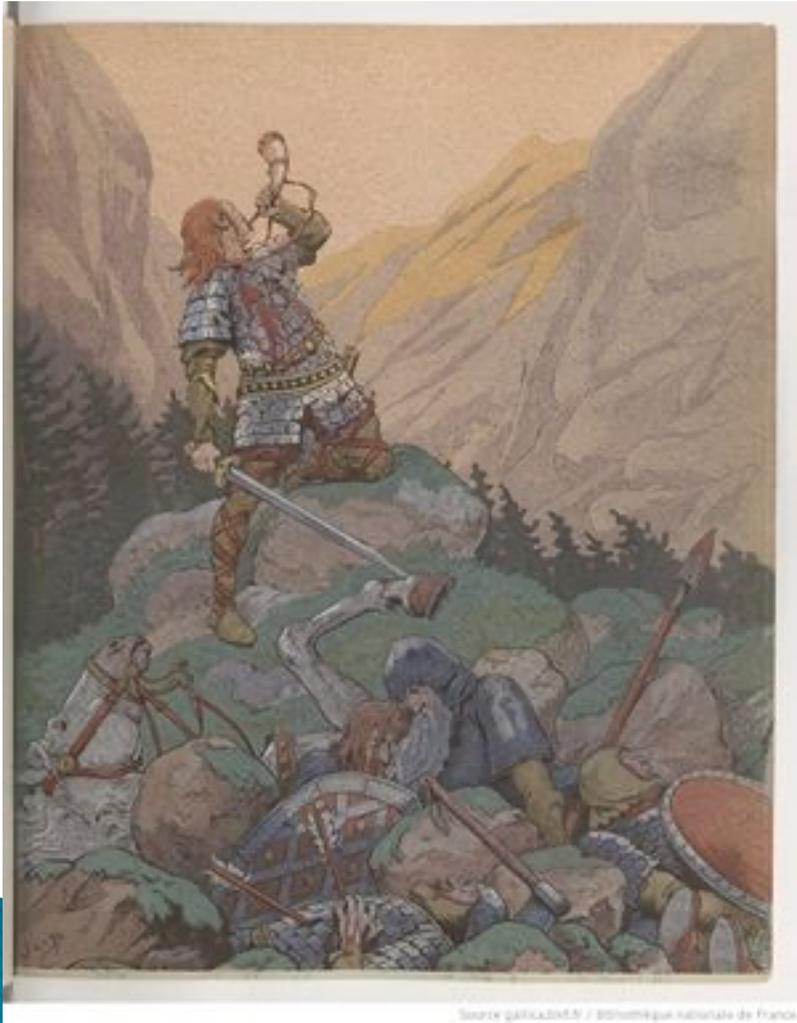
Quand mon mary vient de dehors,
Ma rente est d'estre battue :
Il prend la cuillier du pot
À la teste il me la rue.

J'ay grand peur qu'il ne me tue.
C'est un faux vilain, jaloux
C'est un vilain, rioteux, grommeleux.
Je suis jeune et il est vieux.

Quand l'homme honneste va dehors
Sa femme n'est par la rue
Ainsi à la besogne alors
Menagère elle se rue

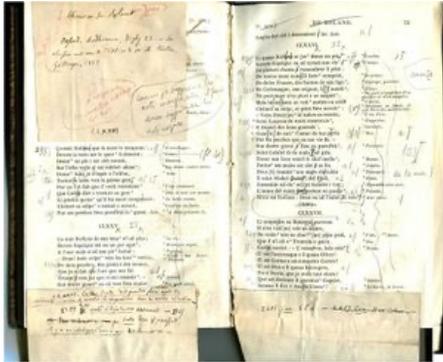
S'il revient elle le salue
Et lui fait accueil gracieux
Il n'est un vilain, rioteux, grommeleux.
Si elle jeune et s'il est vieux.

