

EEXCESS or the challenge of privacy-preserving quality recommendations

Benjamin Habegger, Nadia Bennani, Elöd Egyed-Szigmond, Omar Hassan
Lyon University, CNRS, INSA-Lyon, LIRIS, UMR5205



EEXCESS - Project

Enhancing Europe's eXchange in Cultural Educational and Scientific reSources.

Started: February 2013

Duration: 4 Years

Budget: 7.05 Million EUR

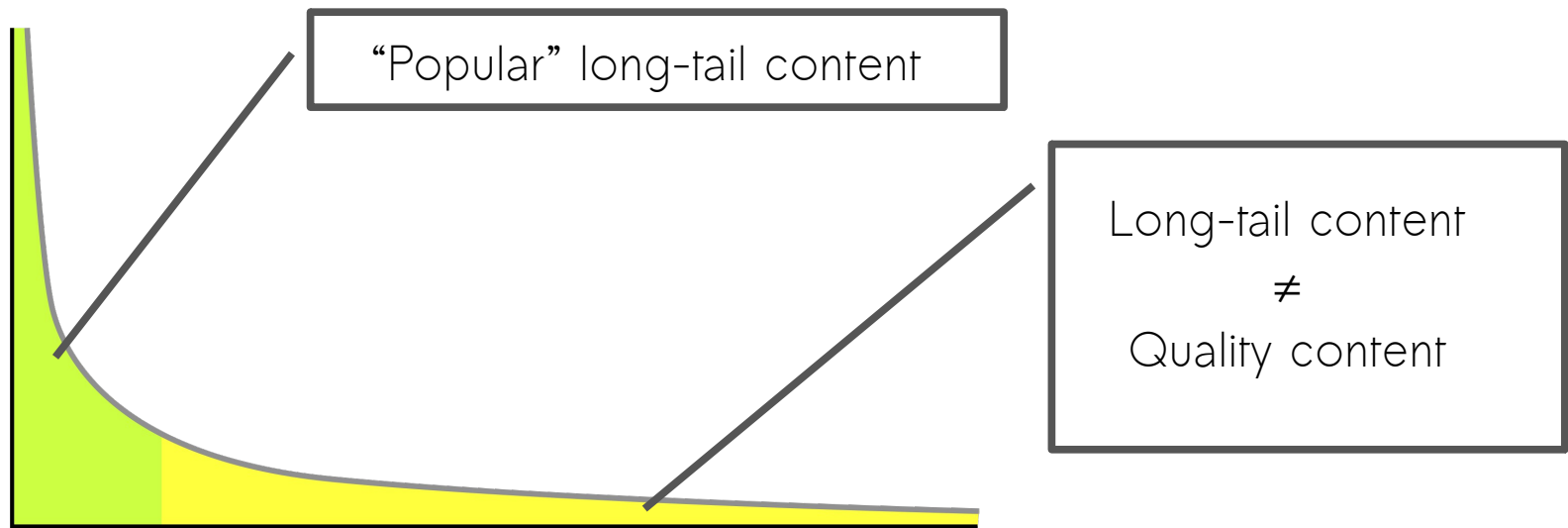
<http://eexcess.eu/>

EEXCESS - Partners



EEXCESS - Main problem

Making existing quality content visible



EEXCESS - Objectives



Content enrichment

Personalized recommendation



Privacy preservation

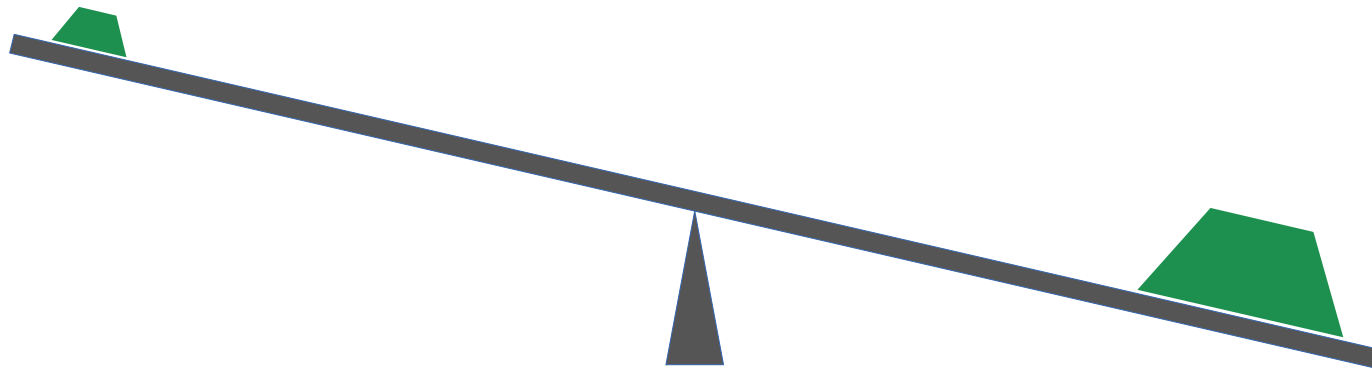
EEXCESS - Global challenges

- Federated recommendation
- Integration of multiple document sources
- Mining for user profiles
- Adapting user interfaces to context
- **Preserving user privacy**

EEXCESS - Tradeoff ?

