



Le futur des logiciels de gestion bibliographique

Frédéric Clavert
LabEx Écrire une Histoire Nouvelle de l'Europe / Paris 4
@inactinique / frederic@clavert.net
<http://histnum.hypotheses.org> / <http://zotero.hypotheses.org/>

Introduction

- Pourquoi s'intéresser aux logiciels de gestion bibliographique? (expérience personnelle)
 - Recherche: Thèse (1999-2006): vide à combler
 - Entrée modeste mais décisive dans les Humanités numériques
 - Répond à des problématiques techniques numériques qui se posent aux chercheurs
 - Structurer ses données comme Monsieur Jourdain faisait de la prose

Introduction

- À partir du moment où l'on structure ses données
 - Logiciels de gestion de bibliographie comme plateformes, comme points de départ.
- On peut « capitaliser » sur ces données structurées
 - À titre individuel,
 - À titre collectif,
 - Si les logiciels ont été prévus pour (Zotero / Mendeley).

Zotero et Mendeley:
des plateformes



Des logiciels prévus pour être des plateformes

L'indispensable condition:
l'API

« Ensemble normalisé de classes, de méthodes ou de fonctions qui sert de façade par laquelle un logiciel offre des services à d'autres logiciels. »

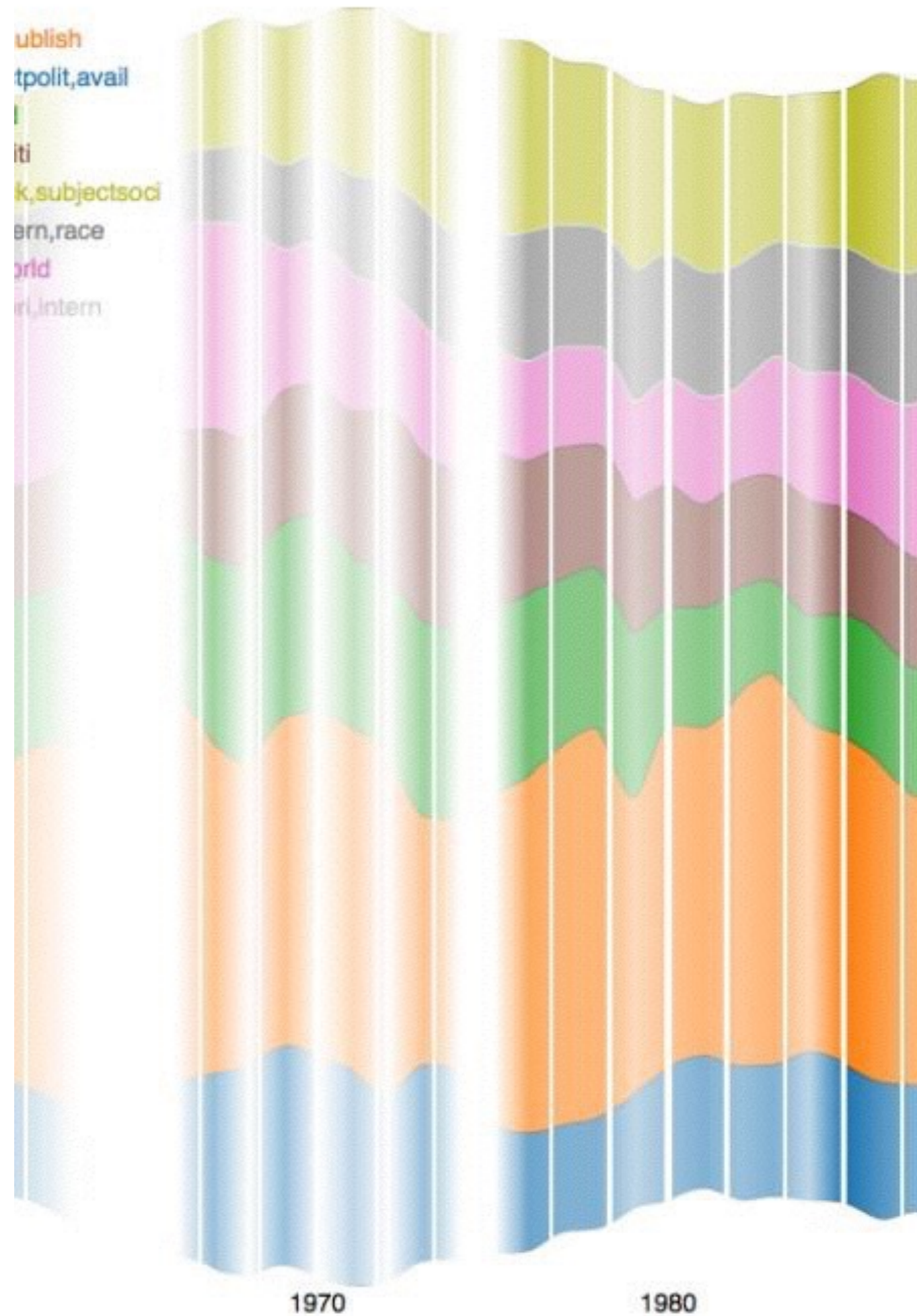
Des logiciels conçus pour communiquer

- Zotero comme Mendeley ont été conçus dès le départ pour « communiquer »
 - Rendent possibles les échanges de données avec des logiciels tiers via leurs API
- A permis de nombreux éléments constitutifs de leurs succès
 - Interaction application locale / site web (y compris couche sociale) / application mobile
- Ex: Zotero

Pertinence d'une API

- Construire sur la compétence des autres / Permettre aux autres de construire sur le logiciel de gestion bibliographique
 - Ex: le « contest » API de Mendeley de 2011
Battle Apps for science contest
- Grande diversité des applications développées:
 - vocabulary.se: montrer le contexte d'un article / d'une recherche
 - <http://www.papercritic.com/>: utiliser Mendeley pour mettre en place un nouveau type de peer-review
 - <http://ropensci.org/> : utiliser Mendeley pour favoriser l'open science

Travailler ses données bibliographiques personnelles



Données bibliographiques?