La Chanson de Roland, une histoire, une épopée

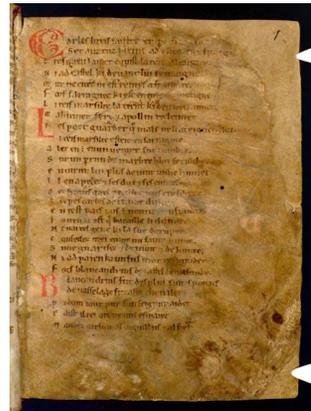


- ▶ Le 14 août 778, Charlemagne, de retour d'Espagne, perd l'arrière garde de son armée au col de Roncevaux
 - Ainsi est mort Roland
- ▶ La légende
 - la trahison de Ganelon
 - L'attitude héroïque de Roland
 - Qui attend de mourir pour sonner de l'olifant
 - Qui brise le rocher avec Durandal son épée
- ▶ La littérature
 - Manuscrits à partir du XIIe siècle
 - Transcriptions, traductions au XIX siècle
 - Des dizaines de milliers d'articles
 - De multiples valorisations (opéras, films...)

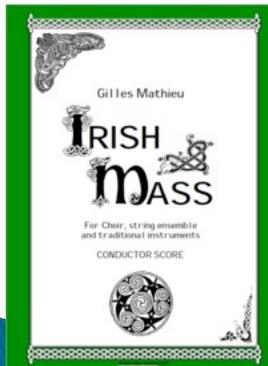
Un sujet de stage... 😊 imaginé tranquille !



Chanson de Roland
Francisque Michel
Annoté par Paul Meyer



Manuscrit d'Oxford,
accessible sur Wikipédia



Gilles Mathieu
Un oratorio sur la Chanson de Roland



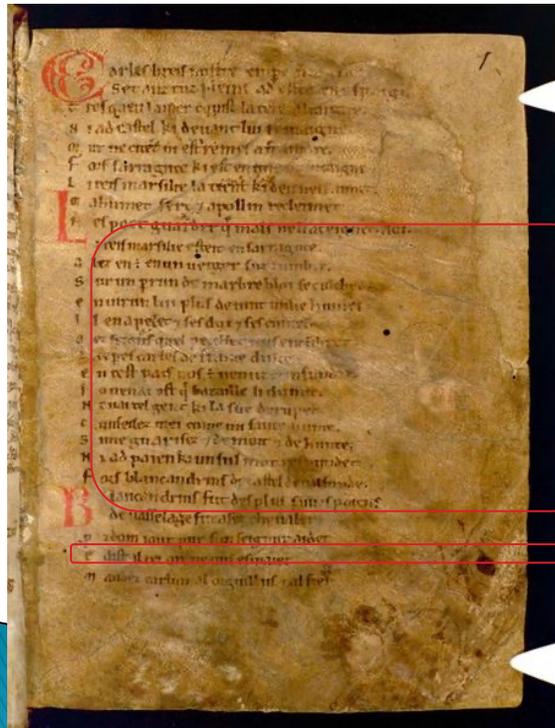
Un invité imprévu
Léon Gautier...



La chanson de Roland, des manuscrits

- ▶ Point de départ : inconnu, nombreuses hypothèses,

Manuscrit d'Oxford (XIIe siècle ?)



72 feuillets
Recto-verso

28?
Couplets
(Laises)