Privacy

Quality

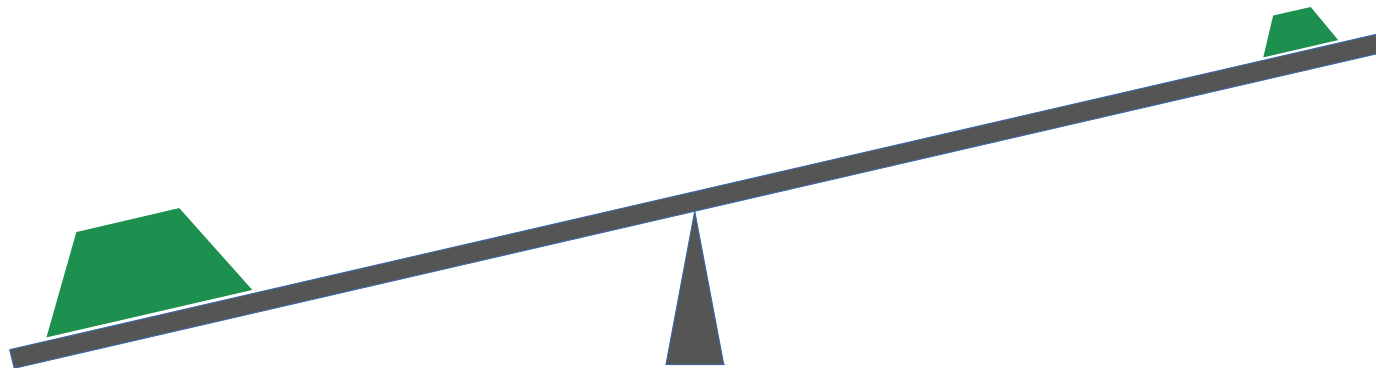


- Unlimited disclosure
 - User privacy is clearly at stake.
 - Does it improve quality ? How much ?

EEXCESS - Tradeoff ?

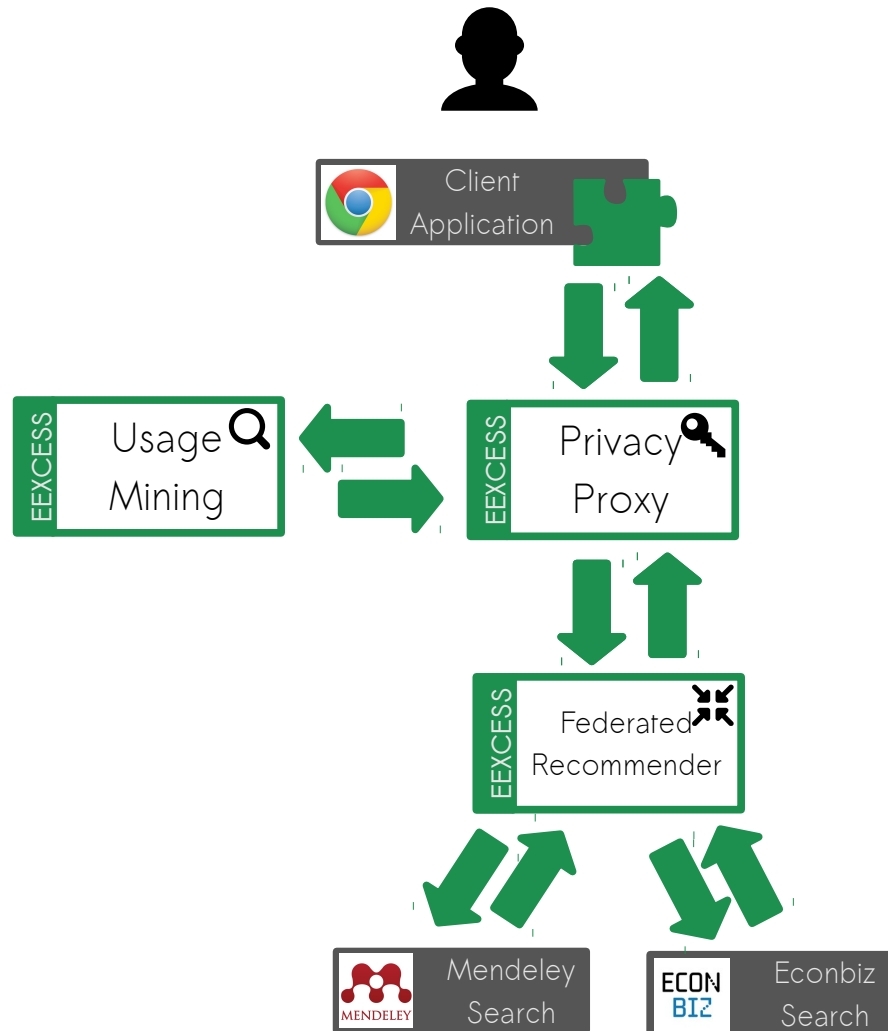
Privacy

Quality



- Limited disclosure
 - Limit recommendation quality ?

EEXCESS - Simplified architecture



Usage mining

→ build detailed user profiles

EEXCESS - User context

- **Activity context**
 - Browsing history
 - Ongoing tasks
 - Reading history
- **Environmental context**
 - Temperature, Humidity, Light
 - Things, Services
- **Personal context**
 - Weight, Pulse, Blood pressure, Mood
- **Social context**
 - Friends
 - Neighbors
 - Co-workers
 - Relatives
- **Spatio-temporal context**
 - Time
 - Location
 - Direction of movement

EEXCESS - User profile

- **Demographic information**
 - Age
 - Gender
 - Relationship status
 - Address
 - ...
- **Interests**
 - Professional interests
 - Personal interests
 - Interest in commercial products
- **Knowledge, background, skills**
 - Knowledge within a domain (e.g. acquired by a student)
 - Professional expertise and skills
- **Goals**
 - Short term goal (current task)
 - Long term goal
- **Behavior**
 - Repetitive behaviors
 - History of user actions

Recommendation

→ find recommendations adapted to
the user profiles

EEXCESS - Recommendation

- Content-based
- Collaborative filtering
- Demographic filtering
- ...