- Ici, données bibliographiques dans un sens large
 - Données issues des sources primaires
 - Données issues des sources secondaires
- Point important
 - Avoir mis en place de bonnes méta-données dans Zotero
 - Avoir lié les documents

Zotero comme outil de lecture distante

Lecture distante?

« [...] what we really need is a little pact with the devil: we know how to read texts, now let's learn how *not* to read them. »

Moretti, Franco, *Distant Reading*, Verso Books, 2013, emplacement 793, édition Kindle.

Zotero comme outil de lecture distante

- Pratiquer une « lecture distante » de ses sources primaires ou secondaires implique
 - Données structurées,
 - Connaissances « techniques » (sens large: statistiques, informatiques notamment) assez poussées,
- Paper Machines (« Visualize your Zotero Library ») simplifie cette démarche

Zotero comme outil de lecture distante

- Sur la base de sa bibliothèque zotero, Paper Machines va ainsi permettre de faire:
 - Nuage de mots,
 - N-grams,
 - Phrase Net,
 - Topic modelling,
 - Mapping,
 - etc.
- Avec intégration des fils RSS, utilisations de ce type seront facilitées.

Capitaliser sur des données collectives



L'eldorado de la data / Le monde mis en données

Les logiciels bibliographiques sont à considérer en fonction des grandes thématiques du moment - notamment celles de la montée en puissance des grands ensembles de données.

«To datafy a phenomenon is to put it in a quantified format it can be tabulated and analysed»

« Big Data as a cultural, technological, and scholarly phenomenon that rests on the interplay of:

- *Technology*: maximizing computation power and algorithmic accuracy to gather, analyze, link, and compare large data sets.
- *Analysis*: drawing on large data sets to identify patterns in order to make economic, social, technical, and legal claims.
- *Mythology*: the widespread belief that large data sets offer a higher form of intelligence and knowledge that can generate insights that were previously impossible, with the aura of truth, objectivity, and accuracy. »

Viktor Mayer-Schönberger and Kenneth Cukier. *Big Data: A Revolution That Will Transform How We Live, Work, and Think*. Boston: Houghton Mifflin Harcourt, 2013, p. 72.

