

EMEARC18

OC LC EMEA
REGIONAL COUNCIL
MEETING

Hello.

I'm the **Smarter** Library.



EzPAARSE and ezMASURE :

Assembling national dashboards from locally generated and fine-grained access events to electronic resources

DOMINIQUE LECHAUDEL - INIST-CNRS

THOMAS JOUNEAU - UNIVERSITE DE LORRAINE



Dominique Lechaudel

Product owner of ezMÉSURE project at INST-CNRS.
The CNRS, the French Center for Scientific Research



Thomas Jouneau

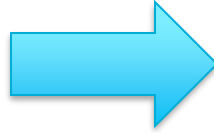
E-librarian @ Université de Lorraine

EzPAARSE and ezMESURE user

Co-animator of the Couperin.org « Indicators » WG, member of the Project COUNTER Executive Committee

EzPAARSE and ezMESURE :
Assembling national dashboards from
locally generated and fine-grained
access events to electronic resources

ezPAARSE and ezMESURE



ezMESURE
AGRÉGEZ COMPAREZ VISUALISEZ VALORISEZ

ezPAARSE

The free and open source software produces **uniform** electronic resources **usage data**

ezMESURE

Our **national** repository and **dashboard tool** to visualize ezpaarse collected data

<http://ezpaarse.org>

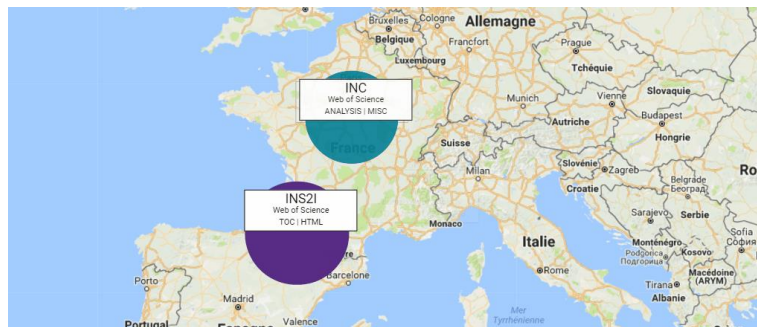
What is ezPAARSE ?

- Free open source software
- Specialized niche software
- Locally installed by institutions
- Produces uniform electronic resources usage data