Privacy preserving recommendation
→ compromise between private
information disclosure and quality

EEXCESS - Privacy questions

- What profile information is useful ?
 - For usage mining ?
 - For quality recommendations ?
- How much and how detailed should this information be disclosed ?
 - To preserve privacy ?
 - To ensure recommendation quality ?
- What control and feedback can we provide ?
 - To „measure“ the impacts of disclosure

EEXCESS - Privacy wish-list

- **Guarantees**
 - Deterministic ?
 - Guarantee that a particular piece of information doesn't leak
 - Risk of disclosure (including inference)
 - Measuring risks and impacts of privacy breaches
- **Flexibility**
 - User-dependant policy
 - Alice and Bob may have different requirements
 - Context-dependant policy
 - Alice may have different requirements at home and at work

EEXCESS - Privacy techniques

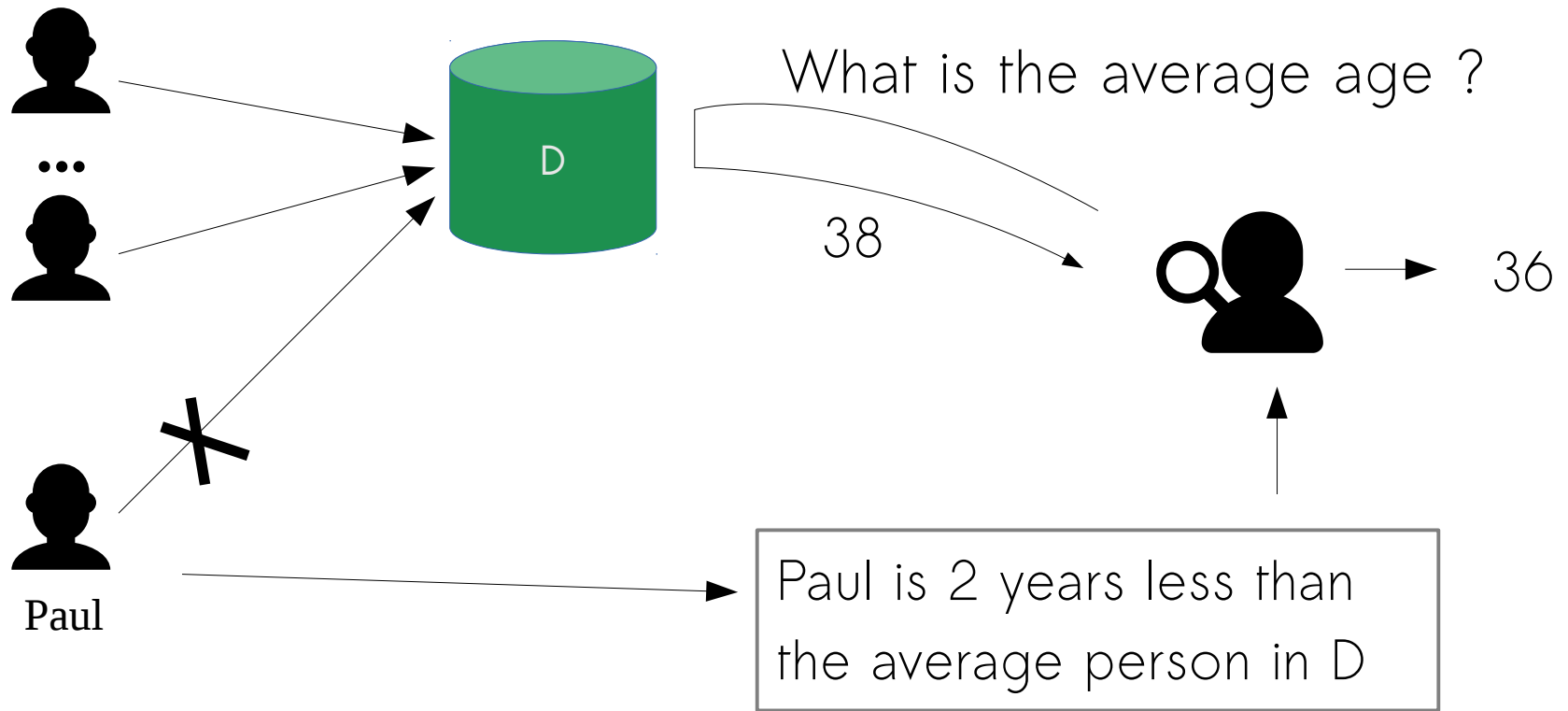
- Anonymization
 - K-anonymity [Sweeney2002]
- Differential Privacy [Dwork2006]
- Hiding in the crowd
- Distributed recommendation
- ...

EEXCESS - Privacy challenges

- Data mining, Big data analytics
- External knowledge
 - De-anonymization [Narayanan2008]
 - Breaches w.o. participation [Dwork2006]

EEXCESS - Privacy challenges

Privacy breach w.o. participation [Dwork2006]