4202 vers

Manuscrit de Paris
XIIIe siècle, incomplet
6828 vers / 275 laisses



Châteauroux



Cambridge



Manuscrit de Venise
XIIIe siècle
6000 vers
419 couplets



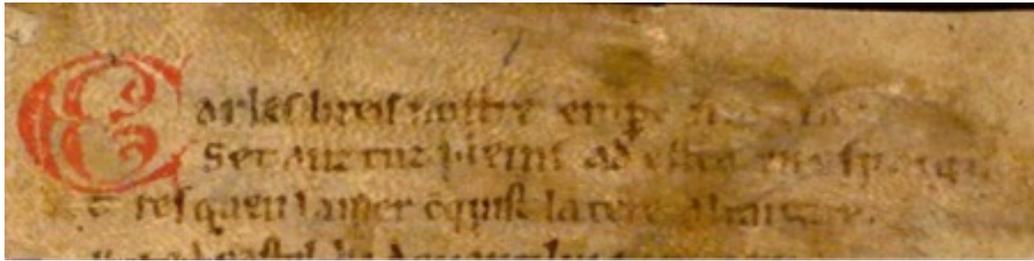
En allemand
Curé Konrad



La Chanson de Roland

Un oratorio de Gilles Mathieu

▶ Des vers en assonance



Carles li reis, nostre emperere magne*,
Set anz tuz pleins* ad ested en Espagne,
Tresqu'en* la mer cunquist la tere altaigne**;
N'i ad castel ki devant lui remaigne*,

▶ Une mise en musique

- En 10 mouvements
- Sur une sélection de 200 vers
- Pour chœurs SATB
- Cor, orchestre à cordes

♩ = 68

Soprano
f Ca - rles li - re - is nostr' em - pe - rer mag - nes Set anz tuz pleins ad e - stet en

Alto
f Ca - rles li - re - is nostre ma - gnes Set anz tuz pleins ad e - stet en

Tenor
f Car - les - li reis nostr' em - pe - rer mag - nes Set anz tuz pleins ad e - stet en

Bass
f Car - les li reis nostr' em - pe - rer mag - nes Set anz tuz pleins ad e - stet en

La Chanson de Roland

Des transcriptions, des traductions (centaines)

Francisque Michel (1837)
Première transcription
(Manuscrit d'Oxford)

LA CHANSON
DE ROLAND
de Roncevaux

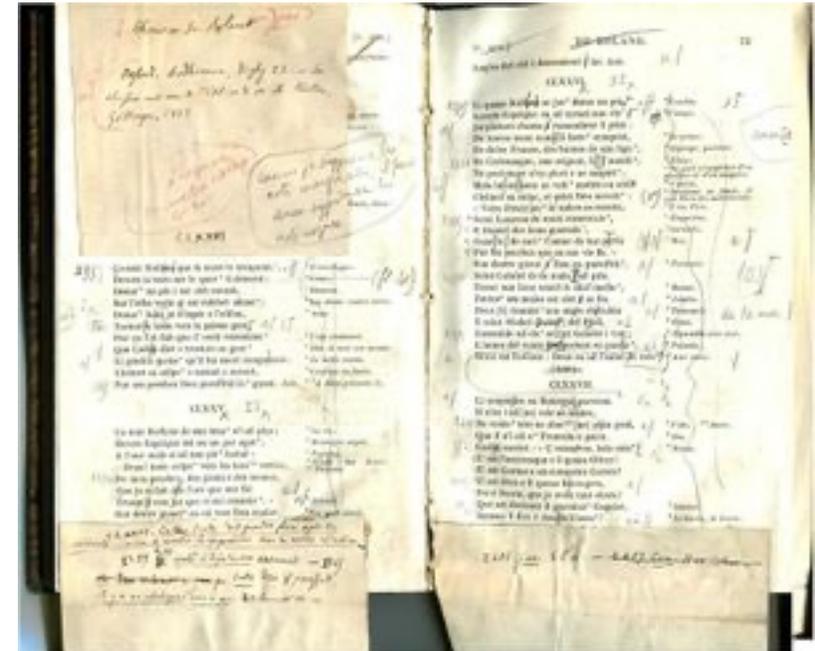
PARIS
CHEZ SOLVAYRE LIBRAIRE

1837

Léon Gautier
Edition populaire
1881



Francisque Michel (1869)
Annoté par Paul Meyer



42.000 articles sur Google Scholar...

La Chanson de Roland, problèmes et divergences

Dis blanches mult fist amener marfil
 Que li tramist li reis de suacie.
 Li freia se dozel seles d'argent misel.
 Cil sunt muntez ki le message firent.
 Enz en lur mains portent branches dolue.
 Vindrent charles ki france ad en barue.
 Li empere se fait y balz y liez.
 Li empere se fait y balz y liez.

e seint michel del pil.
 ensemble el seint gabriel iunt.
 L'anne del cunte portent en pareis.
 Quant se roit d' en ad l'anne el cels.
 Li empere en reuce nat parment.
 Il n'ent ad ne uere ne fenter.