L'eldorado de la data

- Dans le cas des données bibliographiques, quelle exploitation des données?
 - Systèmes de recommandations (cf. RedCube)
 - Bibliométrie « classique »
 - Évaluation de la recherche
 - Analyse de la recherche
 - Bibliométrie alternative
 - « altmetrics »
- Etc...

L'exemple de Zotero

- Analyse de données anonymisées / Altmetrics
 - 21 avril 2014: Annonce d'un nouveau financement (fondation Alfred P. Sloan) / Montréal et Indiana
 - Comparer le lectorat/les utilisateurs de Zotero avec d'autres jeux de données, y compris sociaux et commerciaux
 - API permettant des recherches en bibliométrie et l'intégration de données bibliométriques en provenance de Zotero dans des programmes tiers.
 - Nouvelles métriques pour mesurer les impacts de la recherche, distinctes des fameux impact factor et h-index. L'une des spécificités de ces métriques est de ne pas reposer uniquement sur le nombre de citations dans les livres et articles scientifiques, mais aussi sur leur impact au sein des médias sociaux notamment.

L'exemple de Zotero

- Andrew Mellon Foundation / George Mason et Penn State
- L'intégration de flux RSS ou Atom dans Zotero
 - intégrer dans une bibliothèque zotero des articles / numéros de revues venant de paraître;
 - intégration de zotero avec les dépôts institutionnels en OpenAcces.
 - impliquera développement d'une API "push" d'un côté et un service qui connectera Zotero à des applications web.

Un exemple de « Big Data » via Zotero

- L'Old Bailey On-Line
 - cf. rapport *Data mining with criminal instant*



The Proceedings of the OLD BAILEY London's Central Criminal Court, 1674 to 1913

Home Search About The Proceedings Historical Background API The Project Contact

Celebrating a Decade of Making History Accessible: The Old Bailey Online, 2003-2013

Home Page
Search
About the Proceedings
Historical Background
API
The Project
Copyright & Citation Guide
Contact
Research and Study Guides
Login / Register

ON THIS DAY IN... 1761
Robert Greenstreet allegedly committed petty treason by stabbing his master with a knife and killing him. [read more](#)

The Proceedings of the Old Bailey, 1674-1913
A fully searchable edition of the largest body of texts detailing the lives of non-elite people ever published, containing 197,745 criminal trials held at London's central criminal court. If you are new to this site, you may find the [Getting Started](#) and [Guide to Searching](#) videos and tutorials helpful.

To search the Proceedings use the boxes on the right or go to the [Search Pages](#).

April 2013 Update: Celebrating our Tenth Anniversary
April 2013 marks the tenth anniversary of the launch of the Old Bailey Online, and the 100th anniversary of the last published edition of the *Old Bailey Proceedings*. The year's update includes a few minor corrections. For further information, see [What's New \(April 2013\)](#).

Tales from the Old Bailey

SEARCH
the Proceedings
Keyword(s)
Reference No.
Search in
<All Text>
SEARCH
More Search Options

The Proceedings can also be searched in:
connectedhistories
Integrated search of major

Mendeley, Elsevier et les données

- Une orientation « métriques » très prononcée
 - [‘A Look at Mendeley Readership Statistics’](#)
 - [‘Meet the Analytics Team!’](#)
 - [‘Meet the API Team!’](#)
 - [‘Meet the Data Science Team!’](#)
- Inscrite dans le design même du site web et, surtout, de sa couche sociale.
- La couche sociale du site fournissant probablement elle-même beaucoup de données supplémentaires
- Des équipes API / data / analytics très fournies
- Une participation active à la fouille de données dans le domaine de la bibliométrie
 - [‘Finding Better Ways of Mining Scientific Publications’](#)

Conclusion



Logiciels à insérer dans la recherche / dans la société

- Enjeux de ces logiciels sont ceux de la recherche
 - Open Access
 - L'usage des outils informatiques (lecture distante)
- Enjeux de ces logiciels sont ceux de la société à l'ère numérique
 - Big data (évolution logique: le web comme base bibliographique)
 - Données personnelles
 - Quantification de manière générale (de soi / des autres)

De la gestion bibliographique vers la gestion de la recherche