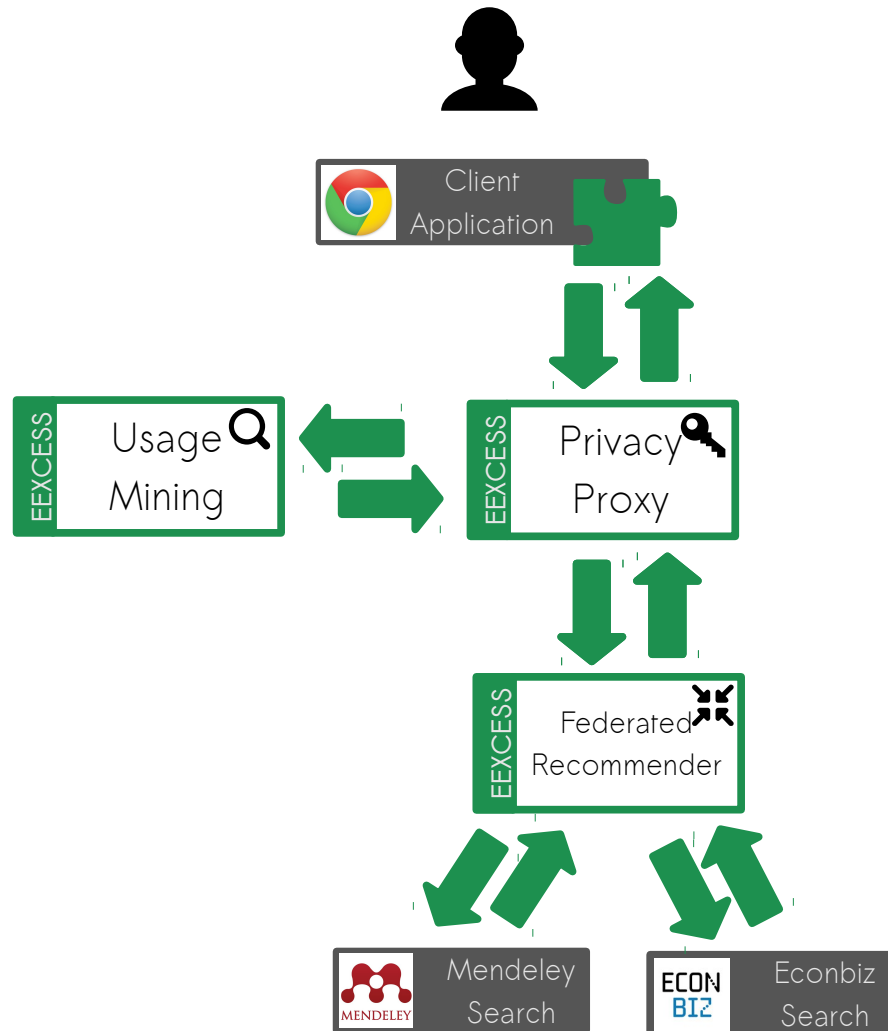


Ongoing work

EEXCESS - Current focus

- Setup a test-framework
 - User interface
 - API's
 - Evaluation tools
- Involve the user
 - Transparency
 - What is going on ?
 - What does the system know ?
 - Control
 - Let users define their own policy
 - Feedback
 - Show the impacts of user's policy

EEXCESS - Simplified architecture



Privacy Plugin - Recommendations

Instagram - Alpine sheep [Chamonix] x chrome-extension://aimdh x

ser.blog.lemonde.fr/2013/09/20/indelicat-sur-instagram-le-president-tchetchene-a-des-problemes-de-copyright/#xtor=RSS-32280:☆

Internet, les photographes sont habitués à ce que, parfois, leur t
pié sans leur autorisation.

sont moins, en revanche, à ce que les indélécats soient des ch

t ce qui est arrivé au début du mois à Herbert Schröer, photogra
'est aperçu qu'un de ses clichés avait été posté, sans son cons
ident tchétchène, Ramzan Kadyrov, sur son compte Instagram.
ialisé PetaPixel.

urait pu comprendre la démarche de M. Kadyrov, un grand ama
al, si l'image originelle — des moutons placides sur fond de mo
igées — avait été prise dans le Caucase tchétchène. Il aurait s
right sur l'autel de la promotion de la petite république de la fédé
sie. Mais les moutons sont savoyards puisque le cliché a été pri
nonix.

ui a forcément déplu à M. Schröer, qui s'est épanché sur Twitte
te !!! [sic] Le président de Tchétchénie m'a volé ma photo Insta
par-dessus."

Herbert Schröer
@HerbertSchroer

oly shit!!! The president of Chechnya has stolen my
stagram photo and put a filter over it.
[instagram.com/p/d06q_SiRqy/](https://www.instagram.com/p/d06q_SiRqy/) @instagram

18 AM - 4 Sept 2013

RETWEETS 10 FAVORIS

EEXCESS
Your traces Synchronize Mendeley Log Out

Welcome Benjamin Habegger
Environnement is set to : work

Home Work

EB [Analyse économique de la vie ++ - Réflexions sur la durée du droit d'auteur](#)
Score : 0.47789204


[Universitaires : "personne ne peut se contenter de l'actuel statut"](#)

EB [DREPTUL DE AUTORSI CONFLICTUL DE LEGI - DROIT D'AUTEUR ET CONFLIT DE LOIS- COPYRIGHT AND THE CONFLICT OF LAWS \(Romanian version\)](#)
Score : 0.45610753

[Settings](#) [More about the EEXCESS project](#)


AIRFRANCE
Pour découvrir toutes nos offres
CLIQUEZ ICI

Privacy Plugin - Profile edition

EEXCESS 

[Traces](#) [Profile](#) [Privacy Sandbox](#) [Settings](#)

Profile



Username: Benjamin Habegger

Email: benjamin.habegger@insa-lyon.fr
▶ Change email

Title: Mister
▶ Change title

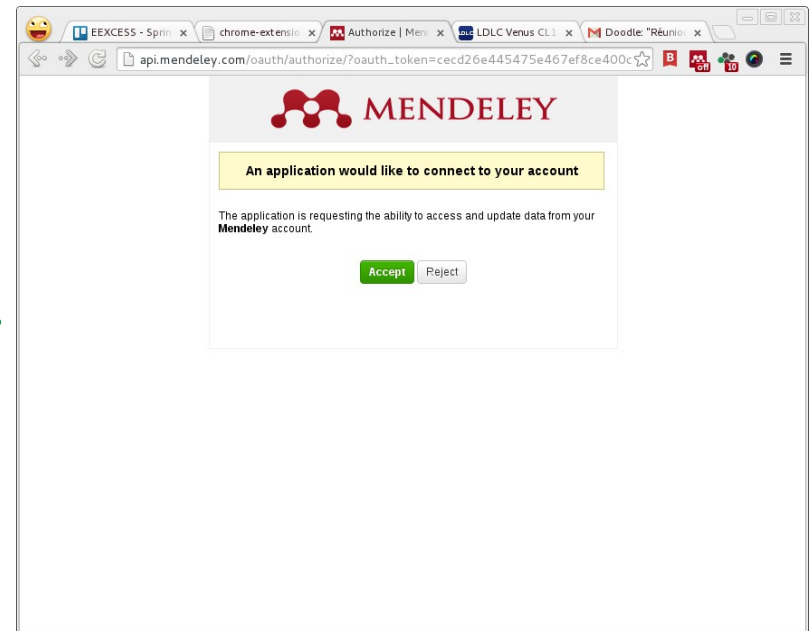
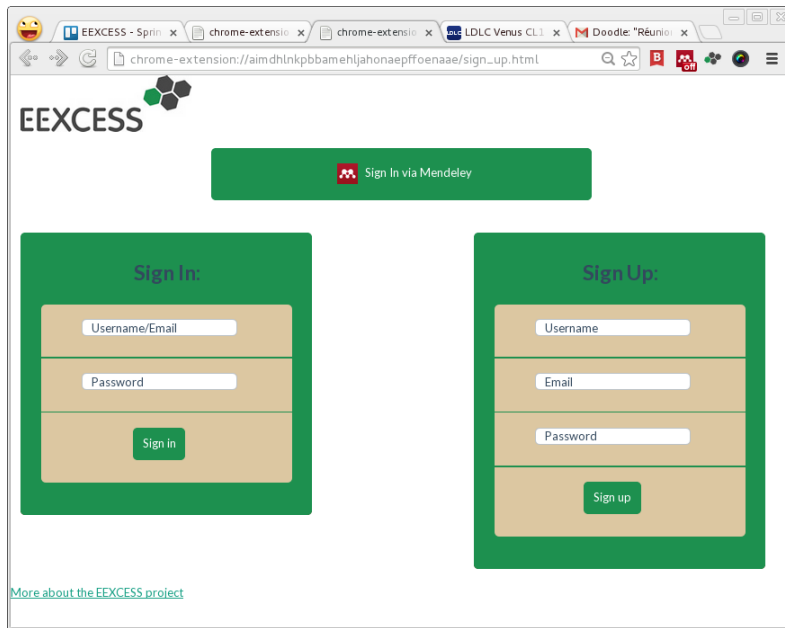
Lastname: Habegger
▶ Change lastname