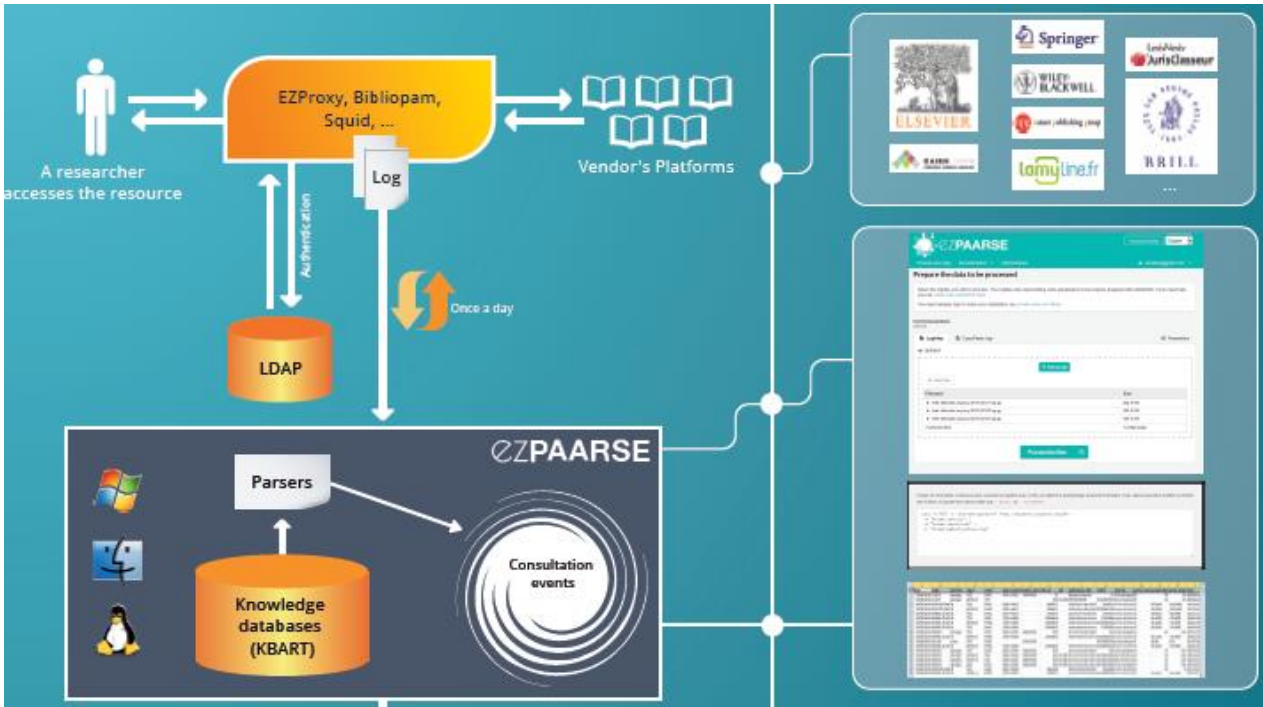


BiblioMap shows
ezPAARSE running live

<http://bibliomap.inist.fr>



What is ezPAARSE ?



Log files



- Filter
- Identify
- Enrich
- Encrypt
- ...

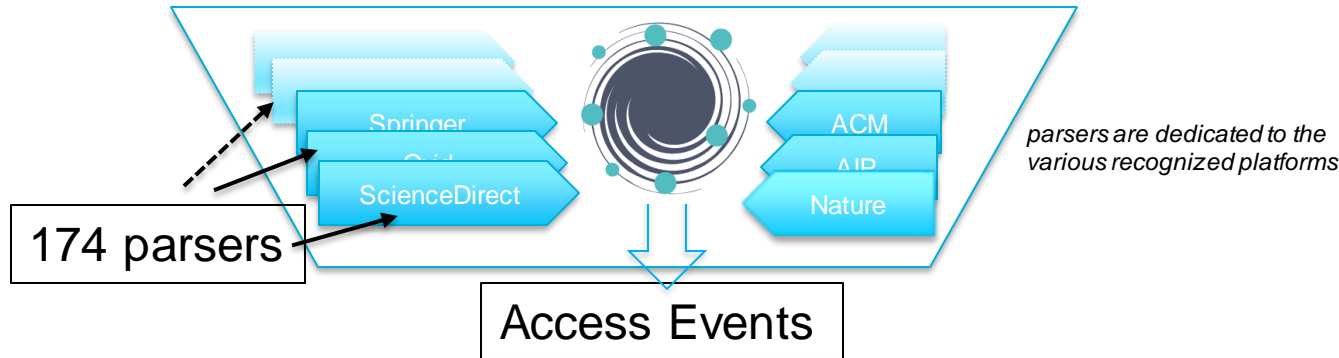


Access Events files

How does ezPAARSE work ?

```
107.206.236.51 - - - [31/Dec/2017:23:06:56] "GET http://insb.bib.cnrs.fr:80/login?url=http://www.sciencedirect.com/science/journal/13698486 HTTP/1.1" 302 0 -
107.206.236.51 OvJrZ - user@biologie.ens.fr_O_CNRS_I_DS53_OU_UMR8197 [31/Dec/2017:23:07:28] " HTTP/1.1 GET https://www.ncbi.nlm.nih.gov:443/pmc/articles/PMC5511345/pdf/ncomms16088.pdf HTTP/1.1 " 200 108723 insb
92.91.207.211 jjGjY9Q - 16SBIUMR7255_O_CNRS_I_DS53_OU_UMR7255 [01/Jan/2018:02:44:07] " GET https://cdn.els-cdn.com:443/sd/css/css_gen_v01_1712R2.css" 200 3410106 insb
83.221.104.173 3aVq1a - greg@ipmc.cnrs.fr_O_CNRS_I_DS53_OU_UMR7275 [01/Jan/2018:05:02:38] "GET http://www.physiology.org:80/doi/pdf/10.1152/ajplung.00348.2002 HTTP/1.1" 200 612360 insb
91.140.193.126 NZBUs1 - bob@unistra.fr_O_OTHER_I_DS53_OU_UMR7213 [01/Jan/2018:06:53:18] "GET http://www.sciencedirect.com:80/science/article/pii/S0009308416300159/pdf?pid=1-s2.0-S0009308416300159-main.pdf HTTP/1.1" 200 7890 insb
```