(v. 2441.) DE ROLAND. 75
 De cels d'Espagne unt les doz tueriez,
 Tenent l'enchalz, tuit en sunt cumunel*.
 Quant veit li reis le vespres* decliner,
 Sur l'erbe verte si descent en un pred;

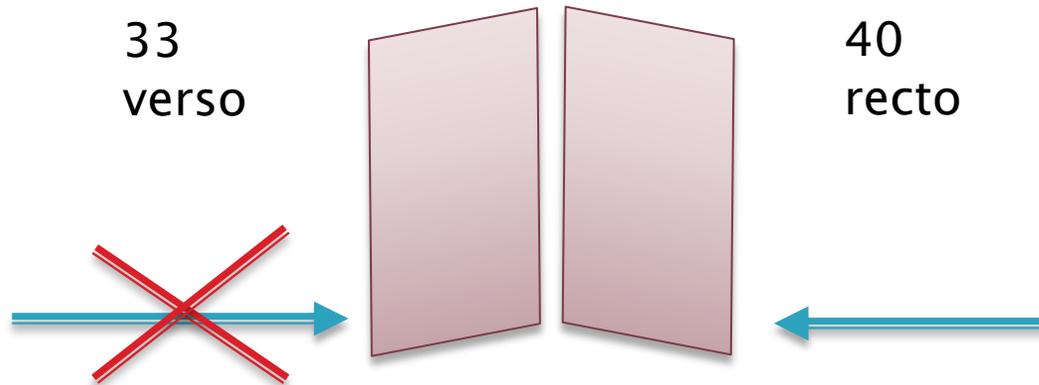
es clos
 * Tiennent la poursuite, tous y prennent part.
 * Soir.

Se cels despaigne unt lur les dos turnez.
 Tenent l'enchalz tuit en sunt cumunel.
 Quant veit li reis le vespres decliner.
 Sur l'erbe verte descent li reis en un pred.

L'accident du vers 2242

- ▶ Le scribe s'est trompé de page !!!
- ▶ Sur un vers facile à analyser

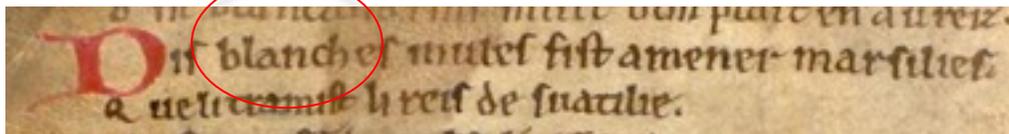
Mors est Turpins, li guerreiers Carlun.

