

Le projet « Guides de Paris » et la production d'un corpus textuel connecté

Emmanuel Château, Marianne Cojannot-Le Blanc

IFT6282, séminaire de Guy Lapalme, 9 novembre 2017



LABEX

Sommaire

- 1.** L'instrumentation d'un corpus textuel en histoire de l'art
- 2.** La création d'une API textuelle
- 3.** L'exposition du corpus sous forme de données liées

1. L'instrumentation d'un corpus textuel en histoire de l'art



LABEX

1.1 La constitution du corpus et ses enjeux



LABEX

Les Guides de Paris

Les historiens d'art & les corpus numériques



les
passés
dans le
présent

LABEX

HAAR

EA4414 HISTOIRE DES ARTS ET REPRÉSENTATIONS

{ BnF |
modyco

Université 
de Montréal

Le corpus

abondance

-- ~10 auteurs

-- du 16e au 18e siècles, 35 textes, + 300 éditions

variété

-- formats, publics

-- taille, illustrations

réemplois & réécritures

-- intertextualité

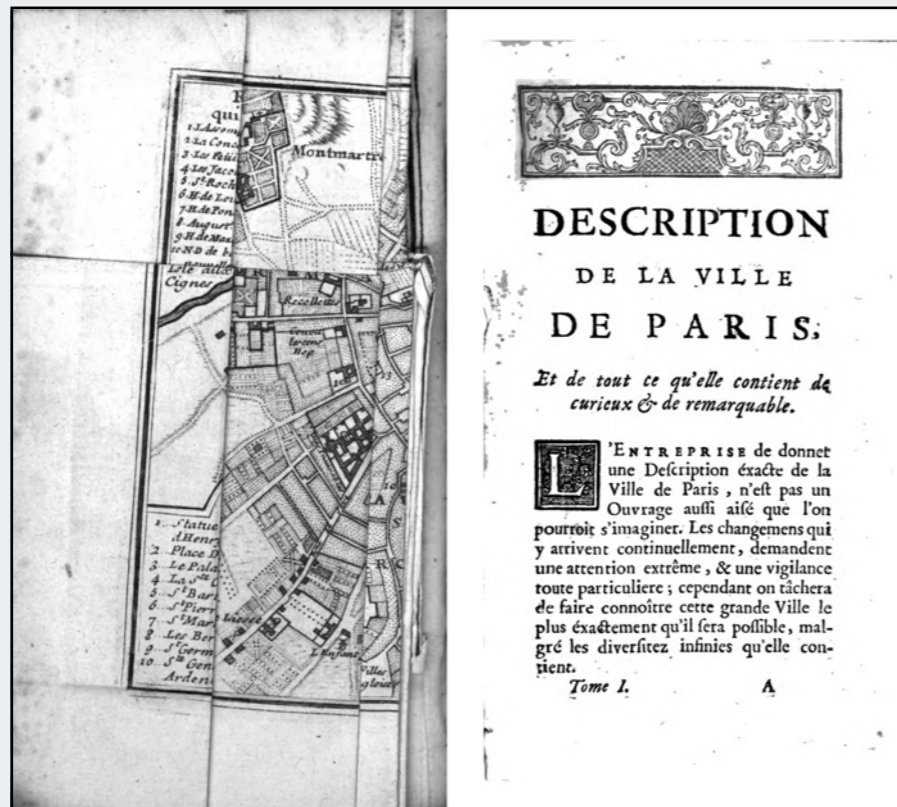
-- plagiat



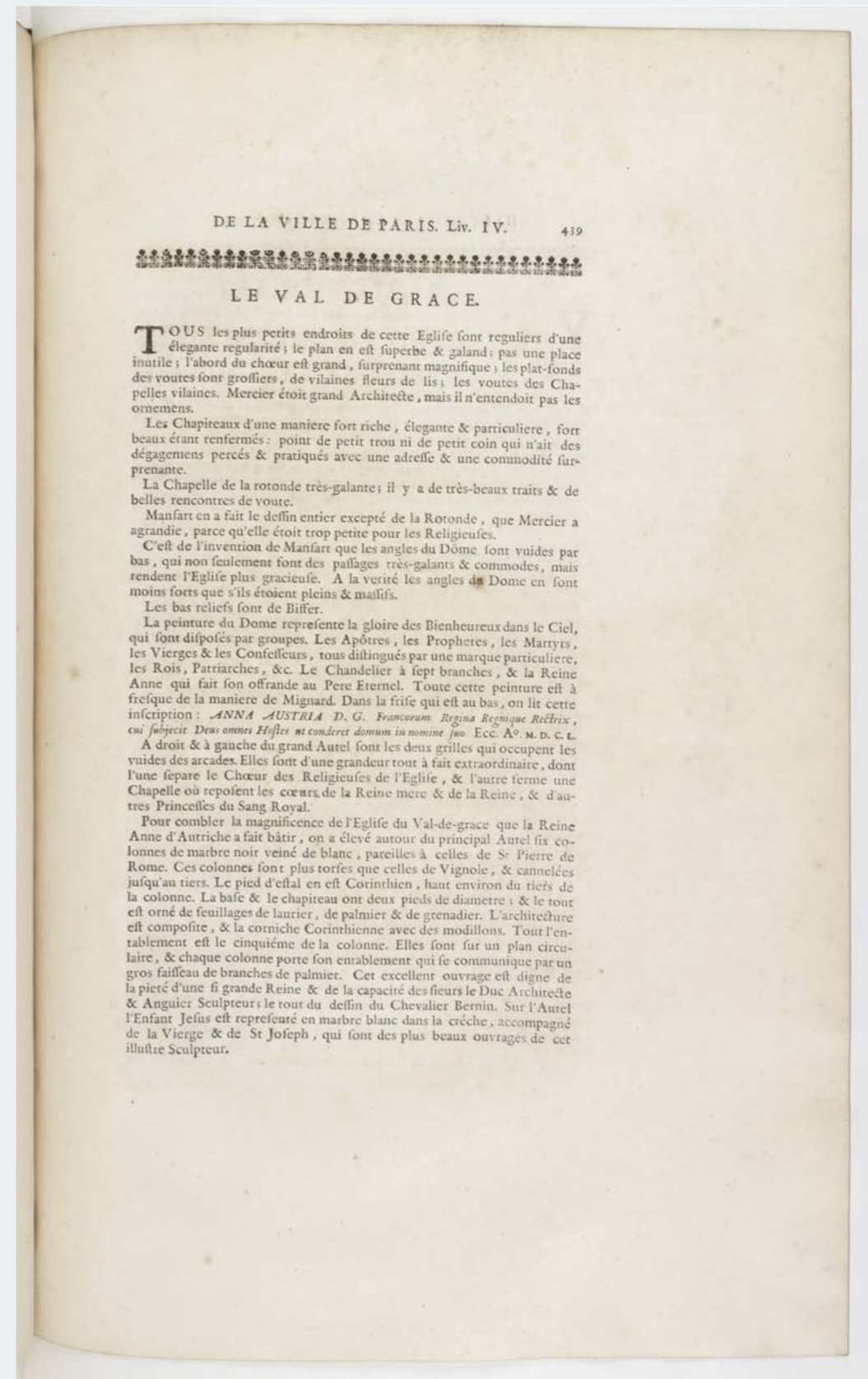
Le corpus cible

**35 titres (XVI^e-XVIII^e siècles)
= 300 éditions (Gilles Chabaud)**

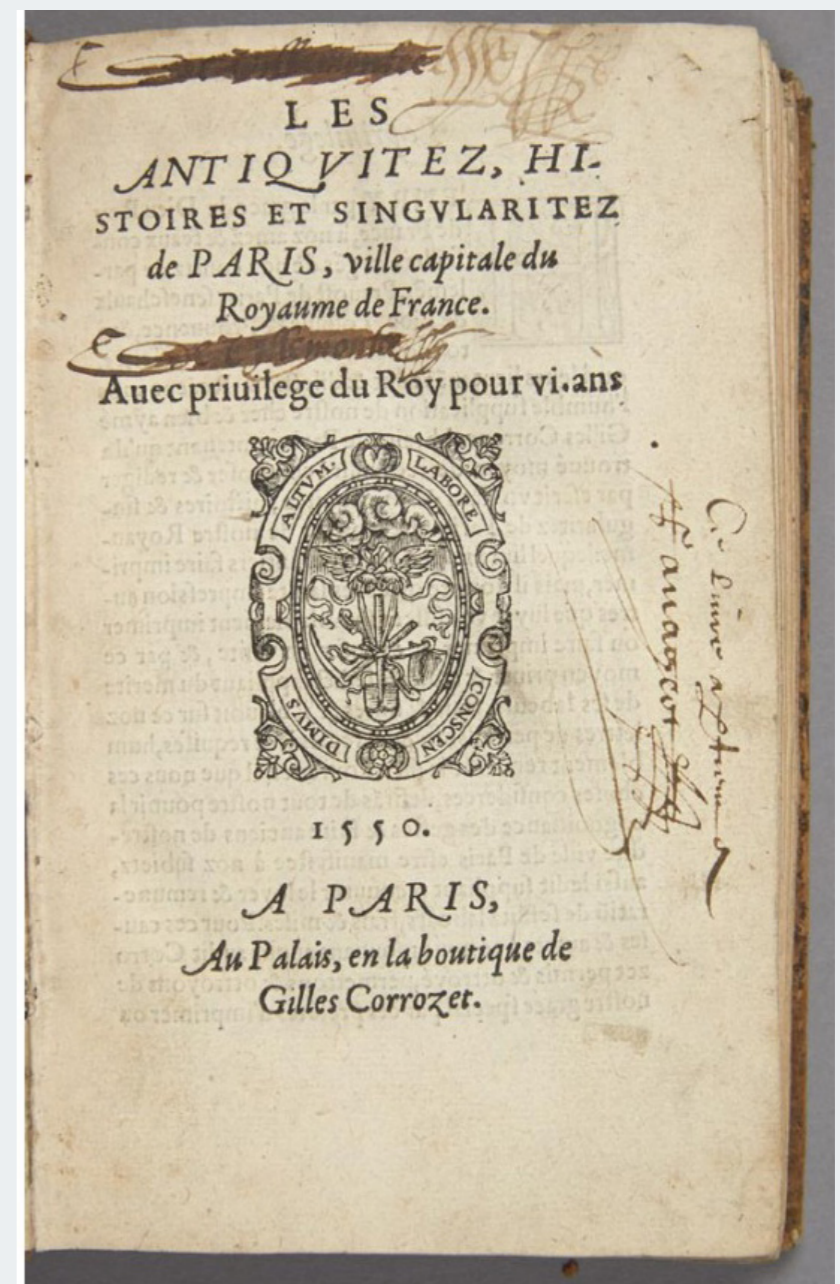
- Gilles Corrozet (1510-1568) - - 19 éditions 1531 - 1586
- Pierre Bonfons (15??-16??) - - 3 éditions 1605 - 1608
- Jacques Du Breul (1528-1614) - - 6 éditions 1608 - 1640
 - 1608, Paris, N. Bonfons
 - 1612, Paris, C. de La Tour
 - 1612, Paris, P. Chevalier
 - 1614, Paris, ex typographia Petipas
 - 1618, Paris, La Tour
 - 1639, Paris, Société des imprimeurs
 - 1640, Paris, P. Rocolet
- Claude Malingre (1580?-1653?) - - 1 édition 1640
- Henri Sauval (1620?-1676) - - 2 éditions 1724 - 1733
- Germain Brice (1653?-1727) - - 9 édition 1684 - 1752
- Charles Le Maire (16..-16..?) - - 2 editions 1685 - 1698
- etc.