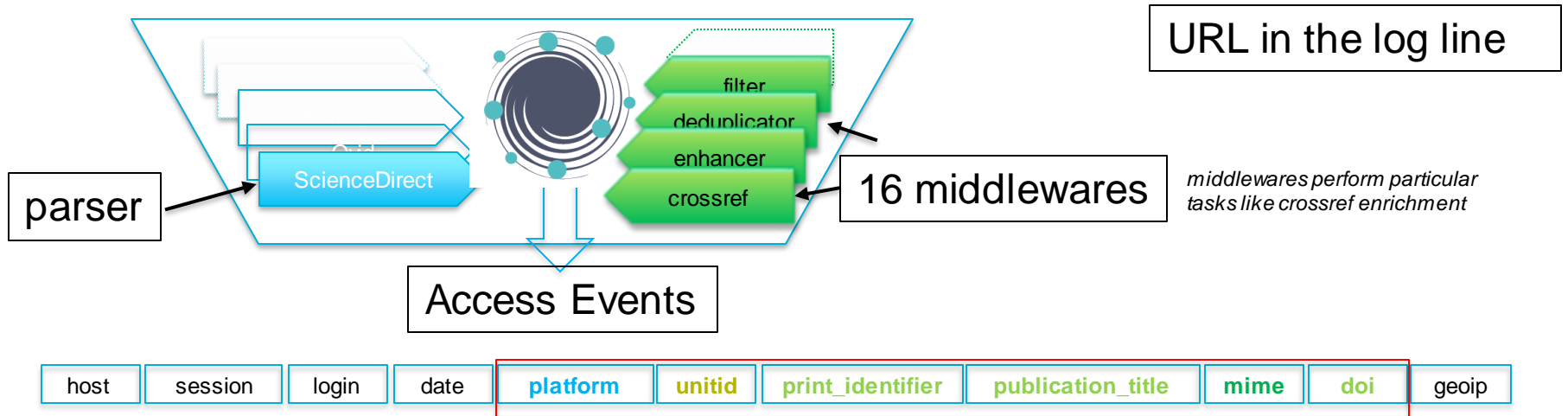
log lines



host	session	login	date	platform	unitid	print_identifier	publication_title	mime	doi	geoip
------	---------	-------	------	----------	--------	------------------	-------------------	------	-----	-------

How does ezPAARSE work ?

<http://www.sciencedirect.com:80/science/article/pii/S0009308416300159/pdf?pid=1-s2.0-S0009308416300159-main.pdf> HTTP/1.1