Firstname: Benjamin
▶ Change firstname

Gender: Male
▶ Change gender

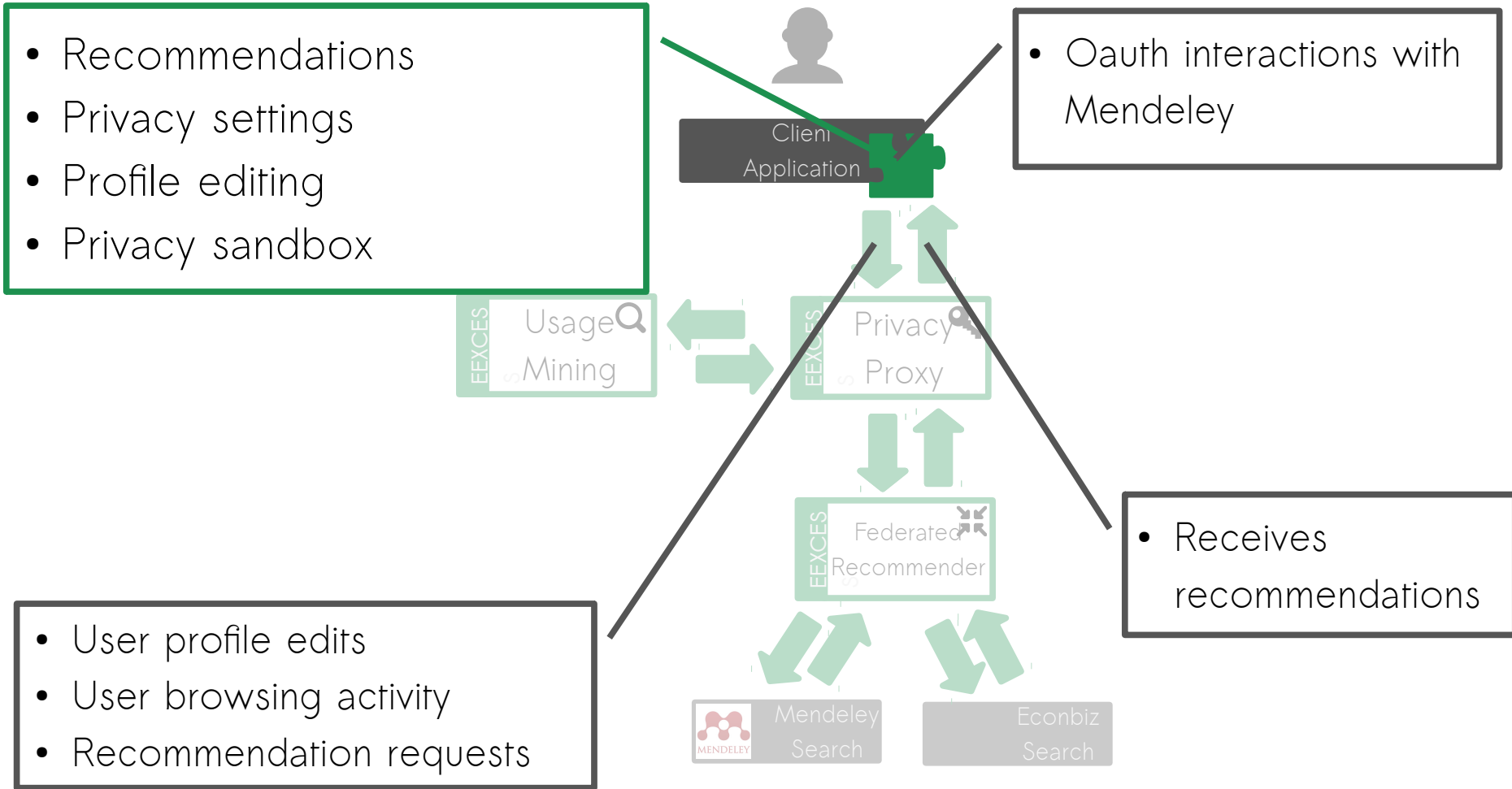
Birthdate: Not saved yet
▶ Change birthdate

Privacy Plugin - Profile data collection



- Mendely auth + profile import

Privacy Plugin



Privacy Plugin - Transparency

The screenshot displays the EEXCESS Privacy Plugin interface. At the top left is the EEXCESS logo, consisting of the word "EEXCESS" in a bold, sans-serif font next to a cluster of five hexagons in shades of green and black. Below the logo is a navigation bar with four tabs: "Traces" (highlighted in green), "Profile", "Privacy Sandbox", and "Settings".