Germain Brice, Description de la ville de Paris, 1752

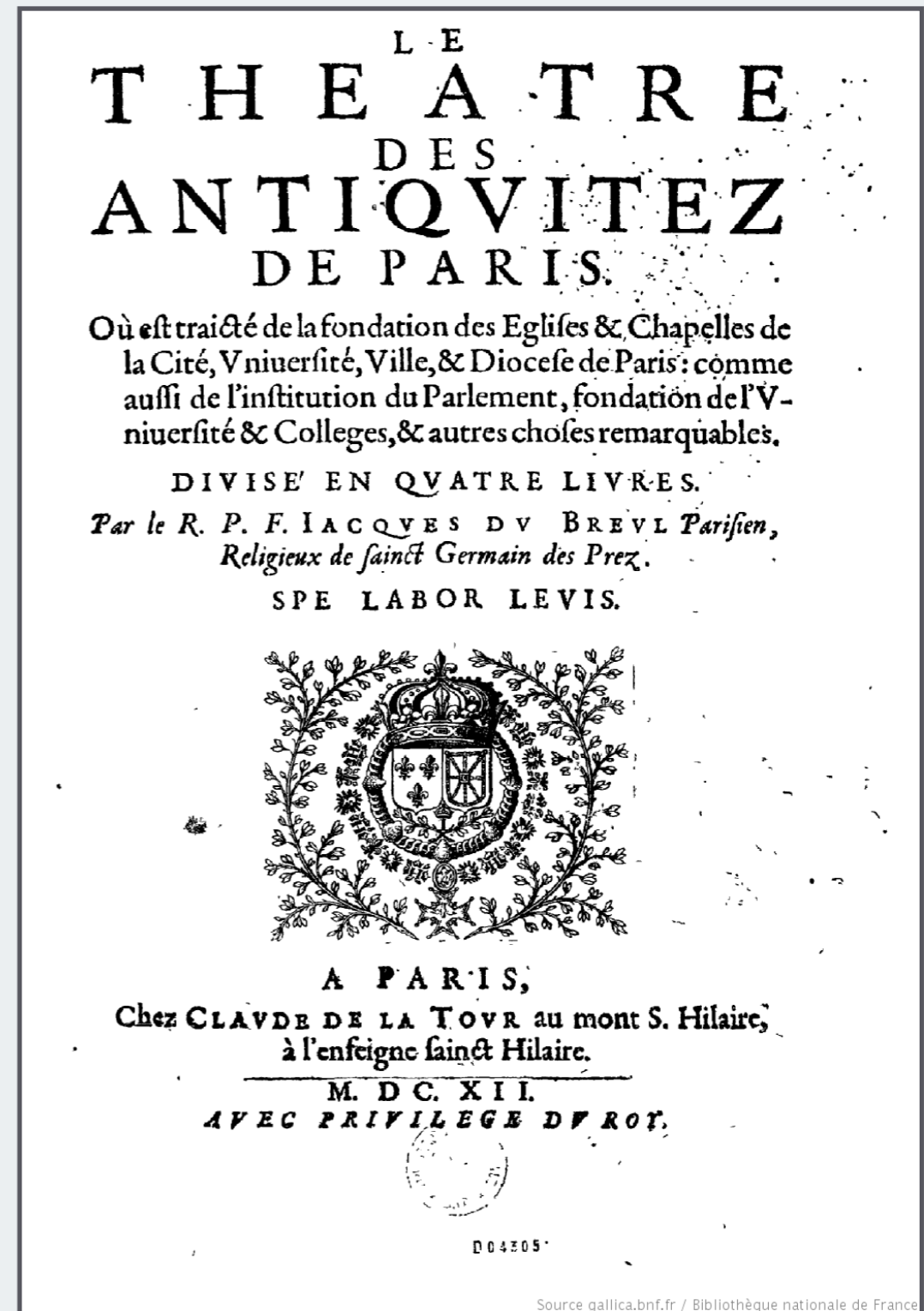
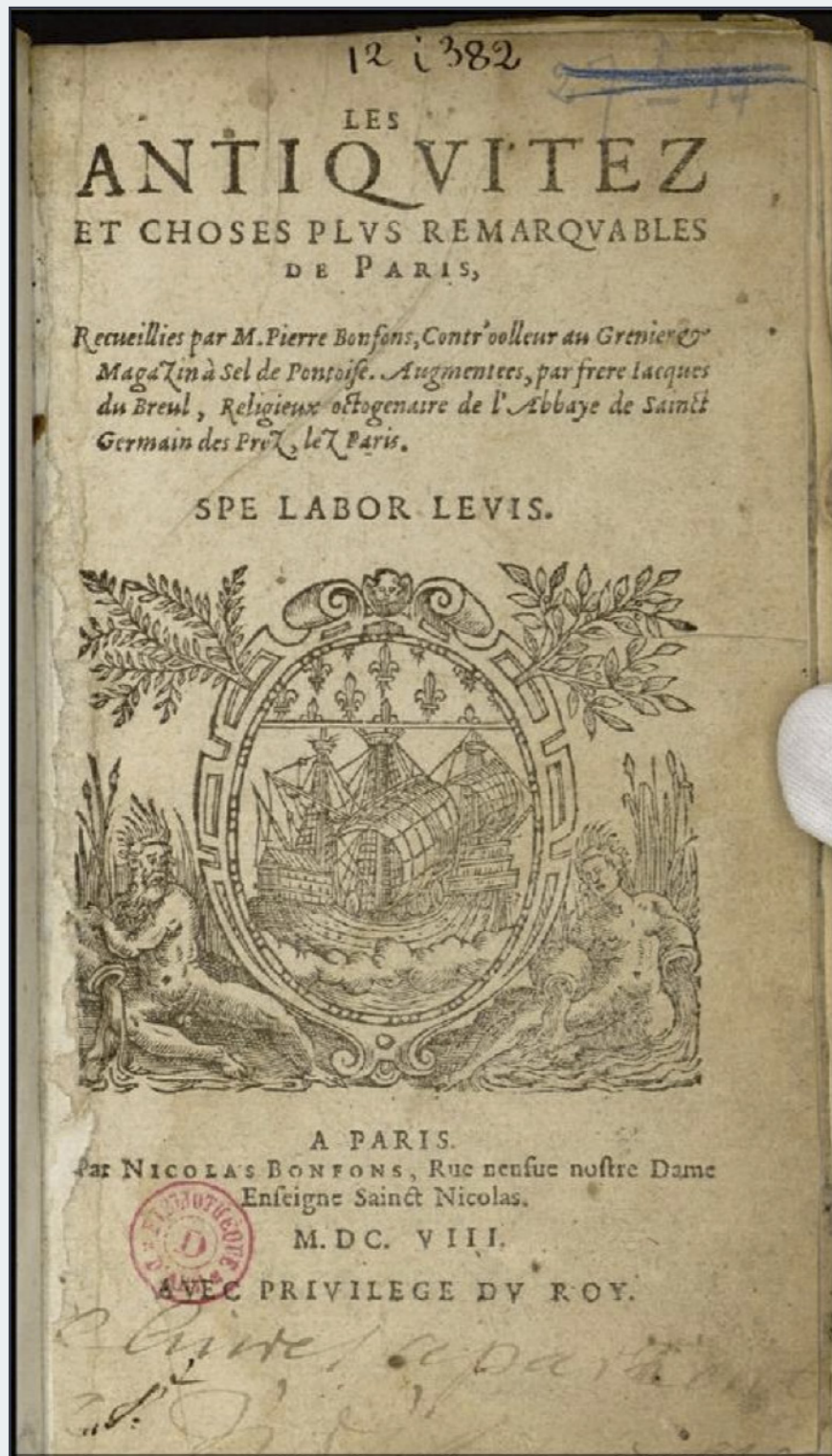


Henri Sauval, Histoire et recherches des antiquités de la ville de Paris, 1724



Gilles Corrozet, *Les Antiquitez...*, 1550.

<http://www.ericgrangeon.com/home/57-corrozet-gilles-.html>



Pierre Bonfons, *Les Antiquitez...*, 1608 Jacques Du Breul, *Le Theatre des antiquitez...*, 1612

LE
THEATRE
DES
ANTIQUITEZ
DE PARIS.

Où est traité de la fondation des Eglises & Chapelles de la Cité, Vniuersité, Ville, & Diocèse de Paris: comme aussi de l'institution du Parlement, fondation de l'Vniuersité & Colleges, & autres choses remarquables.

DIVISE' EN QUATRE LIVRES:

Par le R. P. F. IACQUES DV BREVL Parisien,
Religieux de saint Germain des Prez.

SPE LABOR LEVIS. *Richard.*



A PARIS,

Chez PIERRE CHEVALIER au mont saint Hilaire,
à la Court d'Albret.

M. D. C. XII.

AVEC PRIVILEGE DV ROY.



*tombeau du roi chilperic
tiré de la page 174 de
l'ouvrage de l'abbé de Breuil*

*tombeau de fredegonde
tiré de la page 175.*

LIVRE SECOND.

sternée de cœd & de corps fit son oraison de tant efficace, & si agreable à Dieu, que trois pauvres de long-temps aueugles furent illuminez: *Ad istud miraculum currit Regina, currit & populus: mirantur omnes fidem mulieris, mirantur gloriam confessoris: sed super omnia collaudatur Deus noster, qui tantam virtutem praeiatis suis, ut per eos salus operari dignetur.*

En la mesme Eglise du costé de Septentrion on voit vne autre tombe de pierre, sur laquelle vne statue de noy est couchée avec ces mots grauez à l'entour en lettres anciennes.

Rex Chilperici hoc tegitur lapide.

Belle-forest tome 1. des Annales, liure 1. chap. 23. suivant l'erreur de Nicolas Gilles, escriit qu'en memoire du forfait il tient la main en sa gorge, comme signifiant qu'il mourut de mort violente. Mais il s'abuse. C'estoit son geste ordinaire de tenir la barbe d'une main: Et l'ay veu ainsi insculpé en son grand seel, que me monstra (il y a près de soixante ans) Philippe de Lautier, general des monnoyes, & grand Antiquaire.

Au dessus de son sepulcro vn grand tableau est attaché contre la ceinture du chœur, dans lequel le mesme Roy est encor effigé en plate peinture, avec cet epitaphe sous les pieds.

Chilperic quatriesme Roy des François fils de Clotaire premier & neveu de Childébert, fondateur de ce mesme Monastere, & pere de Clotaire second de ce nom ensepulturé à l'opposite de ce lieu, lequel il engendra de Fredegonde & regna 22. ans, decedant par homicide l'an 587.

De ses vices, cruautez & miserable fin, voyez ce qu'en escriit son contemporain Gregoire de Tours, liure 6. chap. 46. de l'histoire de France: lesquels Gerard poëte comprend en trois carmes,

*Campanos vastavit agros, crudelis & atrox
Vxor em primam ingulat, Fredegunda secunda
Quem ingulat: Mors hac crudeli digna marito.*

Exemple memorable pour les Roys, ne expellice coniugem faciant.

Plus haut & proche du grand Autel on voit encores vne autre tombe, sur laquelle est taillee, par art de marquetterie, l'effigie d'une Roynie, vestue à l'antique où est cet escript, depeinte en vn tableau pendu contre le mur, auquel la mesme Roynie est encore representee.

Cy gist Fredegonde Roynie de France, femme de Chilperic, qui deceda à Paris l'an 601.

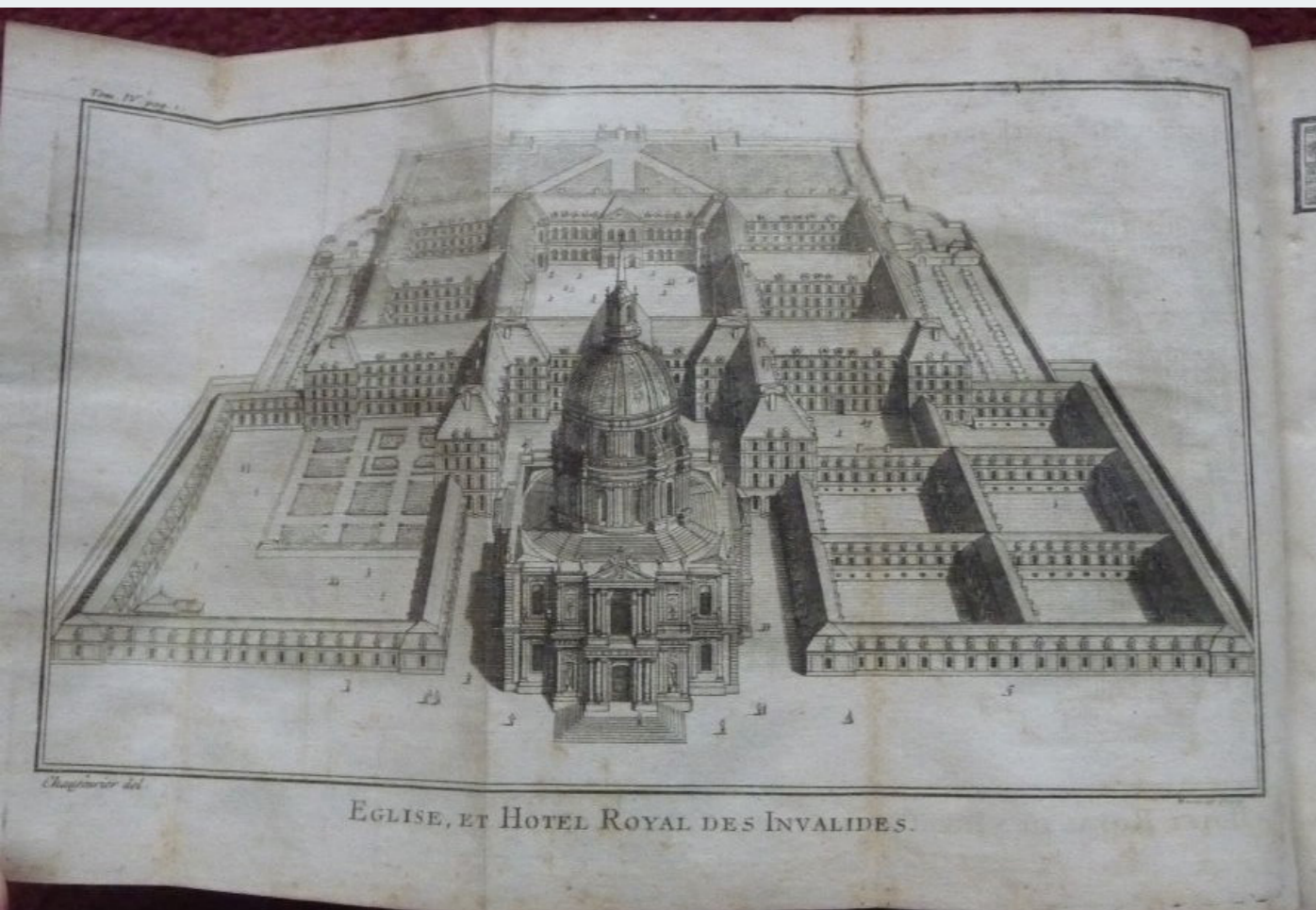
De l'autre costé du chœur vers le Mdy, l'on voit vne autre sepulture de pierre, au dessus de laquelle est peinte en vn tableau la statue d'un Roy avec cest epitaphe.

Cy gist Clotaire second de ce nom, Roy de France, fils de Chilperic & de Fredegonde, ensepulturé à l'opposite d'icy, & pere de Dagobert fondateur de l'Abbaye saint Denys: lequel Clotaire n'auoit que quatre mois d'age, quand son pere fut tué, & mourut ayant regné quarante quatre ans, l'an six cents trente.