Example of an ezPAARSE output

COUNTER
CONSISTENT CREDIBLE COMPARABLE

Deduplicated access events
COUNTER recommendation

KBart fields

Geop fields

A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
date	login	platform	rtype	mime	print_identif	online_ident	id	doi	publication_title	unitid	domain	geop-counb	geop-lattus	geop-longit	host
01/04/2014	JULIET	springer	TOC	MISC	1015-1621	1420-9055		27	Aquatic Sciences		27 link.springer BR	-10		-55	189.84.26
01/04/2014	JULIET	springer	ARTICLE	PDF				364	10.1007/BF00878399		10.1007/BF0 download.sp BR	-10		-55	189.84.26
01/04/2014	KEATON.EMI	sd	TOC	MISC	0195-6671			1956671	Cretaceous Researcl		1956671 www.scienc US	34.0194		-118.4912	74.123.14
01/04/2014	KEATON.EMI	sd	ARTICLE	HTML	0195-6671			1956671	Cretaceous Researcl		S0195667103 www.scienc US	34.0194		-118.4912	74.123.14
01/04/2014	MABEL.OLSC	sd	TOC	MISC	0377-0273			3770273	Journal of Volcanolc		3770273 www.scienc US	39.9612		-82.9988	33.11.15.
01/04/2014	MABEL.OLSC	sd	TOC	MISC	1750-5836			17505836	International Journa		17505836 www.scienc US	42.2238		-72.6409	24.62.194
01/04/2014	MABEL.OLSC	sd	ARTICLE	HTML	1750-5836			17505836	International Journa		S1750583610 www.scienc US	42.2238		-72.6409	24.62.194
01/04/2014	MABEL.OLSC	sd	TOC	MISC	1750-5836			17505836	International Journa		17505836 www.scienc US	42.2238		-72.6409	24.62.194
01/04/2014	GRADY	springer	TOC	MISC	0943-0105	1866-6299		254	Environmental Geol		254 link.springer AU	-27		133	120.22.15
01/04/2014	MABEL.OLSC	sd	ARTICLE	HTML	1750-5836			17505836	International Journa		S1750583613 www.scienc US	42.2238		-72.6409	24.62.194
01/04/2014	VELVA	wiley	TOC	MISC		2169-9100			10.1002/(ISS)online		10.1002/(ISS)online library FR	48.86		2.35	92.187.44
01/04/2014	MABEL.OLSC	sd	ARTICLE	HTML	1750-5836			17505836	International Journa		S1750583614 www.scienc US	42.2238		-72.6409	24.62.194
01/04/2014	GRADY	springer	TOC	MISC	0943-0105	1866-6299		254	Environmental Geol		254 link.springer AU	-27		133	120.22.15
01/04/2014	GRADY	springer	ARTICLE	PDF	0943-0105	1866-6299		254	10.100 Environmental Geol		10.1007/s002 download.sp AU	-27		133	120.22.15
01/04/2014	GRADY	springer	ARTICLE	PDF	0943-0105	1866-6299		254	10.100 Environmental Geol		10.1007/s002 download.sp AU	-27		133	120.22.15
01/04/2014	GRADY	springer	ABS	MISC	0943-0105	1866-6299		254	10.100 Environmental Geol		10.1007/s002 link.springer AU	-27		133	120.22.15
01/04/2014	KEATON.EMI	sd	TOC	MISC	0301-9268			3019268	Precambrian Resear		3019268 www.scienc EU	47		8	151.175.8
01/04/2014	MABEL.OLSC	sd	ARTICLE	HTML	0377-0273			3770273	Journal of Volcanolc		S0377027314 www.scienc US	39.9612		-82.9988	33.11.15.

Text file
(CSV or JSON format)

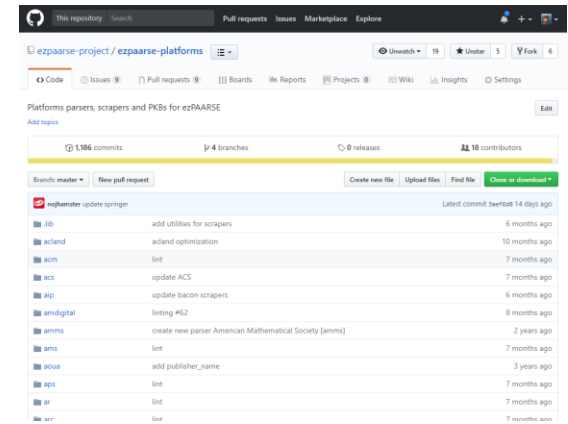
Librarian and computer scientist collaborate to produce parsers