The main content area is titled "141 traces dans l'historique". Below this title are two buttons: "Home" (with a house icon) and "Work" (with a briefcase icon). To the right of these buttons are two toggle switches, both currently turned "ON": "Show this session's traces" and "Show user's traces".

The history list contains three entries, each with a small icon on the left, a title, a URL, and a timestamp. The first entry has a document icon, the title "INDÉLICAT - Sur Instagram, le président tchétchène a des problèmes de...", and the URL "http://bigbrowser.blog.lemonde.fr". It shows two timestamps: "(20 September 2013 at 16:13:31)" and "(20 September 2013 at 16:13:34)", with a "Time spent: 3 seconds" indicator. A "Details" link is visible below the entry.

The second entry has a photo icon, the title "- Alpine sheep [Chamonix, France]", and the URL "http://herbertschroer.com". It shows two timestamps: "(20 September 2013 at 16:13:09)" and "(20 September 2013 at 16:13:31)", with a "Time spent: 22 seconds" indicator. A "Details" link is visible below the entry.

The third entry is partially visible at the bottom, showing the same document icon, title, and URL as the first entry.

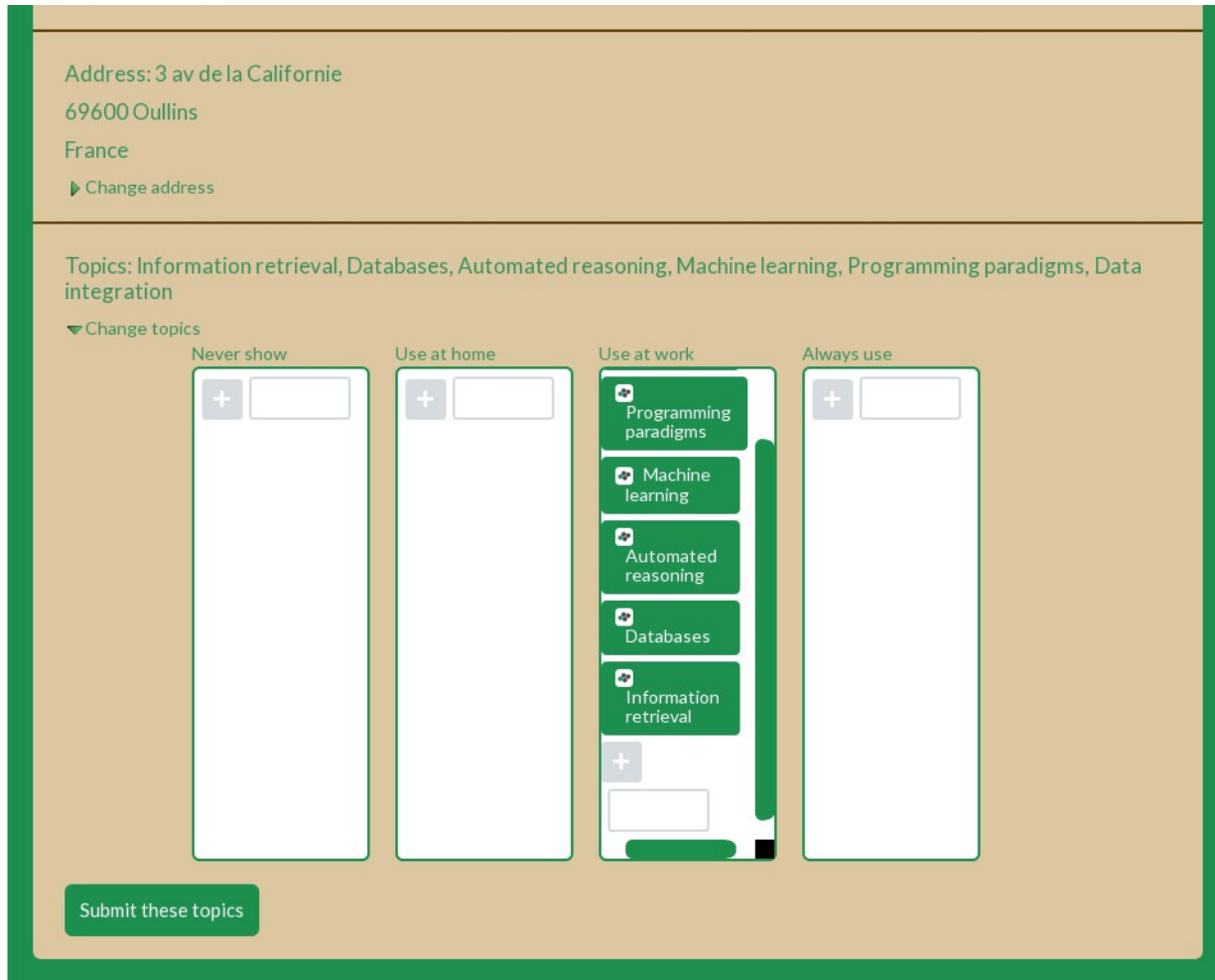
Privacy Plugin - Transparency

Address: 3 av de la Californie
69600 Oullins
France
▶ Change address

Topics: Information retrieval, Databases, Automated reasoning, Machine learning, Programming paradigms, Data integration
▼ Change topics

Never show	Use at home	Use at work	Always use
<input type="text"/>	<input type="text"/>	<input checked="" type="checkbox"/> Programming paradigms <input checked="" type="checkbox"/> Machine learning <input checked="" type="checkbox"/> Automated reasoning <input checked="" type="checkbox"/> Databases <input checked="" type="checkbox"/> Information retrieval <input type="text"/>	<input type="text"/>

Submit these topics

The image shows a user interface for a privacy plugin. At the top, there is an address field with the text "Address: 3 av de la Californie", "69600 Oullins", and "France". Below this is a "Change address" button. The main section is titled "Topics: Information retrieval, Databases, Automated reasoning, Machine learning, Programming paradigms, Data integration" and has a "Change topics" dropdown. There are four columns representing different usage contexts: "Never show", "Use at home", "Use at work", and "Always use". Each column has a list of topics with checkboxes. In the "Use at work" column, all five topics listed in the header are checked. At the bottom, there is a "Submit these topics" button.