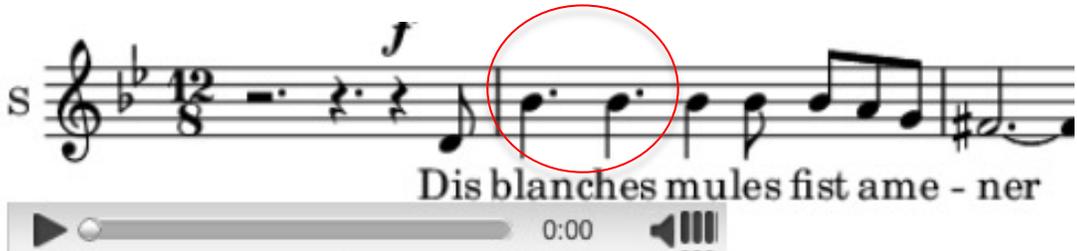


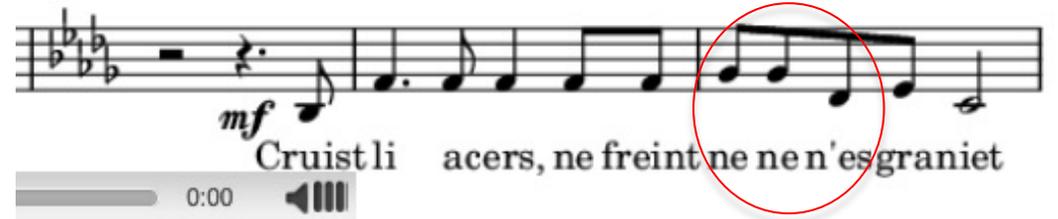
Blog dialogue avec le compositeur

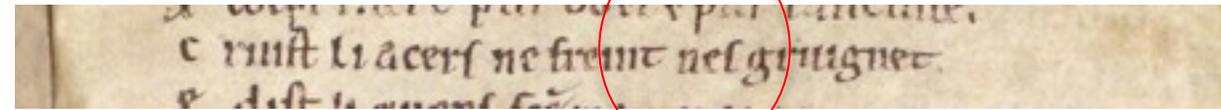


S  *f* Dis mu - les fist ame - ner



S  *f* Dis blanches mules fist ame - ner

 *mf* Cruist li acers, ne freint ne ne n'esgraniet



Leon Gautier :

Cruist li acers, ne freint ne ne s'esgruignet ;

Devant lui est une roche brune ;
Par grande douleur et colère,
il y assène dix forts coups ;
L'acier de Durendal grince :
point ne se rompt, ni ne s'ébrèche :

Dictionnaires hypertextualisés

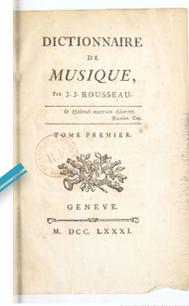
Brossard
1703



Andante



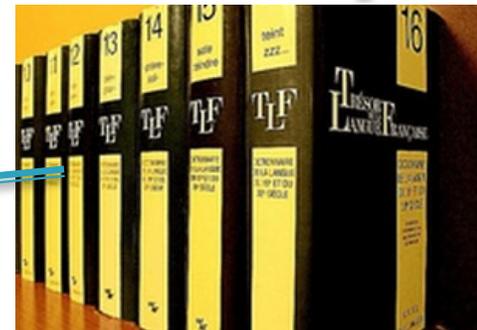
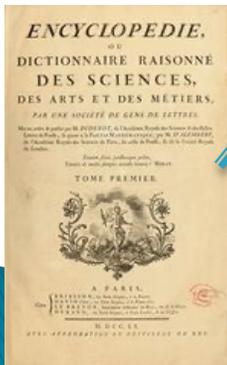
Jean-Jacques Rousseau
1871