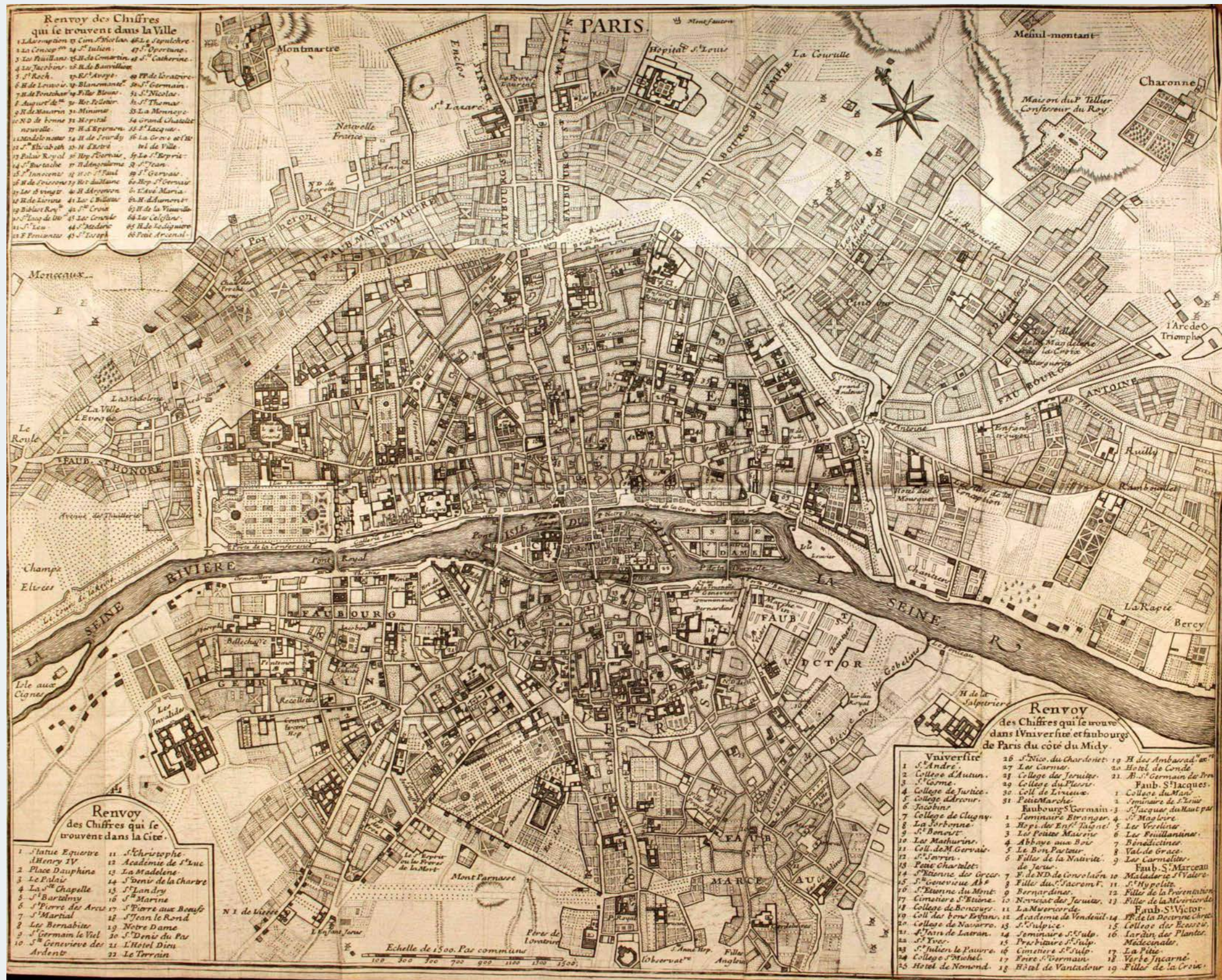
Vn peu plus haut en approchant des marches du grand Autel de la mesme Eglise maintenant dicté de saint Germain des Prez, l'on voit

FF

Jacques Du Breuil, *Le Theatre des antiquitez*,..., 1612



Germain Brice, *Description de la ville de Paris*,... 1752.



Germain Brice, *Description de la ville de Paris*,... 1752.

La guide des chemins de France.



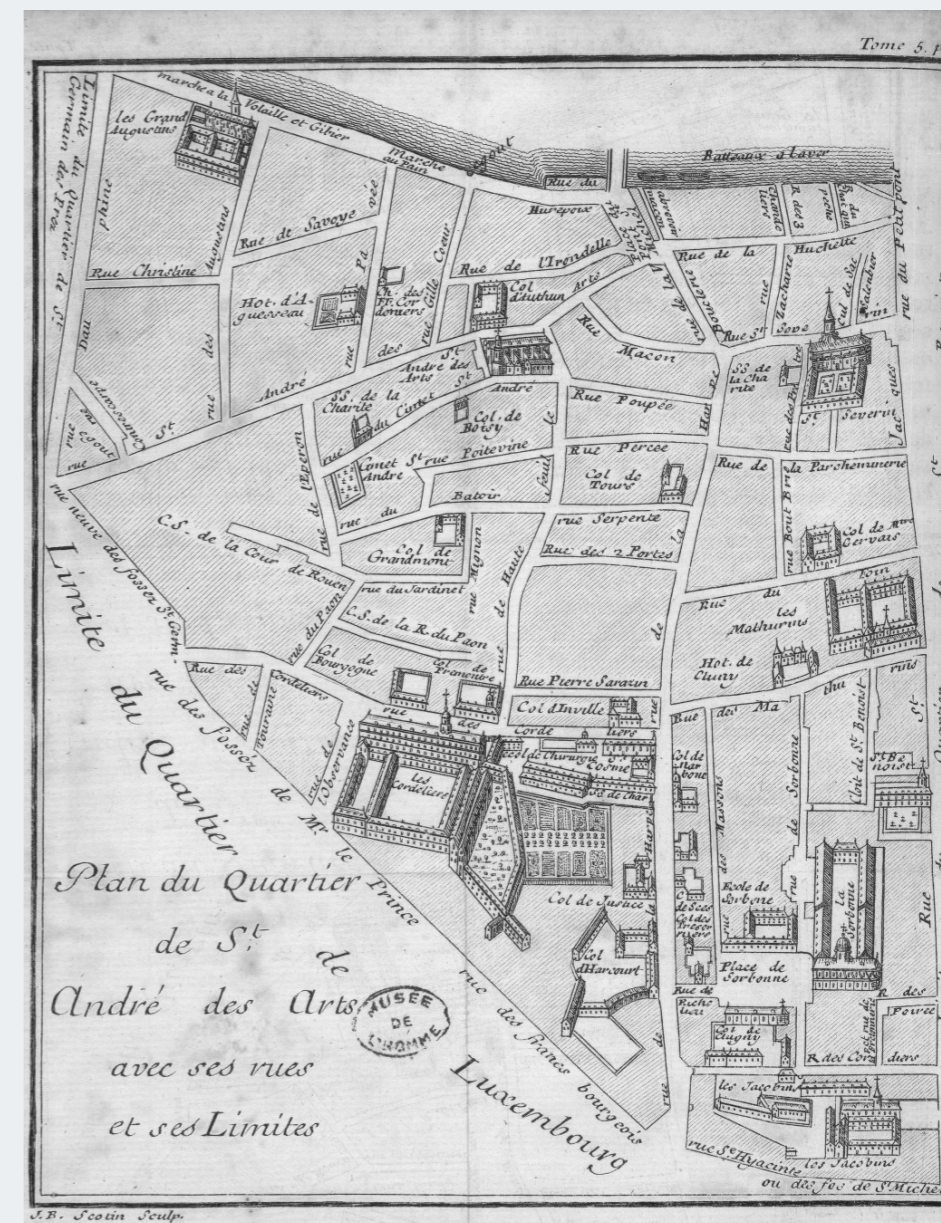
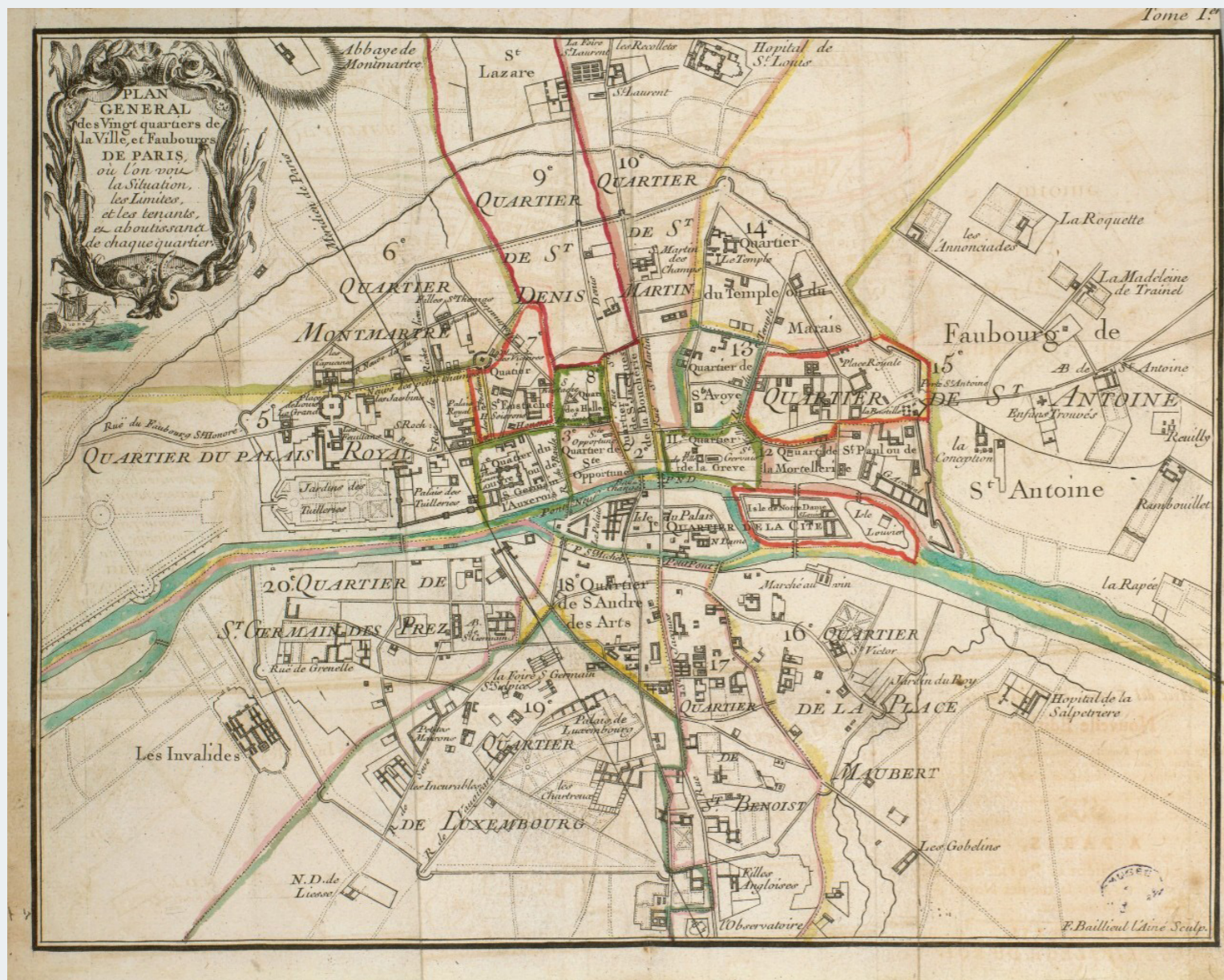
A PARIS,
Chez Charles Estienne, Imprimeur du Roy.
M. D. LII.
Avec priuilege dudit Seigneur

2005

TABLE DES PAYS DESCRIPTS en ce liure, selon l'ordre qu'ils y sont proposez.

Preuosté & viconté de Paris	4
Le Valois	9
La Picardie	16,17,18
Chemins de la haulte Picardie	31
Le Heurepoix	36
La Brye	39,40
La conté de Champagne, & pays adiacents	50,51
Lorraine, & pays adiacents	69,70,71,72
Chemins de la duché de Barrois	77
La Bourgogne	79,80,81,82
Le pays de Gastinois	91
La basse Beausse, & duché d'Orleans, & pays de Solongne	96,97
La Beausse Vendosmoise, dite moyenne	101,102
La haulte Beausse, & pays Chartrain	107
La conté du Perche & du Maine, & la duché d'Aniou	109,110
Normandie	118,119,120
La duché de Bretagne	124,125,126
La duché de Berry	135,136
La duché de Nyvernois	141

Estienne, Charles (1504?-1564). *La guide des chemins de France*. Paris : C. Estienne, 1552.



Piganiol de la Forge, *Description de Paris*,... 1742.

[Volume I. Plan dépl. en reg. du Titre : plan de Paris.] Plan général des vingt quartiers de la ville et faubourgs de Paris (...).

[Volume V. Plan dépl. en reg. p.419 : Paris.] Plan et escription du Quartier de Saint-André-des-Arts (...).