Privacy Plugin - Control & Feedback

The screenshot displays the EEXCESS Privacy Sandbox interface. At the top, the EEXCESS logo is followed by navigation tabs: Traces, Profile, Privacy Sandbox (selected), and Settings. The main content is split into two panels. The left panel, titled 'Privacy Settings', features a search bar with a magnifying glass icon and a 'Submit this setting' button. Below this, there are three settings sections: 1) 'Traces to send : Only the current page will be sent' with a progress slider. 2) 'Email : benjamin.habegger@insa-lyon.fr' with a toggle switch labeled 'ON'. 3) 'Gender: nothing Male' with a toggle switch labeled 'OFF'. The bottom of this panel shows 'Birthdate : 36 years'. The right panel, titled 'Recommendations', has a progress bar at the top and a 'Results' section with a checkmark icon. It lists three recommendations: 1) A green card with an 'EB' icon, a title 'Migration vers OpenOffice.org sous environnement Windows :
analyse d'expériences internationales', and a score of 0.8213317. 2) A tan card with a red icon, a title 'Comparison of analytical and simulation planning models of seaport container terminals'. 3) A green card with an 'EB' icon, a title 'Impediments to the adoption of desktop Open Source Software (OSS) in a South African Provincial Government Department', and a score of 0.32045946.

EEXCESS

Traces Profile **Privacy Sandbox** Settings

Privacy Settings

Submit this setting

Traces to send : Only the current page will be sent

Your account traces on all computers will be sent

Email : benjamin.habegger@insa-lyon.fr
benjamin.habegger@insa-lyon.fr **ON**

Gender: nothing
Male **OFF**

Birthdate : 36 years

Recommendations

Results ✓

EB
[Migration vers OpenOffice.org sous environnement Windows :
analyse d'expériences internationales](#)
Score : 0.8213317

Comparison of analytical and simulation planning models of seaport container terminals

EB
[Impediments to the adoption of desktop Open Source Software \(OSS\) in a South African Provincial Government Department](#)
Score : 0.32045946

Privacy Plugin - Control & Feedback

The screenshot displays a privacy plugin interface with two main sections: user controls on the left and search results on the right.

User Controls (Left Panel):

- Email:** benjamin.habegger@insa-lyon.fr. A toggle switch is set to **ON**.
- Gender:** nothing. A toggle switch is set to **OFF**.
- Birthdate:** 36 years (1977-08-04). A slider control is positioned at 36 years.
- Address:** France. A slider control is positioned at France.
- Geolocation:** Villeurbanne. A slider control is positioned at Villeurbanne.

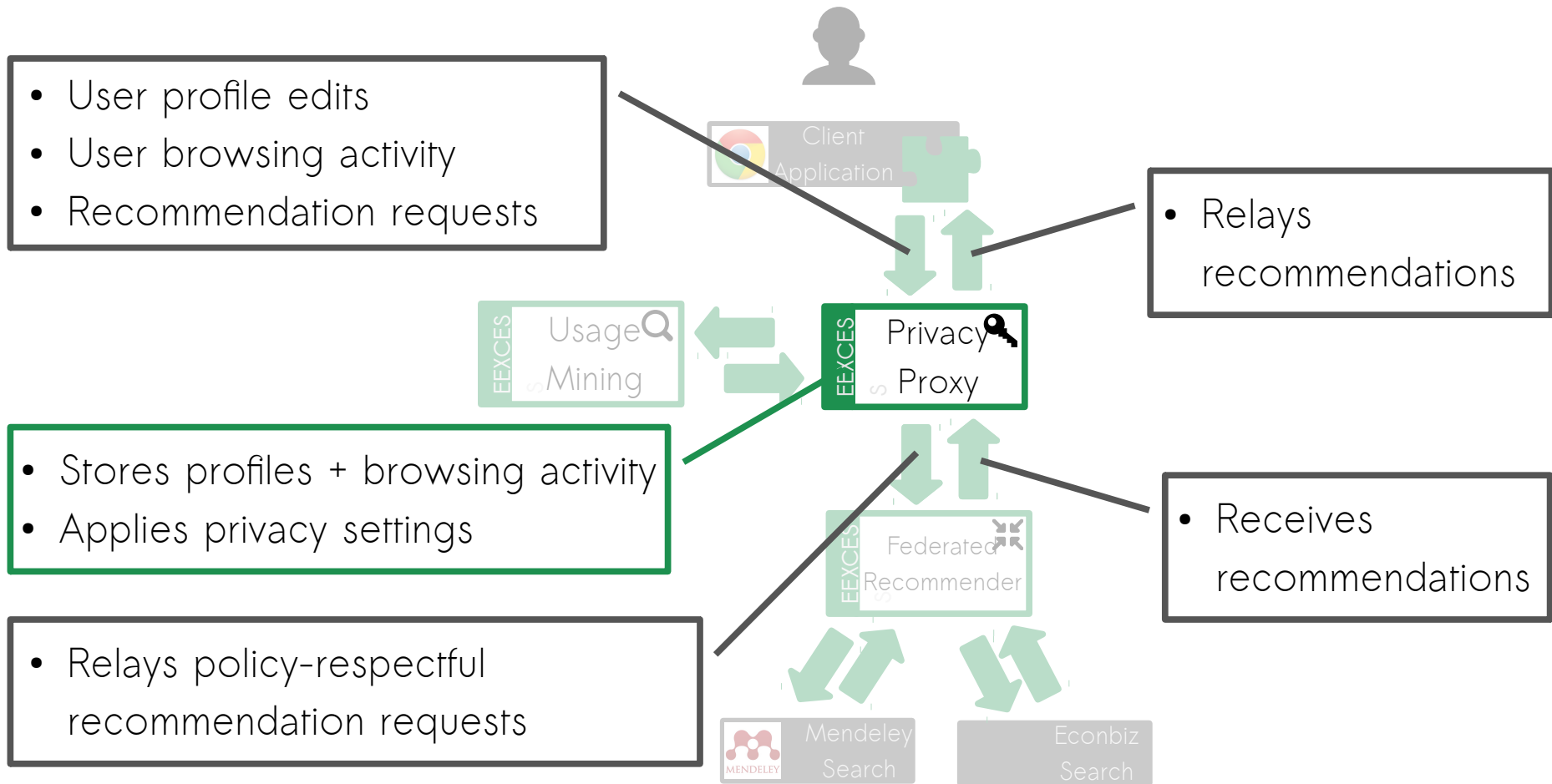
Search Results (Right Panel):