Simple lien



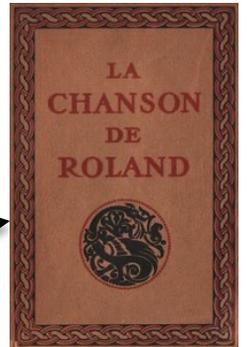
Encyclopédie
1851



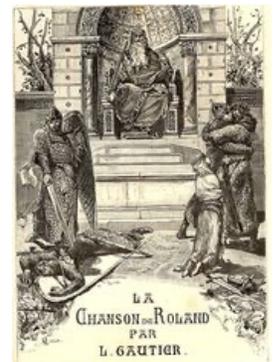
TLF 1970

Exemple TLF

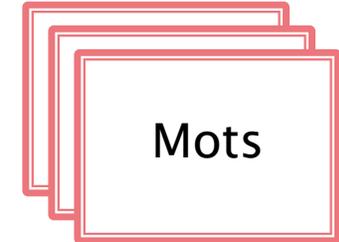
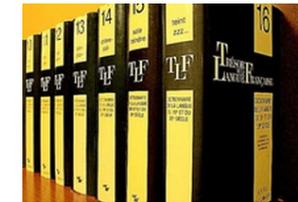
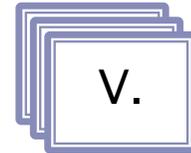
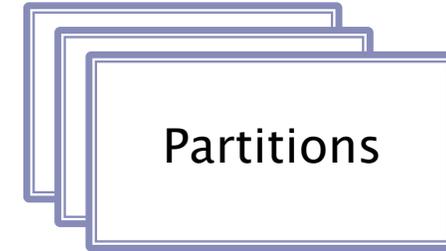
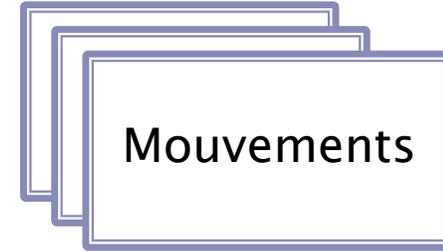
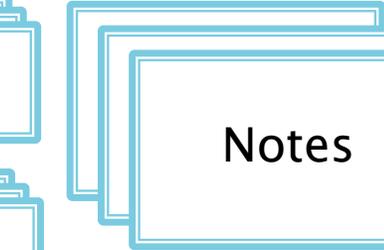
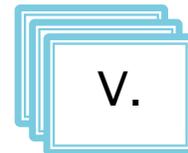
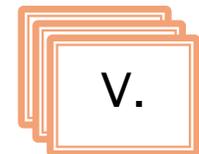
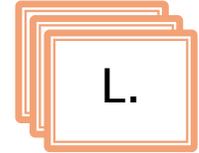
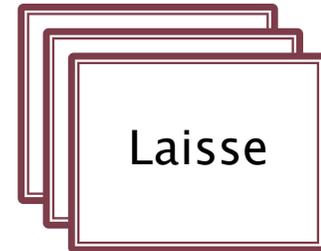
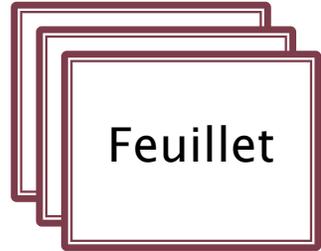
Lien sémantique



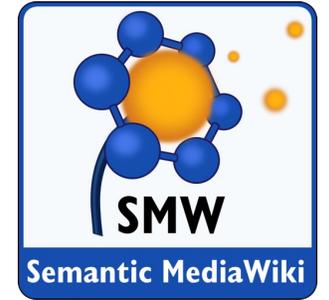
*Nouvelle
Attestation*



La Chanson de Roland un gigantesque hypertexte



Liens sémantiques Autour de Paul Meyer



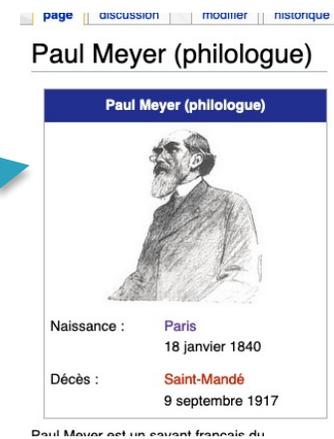
A pour
auteur



Est dans la revue



A pour
créateur



A pour
créateur



Catégories et liens sémantiques structurels

- ▶ Exemple : accès par mosaïque d'images

```
[[Catégorie:Chanson de Roland (Stengel) texte|154]]
```

- ▶ Gestion des actualités

Escarboucle (Trésor de la langue française)

```
;Notes de gestion:  
* signalement : [[A pour date de mise en lecture::8 juin 2022]]  
[[Catégorie:Articles du TLF]]
```

Actualités

```
===Articles du TLF===  
{{#ask:  
[[Category:Articles du TLF]]  
|?A pour date de mise en lecture  
|sort=A pour date de mise en lecture  
|order=descending  
|limit=3}}
```

Articles du TLF [\[modifier\]](#)

	◆ A pour date de mise en lecture ◆
Escarboucle (Trésor de la langue française)	8 juin 2022
Durandal (Trésor de la langue française)	8 mars 2022
Jongleur (Trésor de la langue française)	11 octobre 2021
... autres résultats	

Un wiki dans un réseau

Un réseau de réseaux de communautés



Réseau Wicri
50 familles multilingues
Contributeurs identifiés



Plusieurs communautés
Musicologues
Amateurs
Philologues
Historiens...