La nature du corpus

pas de cohérence interne

- genre non constitué
- usage actuel des chercheurs

difficulté de maniement

- volumes matériels, nombre des éditions, etc.
- absence de numérisation (avant le projet)

pas d'existence en tant que tel

- textes envisagés isolément
- phénomènes d'intertextualité, polémiques
- sollicitations ponctuelles

Des questions préalables

l'émergence des notions d'œuvre d'art et d'artiste

- que désigne-t-on, pourquoi
- naissance de la critique

l'émergence d'une conscience patrimoniale

- production de la valeur
- étude des entrées et des sorties

1.2 Un corpus outillé



LABEX

Systematiquement...

donner le texte à lire

- lecture linéaire
- base de données bibliographique

fournir des points d'accès multipliés

- index (noms, lieux, œuvres, ...)
- recherches plein-texte

autoriser des lectures transversales

- collations, inter-textualité, réemplois, etc.
- entrées et sorties
- géolocalisations

Un corpus instrumenté

permettre une manipulation du matériau textuel et l'outiller au moyen de dispositifs pour favoriser son étude et son exploration

¶ visualisations

- - géolocalisations
- - data-visualisations
- - mashups

¶ explorer et annoter

- - ontologie patrimoniale sur Paris
- - textométrie
- - annotation collaborative

Les « scholarly primitives »

“basic functions common to scholarly activity across disciplines, over time, and independent of theoretical orientation.”

john unsworth,

dean and professor of graduate school library and information science, u.illinois

- - **Discovering**

- - **Comparing**

- - **Sampling**

- - **Representing**

- - **Annotating**

- - **Referring**

- - **Illustrating**

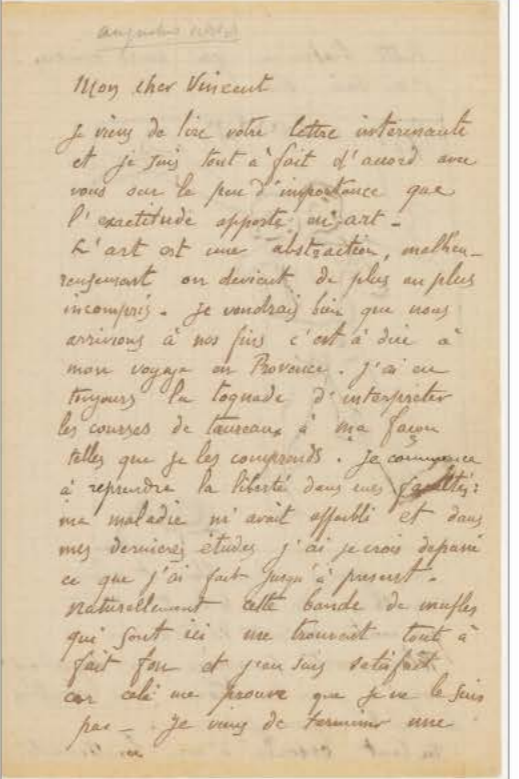
646 Paul Gauguin to Vincent van Gogh. Pont-Aven, on or about Sunday, 22 July 1888. SEARCH THIS LETTER PRINT

original text + line endings facsimile translation notes artworks

ira
My dear Vincent
I've just read your interesting letter and I entirely agree with you on the slight importance that accuracy contributes to art. Art is an abstraction; unfortunately we're becoming increasingly misunderstood. I would very much like it if we were to achieve our aims, that is, my coming to Provence. I've always had an itch to interpret bullfights in my own way, as I understand them. I'm beginning to recover the full use of my faculties: my illness had weakened me, and in my most recent studies I have, I think, gone beyond what I've been doing up to now. Of course, the collection of boors around here think I'm completely mad and I'm happy about that, because it proves to me that I'm not. I've just finished a Breton wrestling that I'm sure you'll like.

646
Br. 1990: 654 | CL: GAC 30
From: Paul Gauguin
To: Vincent van Gogh
Date: Pont-Aven, on or about Sunday, 22 July 1888
more...

original text + line endings facsimile translation notes artworks



original text + line endings facsimile translation notes artworks



(Eugène Henri) Paul Gauguin - The young wrestlers

- Charles Angrand
- Albert Aurier
- Emile Bernard
- Eugène Boch
- Hendrik Bongjer, Hermine Bongjer-Weismann and Anna van Gogh-Carbentus
- Egbert Borchers
- Antoine Philippe Furnée
- Hendrik Jan Furnée
- Paul Gachet Sr
- Paul Gauguin**
- Joseph Ginoux
- Joseph Ginoux and Marie Ginoux-Julien
- Marie Ginoux-Julien
- Anna van Gogh
- Anna van Gogh-Carbentus and Willemien van Gogh
- Cornelis Marinus van Gogh
- Theo van Gogh
- Willemien van Gogh
- Vincent van Gogh and Andries Bongjer
- Jo van Gogh-Bongjer
- Theo van Gogh and Jo van Gogh-Bongjer
- Anna van Gogh-Carbentus
- Theodorus van Gogh and Anna van Gogh-Carbentus
- Anna van Gogh-Carbentus and Theo van Gogh
- Johannes van Homborgh
- Joseph Isaacson
- Anton Kerssemakers
- Arnold Koning
- Henriette Lamblot-Gardin
- Horace Mann Livens
- Octave Maus
- Paul-Eugène Milliet
- Christian Mourier-Petersen
- Théophile Peyron

The letters by correspondent Paul Gauguin

- 576 Paul Gauguin to Vincent van Gogh. Paris, December 1887.
- 581 Paul Gauguin to Vincent van Gogh. Pont-Aven, on or about Wednesday, 29 February 1888.
- 586 Paul Gauguin to Vincent van Gogh. Pont-Aven, between about Saturday, 17 and about Monday, 19 March 1888.
- 646 Paul Gauguin to Vincent van Gogh. Pont-Aven, on or about Sunday, 22 July 1888.
- 675 Paul Gauguin to Vincent van Gogh. Pont-Aven, on or about Saturday, 8 September 1888.
- 688 Paul Gauguin to Vincent van Gogh. Pont-Aven, on or about Wednesday, 26 September 1888.
- 692 Paul Gauguin to Vincent van Gogh. Pont-Aven, Monday, 1 October 1888.
- 695 To Paul Gauguin. Arles, Wednesday, 3 October 1888.
- 706 To Paul Gauguin. Arles, Wednesday, 17 October 1888.
- 730 To Paul Gauguin. Arles, Friday, 4 January 1889.
- 734 Paul Gauguin to Vincent van Gogh. Paris, between Tuesday, 8 and Wednesday, 16 January 1889.
- 737 Paul Gauguin to Vincent van Gogh. Paris, Thursday, 17 January 1889.
- 739 To Paul Gauguin. Arles, Monday, 21 January 1889.
- 817 Paul Gauguin to Vincent van Gogh. Le Pouldu, between about Sunday, 10 and Wednesday, 13 November 1889.
- 828 Paul Gauguin to Vincent van Gogh. Le Pouldu, on or about Friday, 13 December 1889.
- 840 Paul Gauguin to Vincent van Gogh. Le Pouldu, on or about Friday, 17 January 1890.
- 844 Paul Gauguin to Vincent van Gogh. Le Pouldu, Wednesday, 22 or Thursday, 23 January 1890.
- 859 Paul Gauguin to Vincent van Gogh. Paris, on or about Thursday, 20 March 1890.
- 884 Paul Gauguin to Vincent van Gogh. Paris, on or about Friday, 13 June 1890.
- 892 Paul Gauguin to Vincent van Gogh. Le Pouldu, on or about Saturday, 28 June 1890.



http://vangoghletters.org, 2009

English version

Sandrart.NET

Startseite
Die Teutsche Academie im Volltext
Gliederung der Teutschen Academie
Übersetzungen

Suche
Personen
Orte
Kunstwerke
Bibliographie

Editionsinformationen
Projektinformationen
Herausgeber
Kontakt/Impressum

Volltextsuche Gesamtsuche Hilfe

Suche in Text, Personen, Kunstwerken, ...

Editions-Seite **9**

Darstellungsoptionen

Permanente URL:
http://ta.sandrart.net/-text-9

TA 1675, Geleitwort (Leopold Wilhelm Graf von Königsegg) [I]

Spaltenübergreifend

WIR **Leopold von Gottes Gnaden Erwählter Römischer Käyser** zu allen Zeiten Mehrer des Reichs/ in **Germanien/ zu Hungarn/ Böheim/ Dalmatien/ Croatien und Sclavonien etc. König/ Erz Herzog zu Oesterreich/ Herzog zu Burgund/ Steyr/ Kärnten/ Cräin und Würtemberg/ Graf zu Tyrol/** etc. bekennen öffentlich mit diesem Brief/ und thun kundt allemänniglich/ daß Uns Unser und des Reichs lieber getreuer **Joachim von Sandrart/** Unsers lieben Vettern des **Pfalzgrafen Philip Wilhelms zu Neuburg** Ld. Raht in Unterthänigkeit zu vernehmen gegeben/ was maßen er/ aus Lieb zu denen Edlen freyen Künsten/ wie auch zu Nutz und Dienst deroselben Liebhabern/ mit langer Mühe und Fleiß ein Werk zusammen getragen/ und solches in offnen Druck/ unter dem Namen und Titel L'Academia Todesca della Architectura, Scultura & Pittura, oder Teutsche Academie der Edlen Bau- Bild- und Mahlerey Künste/ ausgehen zu lassen/ vorhabens/ auch schon würlklich im Druck verfasst seye/ mit gehorsamster Bitte/ daß Wir ihme/ zu diesem Ende/ und damit von niemand solcher Tractat innerhalb zehen Jahren nachgedruckt werden möchte/ Unser

English version

Sandrart.NET

Startseite
Die Teutsche Academie im Volltext
Gliederung der Teutschen Academie
Übersetzungen

Suche
Personen
Orte
Kunstwerke
Bibliographie

Editionsinformationen
Projektinformationen
Herausgeber
Kontakt/Impressum

Permanente URL:
http://ta.sandrart.net/-person-1568

Leopold <Römisch-Deutsches Reich, Kaiser, I>

Basis-Daten

Kaiser des Heiligen Römischen Reiches (ab 1658), Erzherzog von **Österreich**, König von **Ungarn** (1655) und **Böhmen** (1656), geb. 1640 in **Wien**, gest. 1705 ebda.; Ursprünglich für den geistlichen Stand bestimmt, trat er 1654 nach dem überraschenden Tod seines Bruders **Ferdinand IV.** die Thronfolge an. Er war dreimal verheiratet. Er liebte Theater und Musik und gilt als begabter Komponist. Er übernahm als Erbe von **Erzherzog Leopold Wilhelm** dessen Gemäldesammlung. Vgl. **DBE**, Bd. VI, S. 330–331. (Für die Präzisierungen und Ergänzungen sei Renate Schreiber herzlich gedankt.)

Kategorie(n)
Herrscher und Fürsten

Externe Ressourcen und Referenzdatenbanken

Personennamendatei (PND): 118571869
Virtual Internet Authority File (VIAF): 54156429
Personeneintrag im »Biographisch-Bibliographischen Kirchenlexikon«
Personeneintrag in der »Deutschen Biographie«
Einträge zu dieser Person im »Digitalen Portraitindex«
Mit der Person verknüpfte Werke im »Virtuellen Kupferstichkabinett«

Vorkommen im Text

TA 1675, Titelblatt
»Mit Röm. Käys. Majest. gnädigst-ertheiltem Privilegio. Nürnberg/ Bey Jacob Sandrart/ auch in Franckfurt/ bey Matthaël Merians...«
Titelseite

TA 1675, Geleitwort (Leopold Wilhelm Graf von Königsegg) III

Personen-Bezüge

Mann von

- **Claudia Felicitas <Römisch-Deutsches Reich, Kaiserin>**
- **Eleonore <Römisch-Deutsches Reich, Kaiserin, 1655-1720>**

Gehört zu

- **Habsburger** Personengruppe

Kunstwerk-Bezüge Symbolansicht

Ist dargestellt in

- Kilian/Spillenberger: **Das Glück des Kaisers in Krieg und Frieden**
- Block: **Kaiser Leopold I. im Harnisch mit Feldherrnstab, Dreiviertelporträt**
- Sandrart: **Leopold I.**

Weitere Bezüge einblenden

War Besitzer(in) von

- Oosterwyck: **Vanitas-Stilleben**

Ist Widmungsempfänger(in) auf

- Kilian/Spillenberger: **Das Glück des Kaisers in Krieg und Frieden**

Orts-Bezüge

Steht in Zusammenhang mit

- **Burgund (Herzogtum)**

http://www.sandrart.net, 2012

Collation du texte

The screenshot displays the Juxta software interface for text collation. The window title is 'Juxta - welcome.jxt'. On the left, a 'Comparison Set' panel lists three files: 'BretonvilliersBrice1684.txt', 'BretonvilliersBrice1706.txt' (highlighted in green), and 'BriceHotelBretonvilliers1725.txt'. The main area shows the text from 'BretonvilliersBrice1706.txt' with various words and phrases highlighted in blue to indicate differences between versions. The text includes a description of Paris and a detailed account of 'La Maison de Bretonvilliers'. At the bottom right, a 'Source' panel shows the original text from the 1706 version. The interface also includes a toolbar at the top, a search bar, and a footer with the text 'juxta v1.6 - www.juxtasoftware.org'.

Collation de versions de Brice avec JuXta (<http://www.juxtasoftware.org>)

Voir aussi l'algorithme CollateX (<https://collatex.net>)

Données Outils

Couches

- Vestiges archéologiques
- Tissu urbain ancien
- Topographie historique
- Circonscriptions historiques
 - Limite d'interdiction de bâtir
 - Limite 1724-1729
 - Limite 1672-1674
 - Limite 1638
 - Voies et faubourgs 1724-1729
 - Rue de travers e
 - Grande rue
 - Interdiction de bâtir, bornes
 - Borne 1724 cat 2
 - Borne 1724
 - Borne 1672
 - Borne 1638
 - Censives Vasserot
 - ecclésiastique
 - laïc
 - inconnu
 - Censives source
 - ecclésiastique
 - laïc
 - inconnu
 - Paroisses 1300
 - Régions 1300
 - Cité
 - Outre-Grand-Pont
 - Outre-Petit-Pont
 - Quêtes 1300
- Dynamiques environnementales
- Données Vasserot V1



Vue globale

Géosignet

Légende

- Quartiers XIXe (Vasserot)
- Ass. Vasserot 1 (1810-1836)

<http://alpage.huma-num.fr/fr/>

ex. Pundit



Giulio Andreini Dashboard

Annotation composer Notebook composer

Create new annotation

Verrocchio identifies

verrocchio

Favourites TOPIC Leonardo, da Vinci

Freebase TOPIC Del Verrocchio, Andrea

Korbo TOPIC Ghirlandaio, Domenico

Del Verrocchio, Andrea / A Sleeping Y...