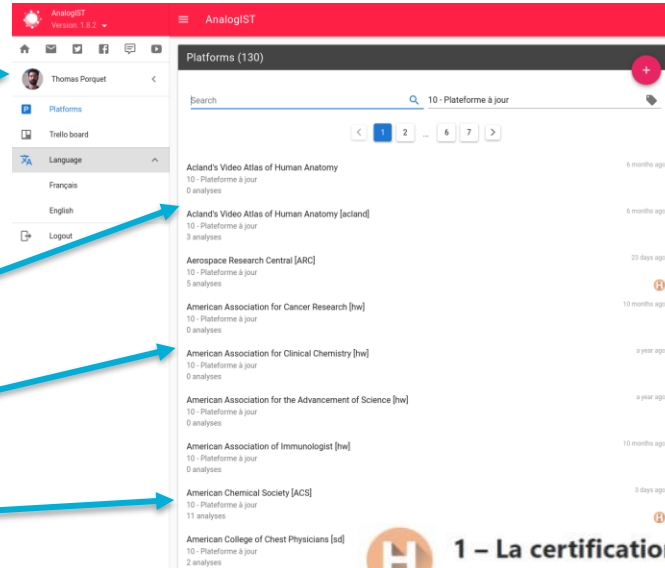
<http://analyses.ezparse.org>



174 parsers



How does ezPAARSE works ?



<http://analyses.ezpaarse.org>

Maintaining and expanding ezPAARSE recognition capacity to a new platform is a collaborative work



Certification Humaine

1 – La certification H (human)

Signifie qu'un contributeur a vérifié que les résultats sous forme d'EC des consultations de la plateforme reconnues par ezPAARSE correspondent à ses consultations réelles.



Certification Editeur

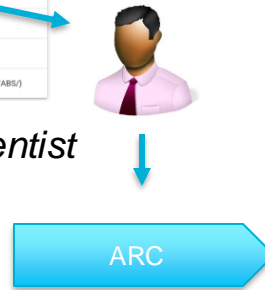
2 – La certification P (publisher)

Signifie qu'un contributeur a comparé le nombre de consultations fournis par les éditeurs avec ceux obtenus avec ezPAARSE et qu'il n'y a pas de différences significatives (<5%).

How does ezPAARSE works ?

Value	Comment
doi	indique que l'on a plus l'identifiant de la revue mais un DOI
pdfplus	format pdfplus
10.2514	indique la premiere partie DOI qui est le même pour toute la plateforme ARC
1.C001500	deuxieme partie du DOI et identifiant article (attention même numero pour abstract seul change le prefixe /PDF/ à la place de /ABS/)

Computer scientist



<http://analyses.ezpaarse.org>

Maintaining and expanding ezPAARSE recognition capacity to a new platform is a collaborative work

ezPAARSE installation - usage

Very easy use :

- 5 minutes for installation from GitHub
- 5 minutes for its configuration
- Fully automatable treatments
- Fully automatable updates
- Web interface
- Command line

- docker container available

Installation quickstart

If you are a Windows user, you can install ezPAARSE on your computer as a docker image. Please refer to the [docker section](#) below.

To install the latest stable version of ezPAARSE on a Unix-type system, open a terminal and type:

```
git clone https://github.com/ezpaarse-project/ezpaarse.git
cd ezpaarse
git checkout `git describe --tags --abbrev=0`
make
```



If you want to install the version in development (unstable), open a terminal and type:

```
git clone https://github.com/ezpaarse-project/ezpaarse.git
cd ezpaarse
make
```

The screenshot shows the GitHub repository page for ezpaarseproject/ezpaarse. The repository is public and has an automated build. The last push was 3 months ago. The page includes a search bar, navigation links (Explore, Help, Sign up, Sign in), and a section for the repository's description and Dockerfile. The Dockerfile section shows the Docker Pull Command: `docker pull ezpaarseproject/ezpaarse`.

Use with docker

ezPAARSE is available as a [docker image](#).