Opéra
Italien
Monteverdi

Charlemagne

Pour des acteurs de la science ouverte

- Ouvrages libres de droits
- Indépendance technologique (appropriation)
 - Sites de travail en évolution permanente

Bilan, quelques chiffres

- ▶ WICRI (10 ans) environ 150 wikis (dont les Mots de l'Agronomie)
 - 100.000 pages dont 30.000 significatives
 - 100 serveurs d'explorations (400.000 documents ISTEEX + PubMed + ...)
- ▶ Wicri/Chanson de Roland (1 an)
 - 6.358 pages, dont 2.229 significatives, 25.000 items sémantiques
 - 2 serveurs d'exploration : Roland (3.182), Paul Meyer (1.500)
 - Premiers ouvrages (squelette / contenu / présentation) :
 - Chanson de Roland par F. Michel, annotée Paul Meyer (100% / 100% / 50%)
 - Roman de Roncevaux par F. Michel, annotée Paul Meyer (100% / 5% / 5%)
 - Chanson de Roland, édition populaire de Léon Gautier 1895 (100% / 100% / 50%)
 - Manuscrit d'Oxford (100% / 30% / 20%)
 - Manuscrit de Paris (100% / 20% / 10%)
 - Oratorio de Gilles Mathieu (99% / 40% / 30%)
 - Manuscrit de Châteauroux (20% / 5% / 2%)

Liens et catégories « imaginaires »

- ▶ Exigence d'érudition pour la nomenclature
 - *Rolandslied* : traduction du terme ou manuscrit de Konrad ou ???
 - Ogier le Danois ou *Ogier von Dänemark* ou *Ogier de Danemarche*
 - un personnage imaginaire présent dans des dizaines de chansons de gestes
 - Un titre de poème, chanson de geste, pièce musicale
- ▶ Relations sémantiques à étudier
 - Charlemagne, Roland : personnalités historiques et personnage imaginaire
 - « a pour personnalité citée » / « a pour personnage cité »
 - Simple lien, ... « A pour personnage chantant »
- ▶ Implique :
 - Un site en évolution permanente (dans sa structure)
 - Cohabitation de zones grand public et... de brouillons

Généralisation de la démarche wiki+xml

Pour les humanistes et les sciences avec observation

- ▶ Protocole de type accompagnement
 - Implique une permanence distanciel + présentiel
- ▶ Au niveau d'un laboratoire
 - À la portée du LORIA
 - Mais difficile pour un laboratoire d'application
- ▶ Au niveau d'une université
 - 5 personnes (1 jour de permanence par semaine)
 - Dont 2 IR experts
 - Avec le modèle Wicri actuel
- ▶ Au niveau national, dans une dynamique internationale
 - Un réseau d'experts
 - Un modèle réseau à approfondir (100 sites 10.000 wikis)

Wicri réflexion alternative et orthogonale

- ▶ Galaxie Wikipédia,
 - Des services classiques
 - encyclopédie, dictionnaire, bibliothèque
 - Gigantisme pluri disciplinaire
 - Mécanismes de régulation souvent obscurs
- ▶ Illusions du Web sémantique
 - La combinaison des triplestores a-t-elle un sens ?
 - A quoi sert un triplestore généré par voie automatique à 60 de fiabilité ?
 - WikiData référentiel universel et éternel
 - Pour les musulmans, les chrétiens, les athées ?
 - Alimenté par les services de Poutine, Trump, le Pape François..



Wicri réflexion alternative et orthogonale

▶ Wicri

- Un réseau de sites
 - thématiques, géographiques ou institutionnels
- Où divers services sont implantés
- Comités scientifiques visibles
 - modération, orientation...



L'école d'Athènes
Socrate, Platon...

Merci

- ▶ Merci pour votre attention et vos questions