Also known as: Schlafender nackter Jüngling Artist also known as: Cione, Andrea; Cioni, Andrea

Filter annotations (31/31)

Verehrteter Herr Professor

Herzlichen Dank für Ihre freundliche Ausk...
namentlich auch für Ihre Beurtheilung uns...
Urtheil treffen Sie, glaube ich, für dieses Ge...
Bilder überhaupt ganz das Richtige: er ist d...
und durch, und daher gelingt ihm häufig die große Aufgabe noch nicht in dem...
Maße, wie den jüngeren, welche auf seinen Schultern stehen: Raphael, Fra...
Bartolommeo, gelglicht selbst Sodoma. Leonardo steht – das vergisst man fast...
immer bei seiner Beurtheilung – ja noch im Quattrocento, ist Altersgenosse von...
den Ghirlandajo's, Filippino, Signorelli, Perugino u.s.f. und steckt als langjähriger...
Gehilfe des Verrocchio in seinen früheren Werken noch sehr in dr Kunst des...
Quattrocento. Lermolieff, der die florentiner Kunst nicht kennt u. durch...
Verwechseln von Meistern wie Filippino u Garbo, Pollajuolo u. Verrocchio leider...
die Kenntniss dieser flor. Schule – wenigstens für uns Deutsche, die wir das...
Neue (wenn es mit Selbstbewusstsein vorgetragen ist) stets bewundern u für...
baare Münze nehmen – weit hinter den durch Cavalcaselle glücklich erzielten...
Standpunkt zurückgeschoben. Von dieser Seite muss daher auch unser Bild...
verdammt werden. In der That schreibt mir Morelli d.T., er könne in dem Bilde...
(soweit er nach einer Photo urtheilen dürfe) nur das Werk eines...
niederländischen Nachahmers von Leonardo sehen, der allerdings, nach «Ohr»,...
Landschaft, Gewandung etc zu urtheilen, theilweise direkt aus Leonardo's

PUNDIT

23 annotations

Web Annotation

Export notebook Edit notebook Delete

Home > Notebooks > Web Annotation

Filter annotations

Comment

Giulio Andreini Mar 17, 2016 · in Web Annotation

You might annotate an otherwise ordinary press release, product information page, technical documentation, a help page, or a speech.

Use cases

Go to the original page: How to annotate the web for the world, your group, or your students - TechRepublic · Delete

Giulio Andreini Dec 14, 2015 · in Web Annotation

personal note taking, peer review, copy editing, post publication discussion, journal clubs, classroom uses, automated classification, deep linking.

A good list of possible applications of web annotation for scholars... to remember

Go to the original page: Annotating All Knowledge · Delete

Annotation Type

Comment (23)

Web Page

Annotating the scholarly web : Nature News & Comment (5)

With Web Annotation, You Can Comment on Any Page—But Should You? - canmua.net (4)

Annotating All Knowledge (2)

Web Annotation Data Model (2)

About | Redux (1)

Annotation Summit at the New York Times – Hypothesis (1)

Annotation platform Genius moves into ads, controversy | Digital Content Next (1)

Exegesis: How early adapters, innovative publishers, legacy media companies and more are pushing toward the annotated web » Nieman Journalism Lab (1)

How to annotate the web for the world, your group, or your students - TechRepublic (1)

<http://thepund.it>

1.3 Un corpus connecté



LABEX

Contexte LOD

contexte de l'ouverture des données

- - contexte réglementaire européen d'ouverture des données de recherche
- - feuille de route web 3.0 du MCC

politique interne du Labex

- - politique d'ouverture des données assumée
- - définition d'un modèle de données commun
- - groupe ModRef

Qu'est-ce que le LOD ?

LOD Linked Open Data

répose sur le web sémantique

- - ressources & représentation
- - identifiants (URI)
- - utilisation de RDF et d'ontologies (RDFs, OWL, ...)

une manière d'utiliser le web

- - pour connecter des données en rapport
- - pour exposer des données et des informations sur le web sémantique

Un corpus connecté



**Semantic
Web**

Faire partie du Linked Open Data

- - **pour la distribution et les réutilisations**
- - **pour l'enrichissement**



Vue du Pont neuf et de l'Isle du Palais a Paris

Israel silvestre sculp. et ex.

3

Vue du Pont neuf et de l'Isle du Palais a Paris, Israël Silvestre sculp. et ex., 154 x 109

Cette image fait partie de la suite 53 : Vues de Paris et de France.

2. La création d'une API textuelle



LABEX

2.1 Structuration du corpus avec XML-TEI



LABEX

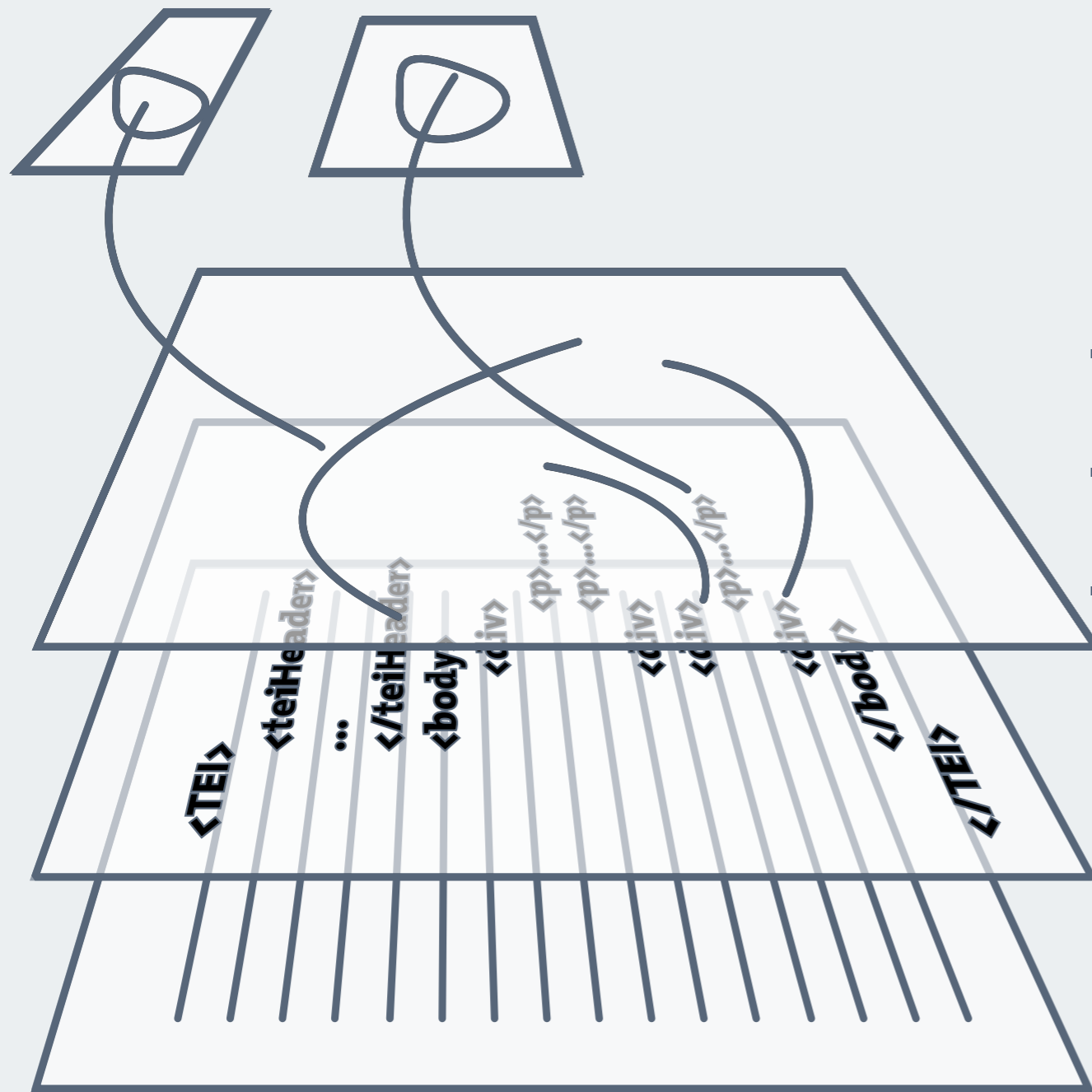
La TEI comme véhicule

- ¶ modélisation ODD
- ¶ guide d'encodage



- ¶ modélisation centrée sur
 - - transcription diplomatique (typographie, pages de titre, etc.)
 - - structuration du corpus
 - - identifiants canoniques (liages, etc.)
 - - entités-nommées et référentiel patrimonial

Les documents TEI ont trois couches sémantiques



- texte unicode
- balisage sémantique
- méta-structures

Hugh A. Cayless, TEI Journal, 2013

Pourquoi encoder en XML-TEI ?

¶ production de métadonnées

- - documenter le texte qu'on édite
- - pour la publication, le moissonnage

¶ enrichissement éditorial

- - signaler la structure du document
- - changements de pages, ratures, etc.

¶ enrichissement sémantique

- - nom de personnes, lieux, dates, abréviations, etc.

Tirer parti de la TEI ?

¶ documenter son travail

- - spécification formelle
- - production d'une documentation

¶ validation contre un schéma

- - contrôle du travail des éditeurs (collectif)
- - assurer la consistance des fichiers

¶ utilisation comme format pivot

- - chaîne d'édition multimodale
- - formats de sorties (html, epub, etc.)

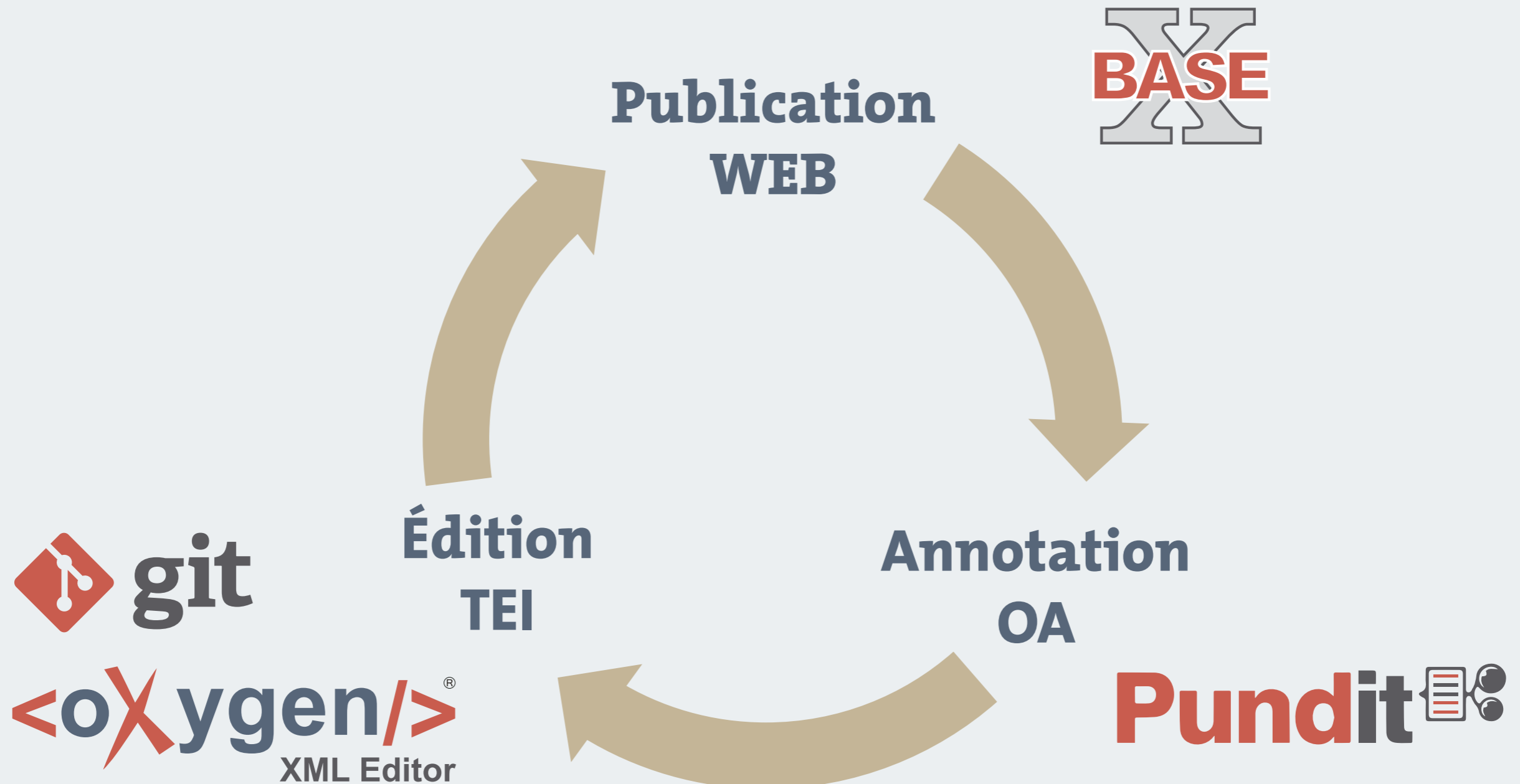
Présentation du corpus XML-TEI

2.2 Un cadre léger de publication avec XQuery



LABEX

Outils mis en œuvre



SynopsX

<http://synopsx.github.io>

cadre léger de publication XQuery

- - publier et exposer des corpus XML sur le web
- - application XQuery avec BaseX
- - scalabilité et personnalisation
- - système de templating

¶ exemples fonctionnels

- - *<http://rey.huma-num.fr>*
- - *<http://montesquieu.huma-num.fr>*
- - **<http://editions.ihpc.huma-num.fr/skepsis>**

collaboration non-institutionnelle

- - initié par l'Atelier des Humanités Numériques (AHN) de ENS Lyon

<http://ahnenslyon.hypotheses.org>

- - plusieurs projets de recherche (Ampère, Guides de Paris, etc.)