- Item 1:** [Comparison of analytical and simulation planning models of seaport container terminals](#)
- Item 2:** [Impediments to the adoption of desktop Open Source Software \(OSS\) in a South African Provincial Government Department](#)
Score : 0.32045946
- Item 3:** [Stretch blow molding method of plastic bottle](#)

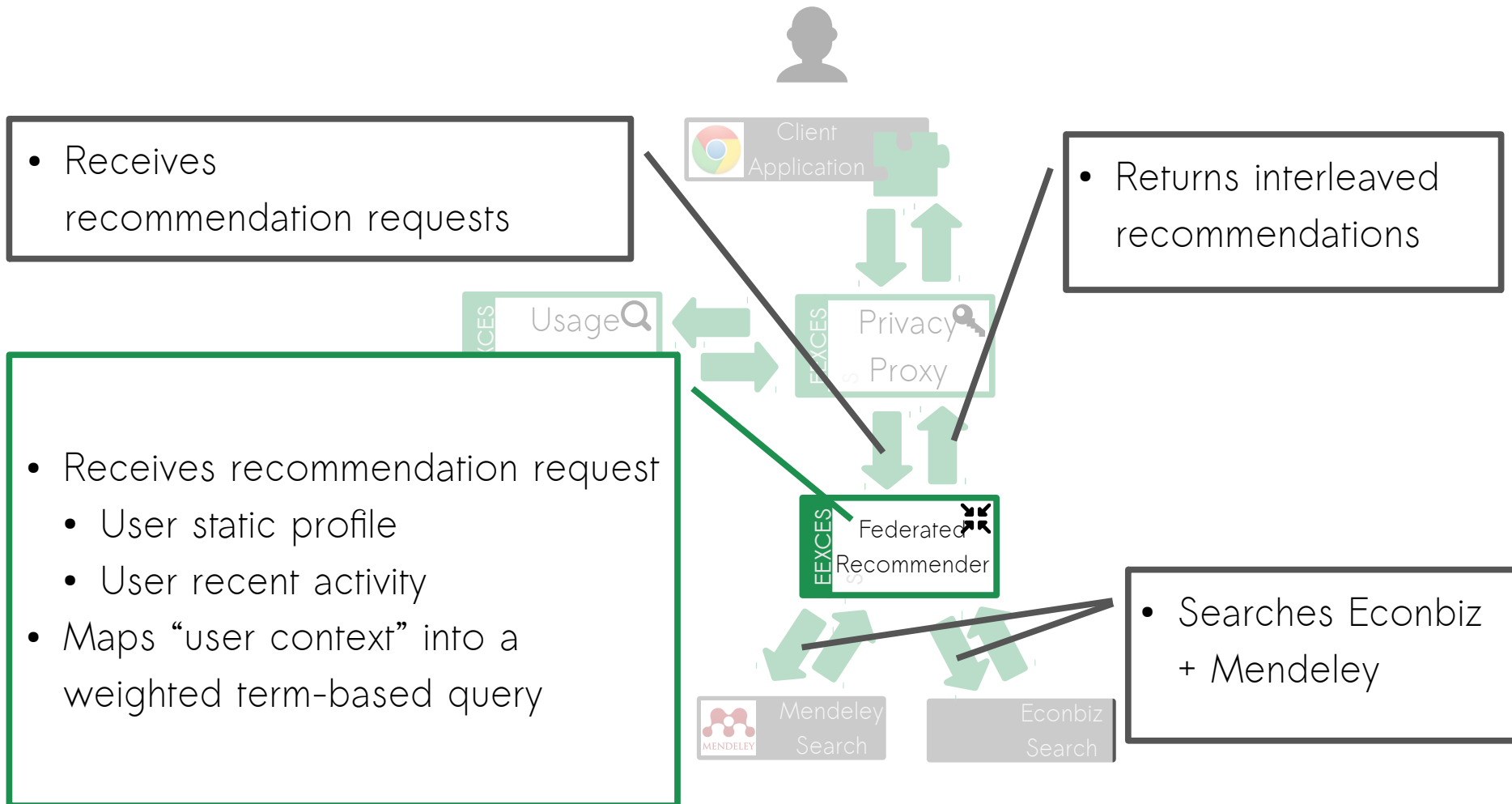
Query Results:

Query	Count
draw	10
equivalent	10
forum	10
container	10
inside	10
mask	10
enoffice.org	10
place	10
clipping	10
ion retrieval	1
Databases	1
ng paradigms	1
a integration	1
nine learning	1
ed reasoning	1

Privacy Proxy - Protection



Basic Federated Recommender



EEXCESS - Summary

- **Quality**
 - Recommendations adapted to the user
- **Preserving privacy**
 - Guaranteeing respect of privacy policy
- **Scalability**
 - Ensuring the whole system scales
- **Quantifiable measures**

EEXCESS - Future work

- Privacy-preserving search
- Measuring
 - recommendation quality
 - privacy impacts of disclosure
- Impacts of trustworthiness in peers

Contact



benjamin.habegger@insa-lyon.fr



<http://www.linkedin.com/in/benjaminhabegger>



@b_habegger