¶ modèle du logiciel libre

- - GNU General Public License >3

- - *<http://synopsx.github.io>*



Pourquoi SynopsX ?

¶ Collaborer

- - labex = recherche par projet
- - qualité du code & bonnes pratiques

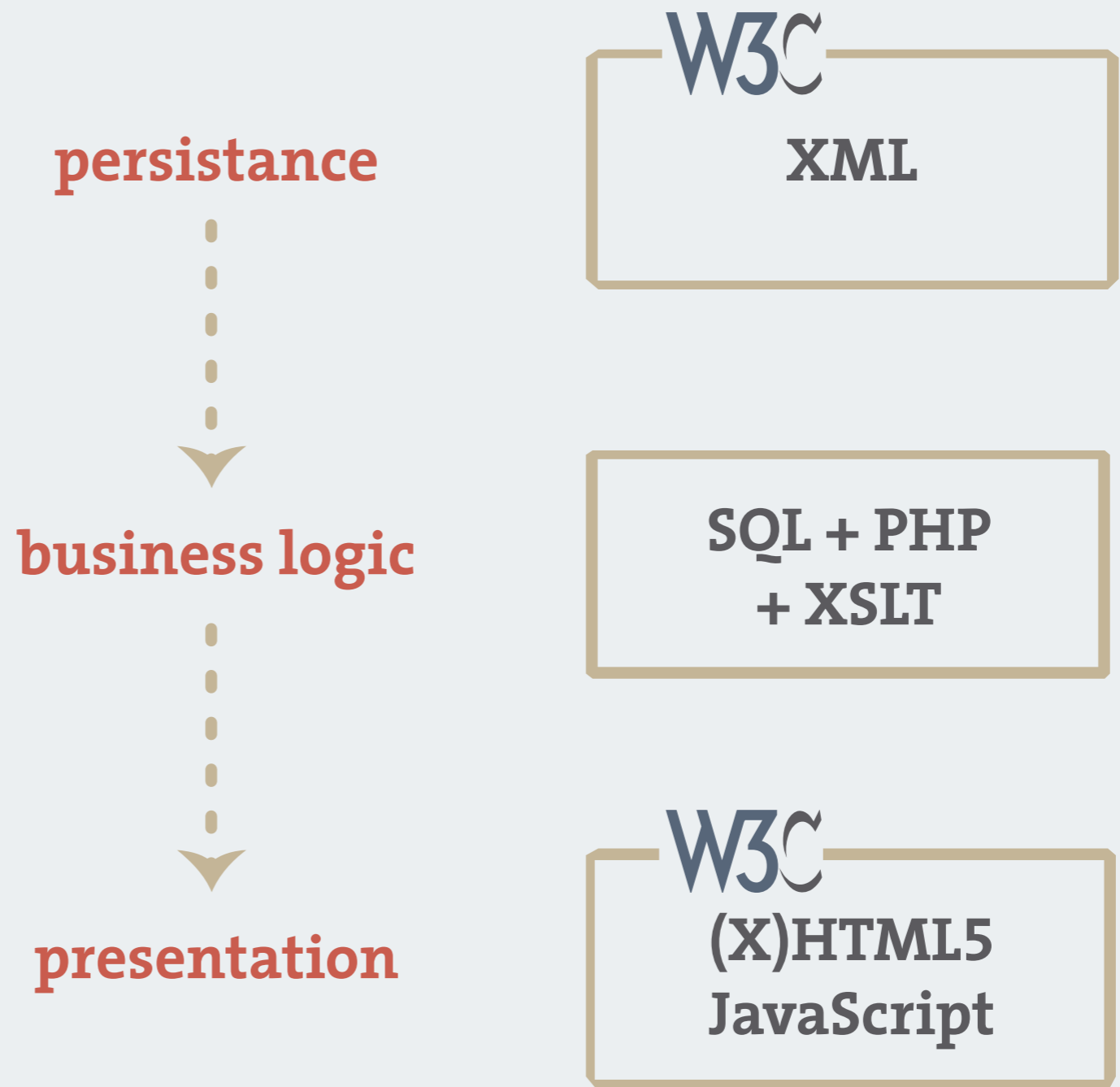
¶ Mutualiser

- - ne pas réinventer la roue
- - soutenabilité

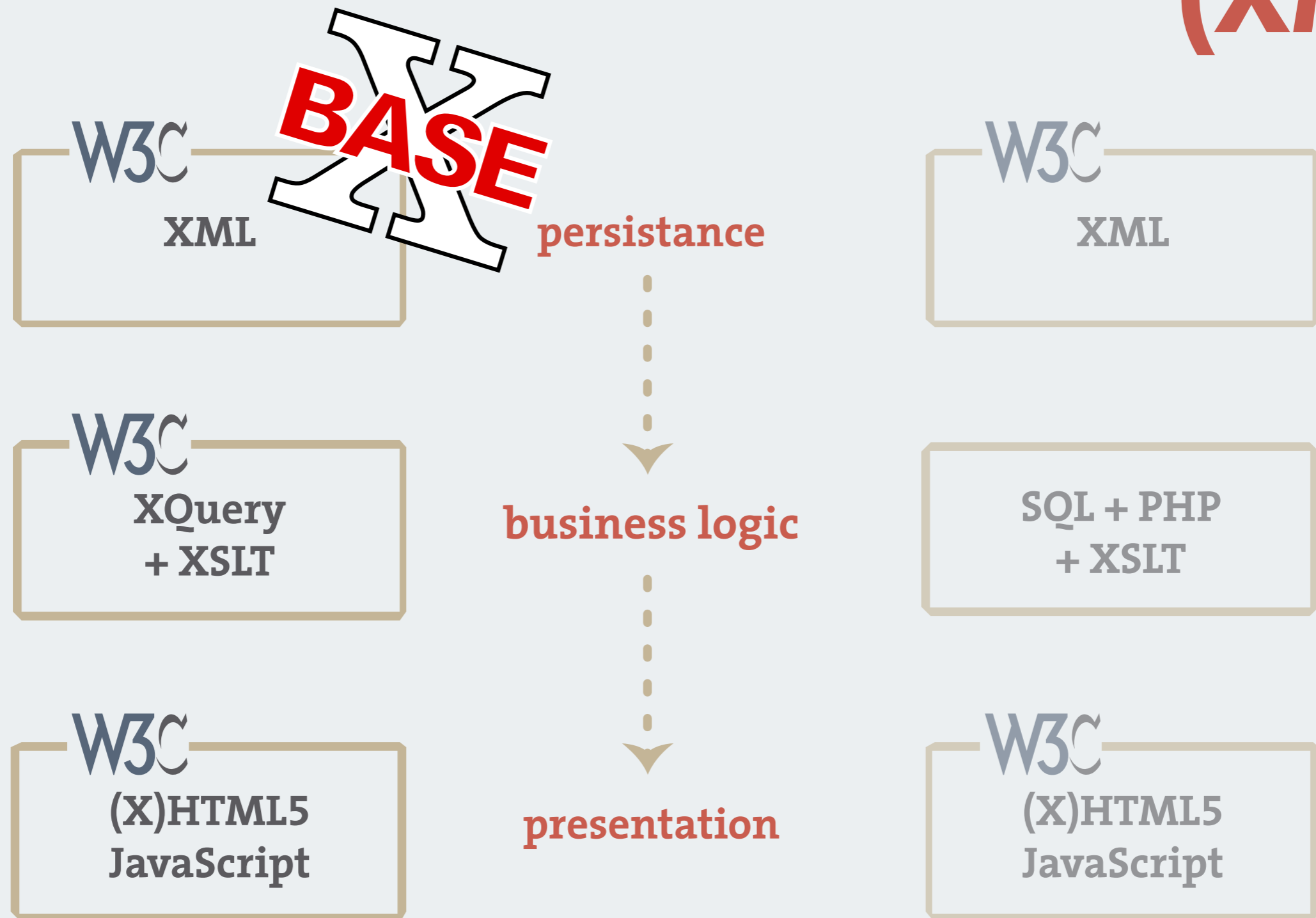
¶ Génériciser

- - maintenabilité
- - réutilisabilité

Structures traditionnelles des applications



des technologies uniformes (XML)



Principes de dévt de SynopsX

¶ keep it simple !

- un cadre de publication léger
- facile d'accès

¶ entièrement personnalisable

- ressources & représentation définies par l'utilisateur
- libre choix du scheme d'URIs
- surcharge et mécanismes d'héritage

¶ ~ MVC

- séparation du contenu et du rendu
- mécanisme de templating complet (mappings)

Bases de données XML

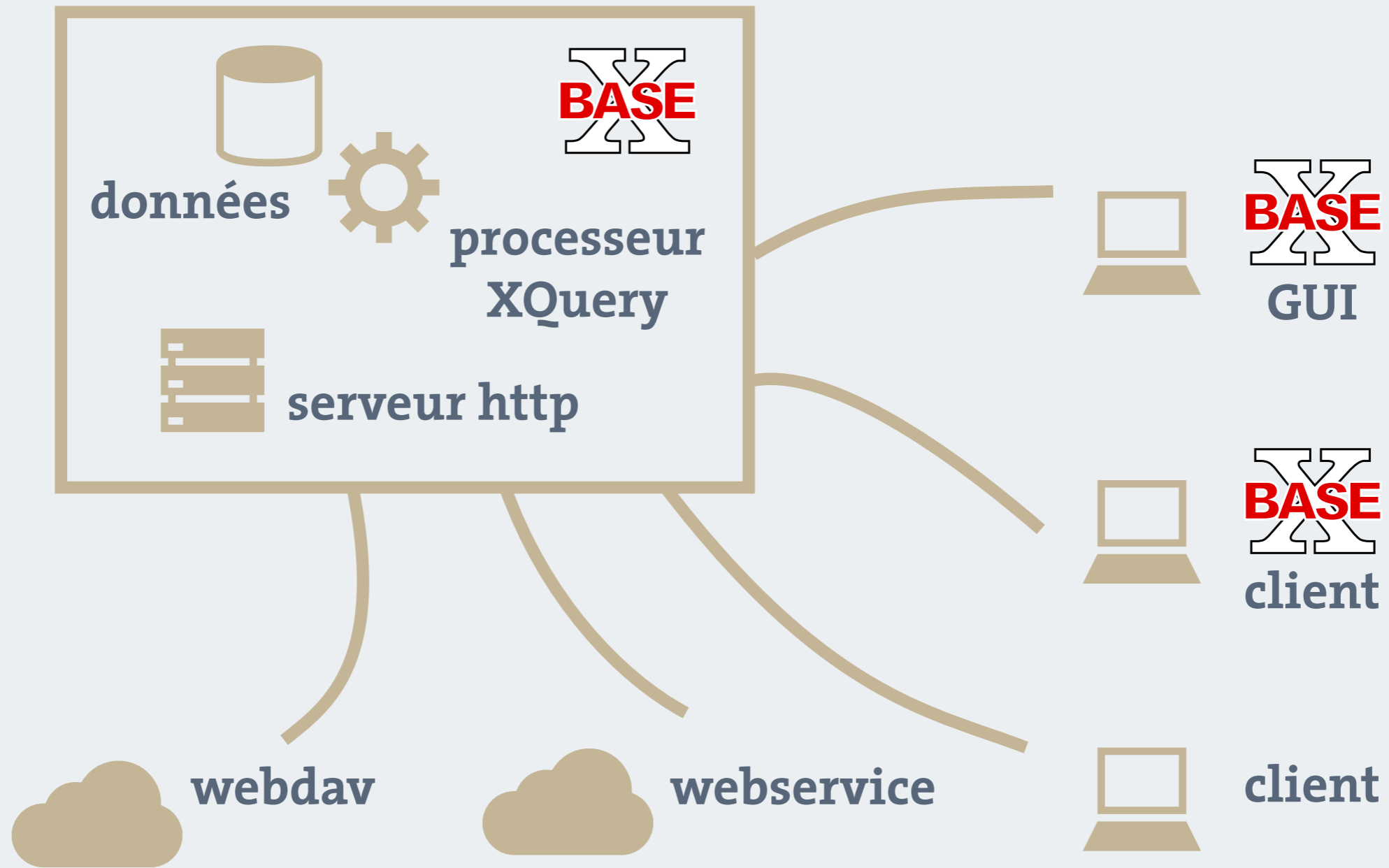
¶ avantages

- - préservation des structures de données XML
- - utilisation des spécifications de schémas
- - indexation des structures XML
- - accès à travers des API XML (XPath, XQuery, etc.)

¶ solutions libres & open source

- - **BaseX**
<http://basex.org>
- - **eXistDB**
<http://exist-db.org>

Composants & APIs de BaseX



¶ proposition de spécification

- - équivalent à l'API Java JAX-RS
- - annotation XQuery
- - proposition de spécification d'Adam Retter

<http://exquery.github.io/exquery/exquery-restxq-specification/restxq-1.0-specification.html>

¶ basé sur le protocole HTTP

- - RFC7230, RFC7232, RFC7233, RFC7234, RFC7235
- - cf. Roy Thomas Felding

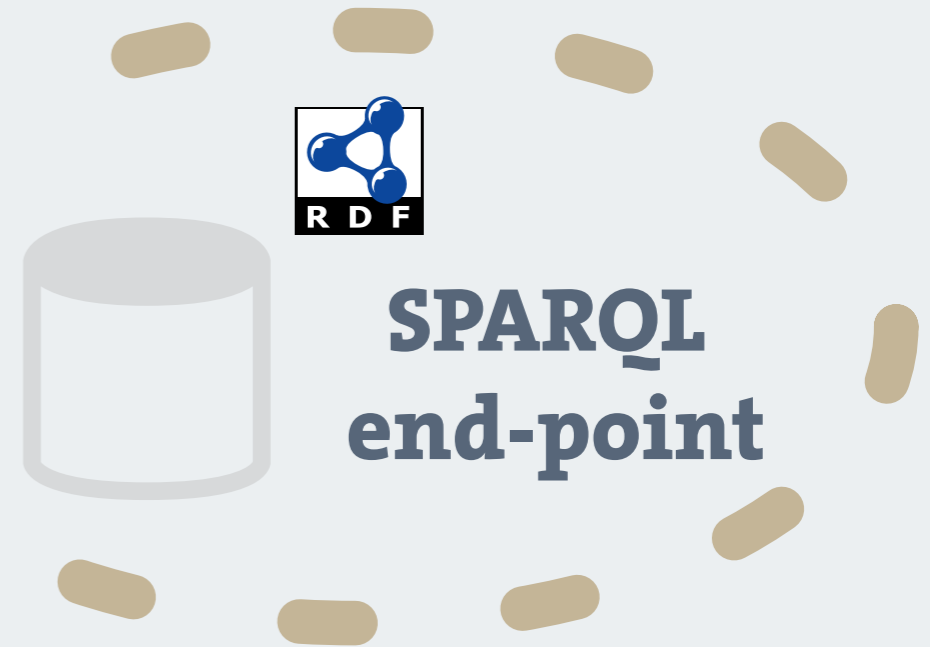
<http://www.ics.uci.edu/%7Efielding/pubs/dissertation>

ex. de fonction ressource en XQuery

```
xquery version '3.0' ;
module namespace demo = 'demo' ;
declare default function namespace 'demo' ;

declare
  %restxq:path('/demo')
  %rest:produces('text/html')
  %output:method('html')
  %output:html-version('5.0')
function titres() {
  <p>Hello World !</p>
};
```

De multiples formats de sortie



XML sources

representations

Resource, URI, représentations

URI

`http://guidesdeparis.net/corpus/sauval1724`

identifie



Ressource

`sauval1724`

Représentation

Métadonnées

Content-Type: text/html; charset=UTF-8

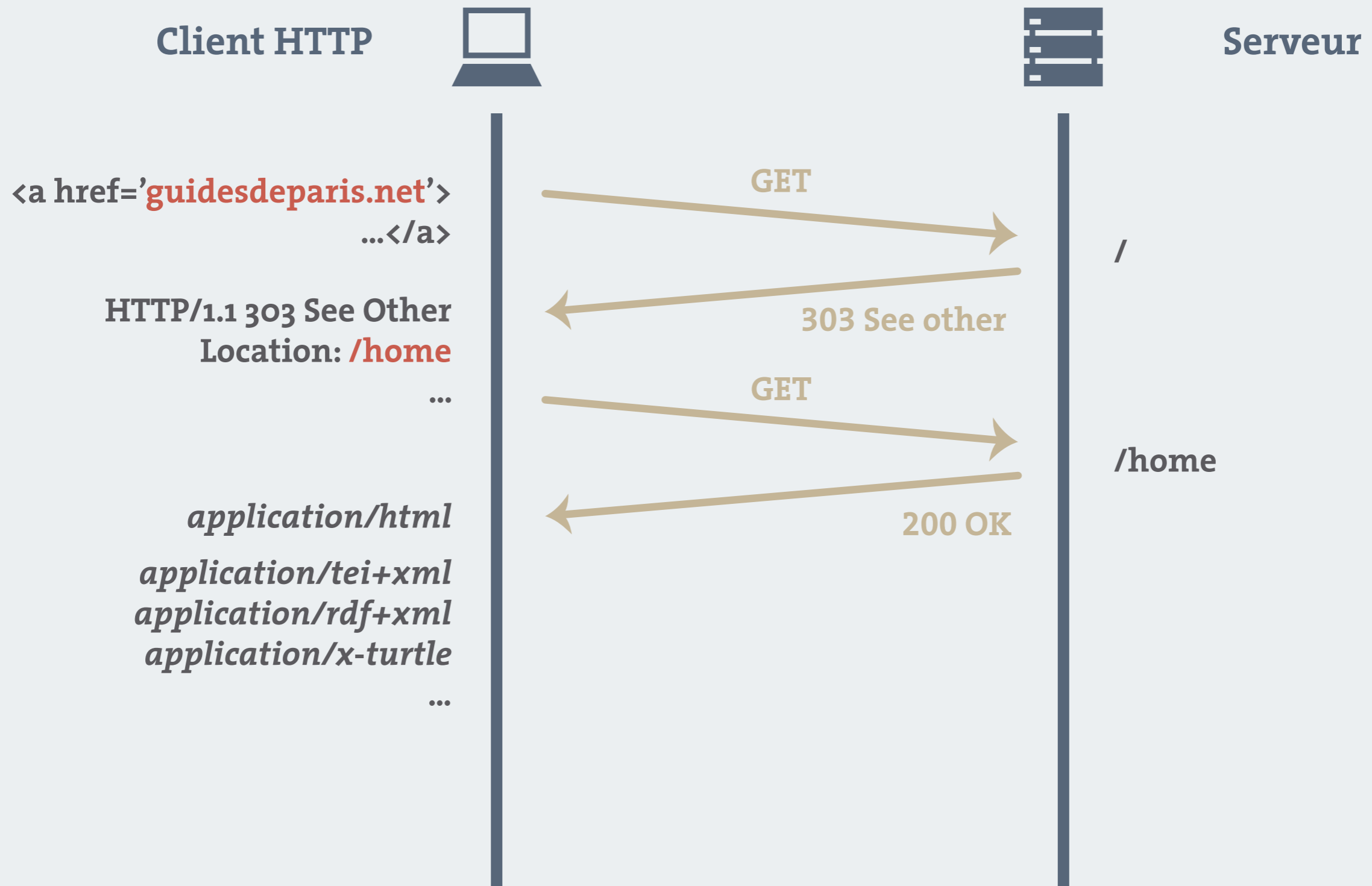
Données

```
<html xmlns="http://www.w3.org/1999/xhtml"
  lang="fr" xml:lang="fr">
  <head>
    <meta charset="utf-8"/>
    <title > { Titre du document } </title>
  </head>
  <body> { page content } </body>
</html>
```



représente

Négociation de contenu



démonstration de l'application SynopsX

3. L'exposition du coprus sous forme de données liées



LABEX

3.1 Le web de données culturel



LABEX

Données disponibles

	IDs	URIs	LOD	
	yes	yes	yes	+++
Institut national d'histoire de l'art	yes	--	next	++
	yes	yes	yes	++++
geo	yes	yes	yes	+++
ISNI	yes	yes	no	++
AN	yes	--	--	++
AKL	yes	no	no	+++

Web de données culturel quelques ex. structurants

¶ thesaurus

- - **Linked Open data Getty vocabularies (AAT, TGN, ULAN)**

<http://www.getty.edu/research/tools/vocabularies/lod/>

- - **Vocabulaires du MCC (France)**

<http://data.culture.fr/thesaurus/>

¶ jeux de données

- - *<http://data.bnf.fr>*
- - *<https://www.wikidata.org>*
- - *<http://www.geonames.org>*

Web de données culturel ex. domaine muséal

-- Amsterdam Museum

<https://hart.amsterdam/nl/page/12686/amsterdam-museum-en-linked-open-data>

-- British Museum

<https://collection.britishmuseum.org/resource/About>

-- American Art collaboratory (AAC)

<http://americanartcollaborative.org>

-- Musée du Prado

<https://www.museodelprado.es/en/modelo-semanticodigital/modelo-ontologico>

Web de données culturel biblio

Januals, B. et Minel, J-L. (2016). “ La construction d’un espace patrimonial partagé dans le Web de données ouvert ”. Communication, 34/1. En ligne : <http://communication.revues.org/6650>. DOI : 10.4000/communication.6650

Pascaline Tchienehom, «ModRef Project: from Creation to Exploitation of CIDOC-CRM Triplestores», The Fifth International Conference on Building and Exploring Web Based Environments (WEB 2017), May 21 - 25, 2017 - Barcelona, Spain - Regular paper - <https://hal.archives-ouvertes.fr/hal-01578878>

Pascaline Tchienehom, «Humanités Numériques et Web Sémantique : du langage naturel à une représentation computationnelle structurée et sémantique des données», Digital Humanities - Alliance of Digital Humanities Organizations (ADHO), August 8 - 11, 2017 - Montréal, Canada - Résumé long - <https://hal.archives-ouvertes.fr/hal-01579338>

EDM

Europeana Data Model

<https://pro.europeana.eu/resources/standardization-tools/edm-documentation>

Harmonisation des données culturelles (MCC, France)

*[http://www.culturecommunication.gouv.fr/Divers/
Harmonisation-des-donnees-culturelles](http://www.culturecommunication.gouv.fr/Divers/Harmonisation-des-donnees-culturelles)*

- **Ginco**
- **Onoma**

Feuille de route web 3.0 (MCC, France)

Métadonnées culturelles & transition Web 3.0

<http://www.enssib.fr/bibliotheque-numerique/documents/64776-feuille-de-route-strategique-metadonnees-culturelles-et-transition-web-3-0.pdf>

<https://www.inha.fr/fr/agenda/parcourir-par-annee/en-2016/decembre-2016/le-web-semantique-pour-les-donnees-culturelles.html>

Biblio

Sajus, B. & Leroi, M. (2016). Le développement du web des données culturelles: Les enjeux pour le ministère de la Culture et de la Communication. *I2D – Information, données & documents*, volume 53,(2), 46-47. <https://www.cairn.info/revue-izd-information-donnees-et-documents-2016-2-page-46.htm>

POC du Labex les Passés dans le Présent

<http://triplestore.modyco.fr:8080/ModRef>

<https://hal.archives-ouvertes.fr/hal-01578878>

3.2 La représentation des entités historiques du corpus



LABEX

Entités & formats

textes

- RDA (FRBR)
- XML-TEI

agent historiques

- dénominations, informations historiques
- XML-TEI ou EAC-CPF

lieux

- dénominations, informations historiques
- XML-TEI

objects culturels

- dénominations, informations historiques
- LIDO (CIDOC-CRM), XML-TEI (object) à venir ?

Un modèle conceptuel de référence

- - **conçu pour l'interopérabilité**
- - **une ontologie de domaine**
propose des définitions et des structures formelles pour décrire explicitement ou implicitement des concepts et leurs relations pour une utilisation dans la documentation du patrimoine culturel
- - **version OWL** *<http://erlangen-crm.org>*

Mappings

- - **avec EAD (2001)**
- - **convergence avec FRBR (FRBR-oo)**
- - **SIG TEI ontology**

TEI & CIDOC-CRM

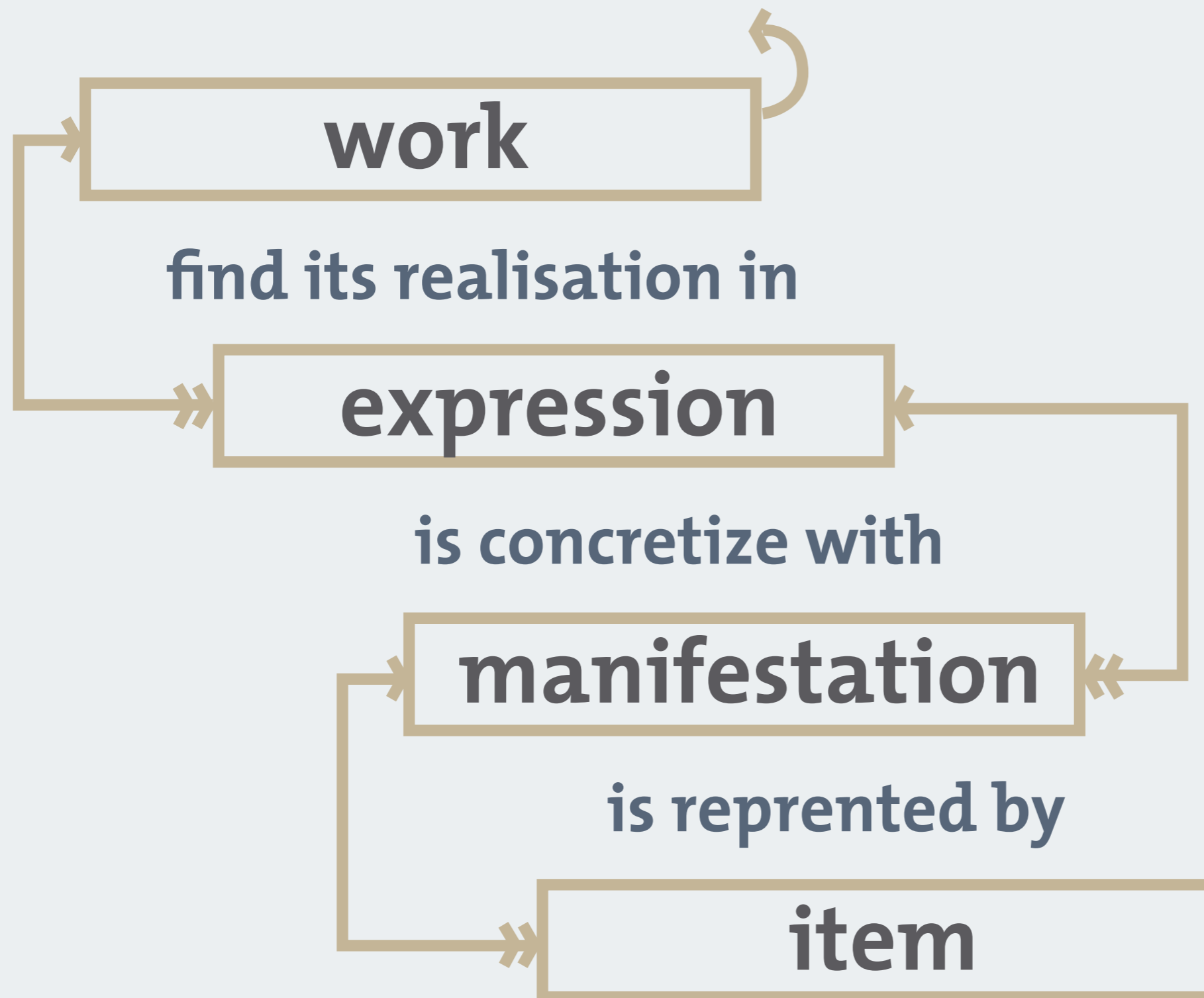
Mappings

- pour les métadonnées
- pour les entités place, person, org
- intégrées aux XSL Stylesheets

Biblio

Øyvind Eide (2014). Ontologies, Data Modeling, and TEI. TEI-Journal, n° 8. <https://jtei.revues.org/1191>

FRBR group 1



```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <listBibl xml:id="sauvalBiblio">
3   <head>Henri Sauval, Histoire et recherches des antiquités de la ville de Paris, ... </head>
4   <bibl xml:id="sauvalWork" type="work"> <author> <persName ref="#sauvalHenri"> <forename>Henri</forename> <surname>Sauval</surname> </persName> </author>
5   </author>
6   <title level="m">
7     > Histoire et recherches des antiquités de la ville de Paris, par Me Henri Sauval, ... </title> </bibl>
8   <biblStruct xml:id="sauvalExpression" type="expression">
9     <monogr>
10    <author>
11      <persName ref="#sauvalHenri"> <forename>Henri</forename> <surname>Sauval</surname> </persName>
12    </author>
13    <title level="m"> Histoire et recherches des antiquités de la ville de Paris, par Me Henri Sauval, ... </title>
14  </monogr>
15  <listRelation>
16    <relation type="work" corresp="#sauvalWork"/>
17  </listRelation>
18 </biblStruct>
19 <biblStruct xml:id="biblioSauval1724" type="manifestation">
20 <monogr>
21 <author>
22   <persName ref="#sauvalHenri"> <forename>Henri</forename> <surname>Sauval</surname> </persName>
23 </author>
24 <title level="m"> Histoire et recherches des antiquités de la ville de Paris, par Me Henri Sauval, ... </title>
25 <imprint>
26   <pubPlace>À Paris</pubPlace>
27   <publisher>Chés Charles Moette & Jacques Chardon</publisher>
28   <date when="1724">1724</date>
29 </imprint>
30 </monogr>
31 <idno type="catBnf"> http://catalogue.bnf.fr/ark:/12148/cb34108612p/ </idno>
32 <listRelation>
33   <relation type="expression" corresp="#sauvalExpression"/>
34 </listRelation>
```

**use relatedItem
instead ?**

La TEI est-elle appropriée pour les objets culturels ?



??



LIDO overview

Descriptive and administrative groups of information of a LIDO record

– Object Classifications –

Object / Work Type Classification

– Object Identifications –

Title / Name (mandatory) Inscriptions

Repository / Location

State / Edition

Object Description Measurements

– Events –

EventSet

– Relations –

Subject Set

Related Works

– Administrative Metadata – Rights

Record (mandatory)

Resource

Events in LIDO

Event

- Event Identifier
- Event Type
- Role in Event
- Event Name
- Event Actor
- Culture
- Event Date
- Period
- Event Place
- Event Method
- Materials / Technique
- Thing Present
- Event Related
- Event Description

Content / Subject in LIDO

Subject

- Extent Subject
- Subject Concept
- Subject Actor
- Subject Date
- Subject Place
- Subject Event
- Subject Object

enregistrement LIDO

```
1 <?xml version="1.0" encoding="UTF-8"?>
2 <lido:lidoWrap xmlns:lido="http://www.lido-schema.org" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:s
3 <lido:lido>
4 <lido:lidoRecID lido:source="Deutsches Dokumentationszentrum für Kunstgeschichte - Bildarchiv Fotothek Bonn" lido:ty
5 <lido:category>
6 <lido:conceptID lido:type="URI">http://www.cidoc-crm.org/crm-concepts/E22 </lido:conceptID>
7 <lido:term xml:lang="en">Man-Made Object</lido:term>
8 </lido:category>
9 <lido:descriptiveMetadata xml:lang="en">
10 <lido:objectClassificationWrap>
11 <lido:objectWorkTypeWrap>
12 <lido:objectWorkType>
13 <lido:term lido:encodinganalog="5230/5230">fountain</lido:term>
14 <lido:term lido:addedSearchTerm="yes" lido:encodinganalog="5230/523a">water distribution structure</lido:t
15 </lido:objectWorkType>
16 </lido:objectWorkTypeWrap>
17 <lido:classificationWrap>
18 <lido:classification lido:type="europeana:type">
19 <lido:term>IMAGE</lido:term>
20 </lido:classification>
21 <lido:classification>
22 <lido:term lido:encodinganalog="5220/5220">fountain sculpture</lido:term>
23 <lido:term lido:addedSearchTerm="yes" lido:encodinganalog="5220/520a">sculpture</lido:term>
24 </lido:classification>
25 </lido:classificationWrap>
26 </lido:objectClassificationWrap>
27 <lido:objectIdentificationWrap>
28 <lido:titleWrap>
29 <lido:titleSet>
30 <lido:appellationValue lido:pref="preferred">Fontana del Moro</lido:appellationValue>
31 </lido:titleSet>
32 </lido:titleWrap>
33 <lido:inscriptionsWrap/>
34 <lido:repositoryWrap>
35 <lido:repositorySet lido:type="current">
```

Créer un élément TEI personnalisé pour les objets culturels ?

LIDO un format d'échange

Agents et Lieux déjà dans la TEI

- ensemble du corpus édité en TEI
- modèle de contenu person suffisant pour nous
- modèle de contenu places suffisant pour nous

'traits', 'events', 'states' paraissent
bien correspondre aux besoins

Objectifs du projet

produire des mappings & construire une API

- - donner accès aux ressources via HTTP
- - offrir différentes représentations des ressources (html, tei, DC, rdf, txt)

offrir un SPARQL-endpoint

- - nourrir un SPARQL endpoint avec les triplets issus des mappings
- - enrichissements

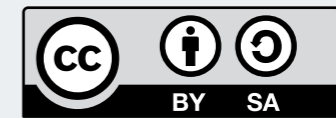
Ressources publiées

transcriptions des textes

- toutes dans le domaine public

édition critique

- textes annotés
- textes encodés en XML-TEI



référentiels & métadonnées

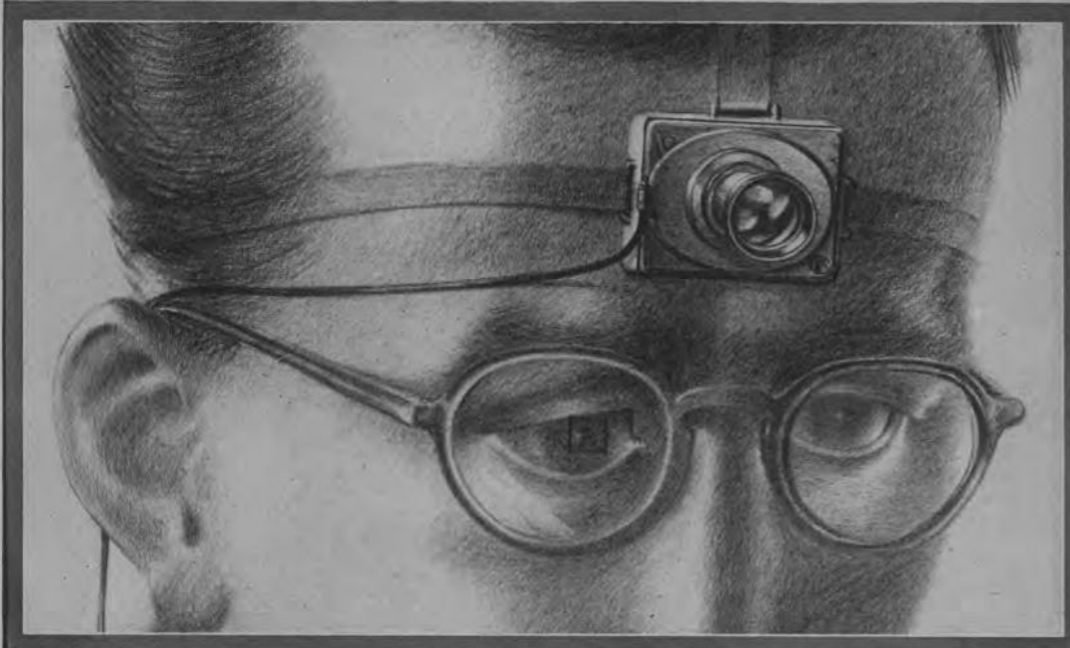
- référentiel sur le patrimoine parisien
- métadonnées réutilisables
(MCC, Bnf)

Conclusions



les
passés
dans le
présent

LABEX



A SCIENTIST OF THE FUTURE RECORDS EXPERIMENTS WITH A TINY CAMERA FITTED WITH UNIVERSAL-FOCUS LENS. THE SMALL SQUARE IN THE EYEGASS AT THE LEFT SIGHTS THE OBJECT

AS WE MAY THINK

A TOP U. S. SCIENTIST FORESEES A POSSIBLE FUTURE WORLD IN WHICH MAN-MADE MACHINES WILL START TO THINK

by VANNEVAR BUSH

DIRECTOR OF THE OFFICE OF SCIENTIFIC RESEARCH AND DEVELOPMENT

Condensed from the *Atlantic Monthly*, July 1945

This has not been a scientists' 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. What are the scientists to do next?

For the biologists, and particularly for the medical scientists, there can be little indecision, for their war work 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. They have worked in combined effort with the physicists of our allies. They have felt within themselves the stir of achievement. They have been part of a great team. Now one asks where they will find objectives worthy of their best.

* * *

There is a growing mountain of research. But there is increased evidence that we are being bogged down today as specialization extends. The investigator is staggered by the findings and conclusions of thousands of other workers—conclusions which he cannot find time to grasp, much less to remember, as they appear. Yet specialization becomes increasingly necessary for prog-

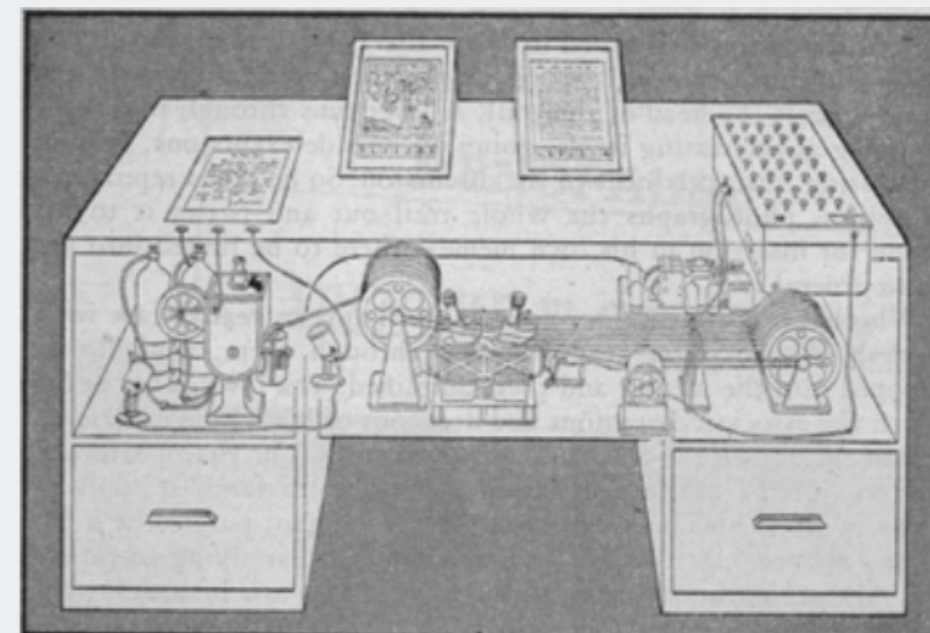
ress, and the effort to bridge between disciplines is correspondingly superficial.

Professionally our methods of transmitting and reviewing the results of research are generations old and by now are totally inadequate for their purpose. If the aggregate time spent in writing scholarly works and in reading them could be evaluated, the ratio between these amounts of time might well be startling. Those who conscientiously attempt to keep abreast of current thought, even in restricted fields, by close and continuous reading might well shy away from an examination calculated to show how much of the previous month's efforts could be produced on call.

Mendel's concept of the laws of genetics was lost to the world for a generation because his publication did not reach the few who were capable of grasping and extending it. This sort of catastrophe is undoubtedly being repeated all about us as truly significant attainments become lost in the mass of the inconsequential.

Publication has been extended far beyond our present ability to make real use of the record. The summation of human experience is being expanded at a prodigious rate, and the means we use for threading through the consequent maze to the momentarily important item is the same as was used in the days of square-rigged ships.

But there are signs of a change as new and powerful instrumentalities come into use. Photocells capable of seeing things in a physical sense, advanced photography which can record what is seen or even what is not, thermionic tubes capable of controlling potent forces under the guidance of



Vannevar Bush, « As we May Think », *Atlantic Monthly*, Juillet 1945.

<http://www.guidesdeparis.net> (bientôt !)

mleblanc@parisnanterre.fr

[@emchateau](#)



LABEX

<http://passes-present.eu>