You need:

- Docker (Version >= 1.12)
- Docker Compose (Version >= 1.7)

Then, you can run the dockerized ezpaarse this way:

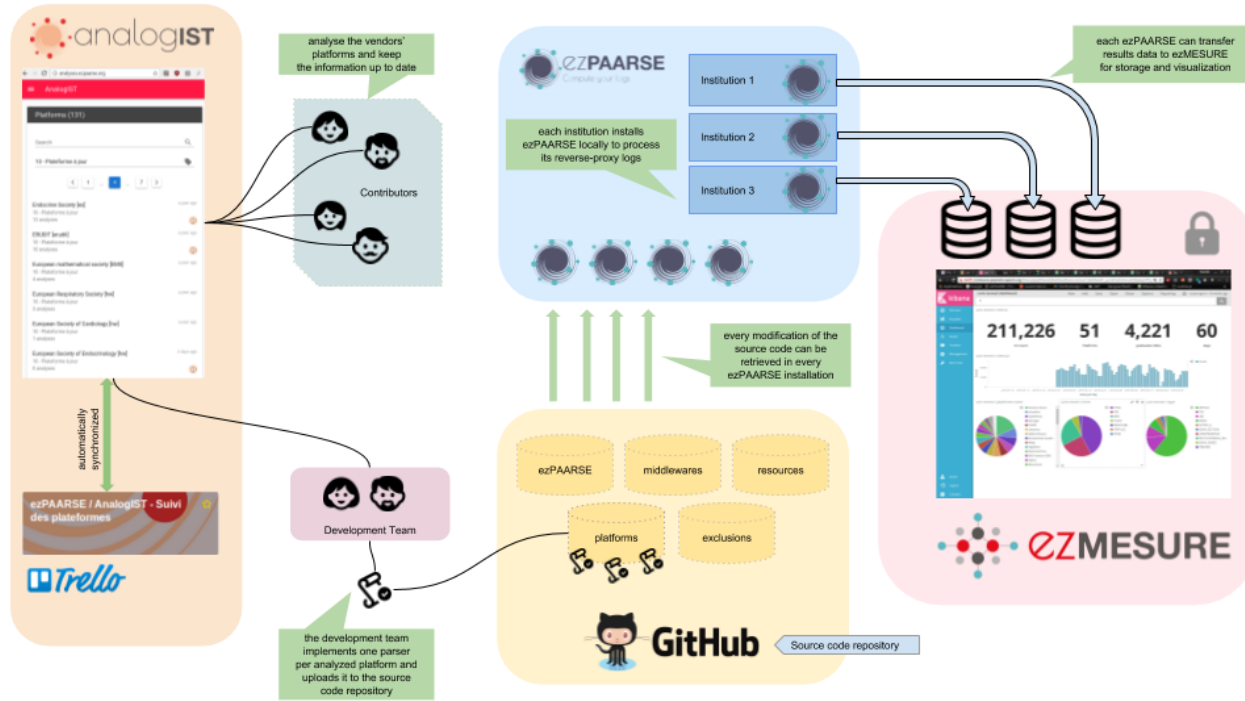
```
mkdir ezpaarse/
wget --no-check-certificate https://raw.githubusercontent.com/ezpaarse-project/ezpaarse/#
docker-compose up -d
```

Then ezpaarse is available at this URL: <http://127.0.0.1:59599>

ezPAARSE worldwide

- We target ~**130 french institutions** (with a majority of universities) declaring using a reverse-proxy
- **80** of them explicitly declared being interested
- **40** have a proper logformat parameter defined and tested at least a log sample
- **50** have installed and use ezPAARSE on a regular basis
- ~**120** instance **installations out of France**
- **60** in the USA

ezPAARSE / ezMESURE Ecosystem



What is ezMESURE ?

Universities Research Institutes Grandes Écoles Others

Each ezPAARSE can upload data results to ezMESURE to be visualized and archived.

Each institution can have access to a dynamic and tweakable dashboard.

The ezPAARSE data which comes from the HER institutions is indexed and archived by ezMESURE.

Tableau de bord agrégé des organismes :

7,699,144 116 46,427 649 22

2,075,714 275 1,168

Physical Science Lectures



Access Events files



- Stock
- Aggregate
- Compare
- Visualize
- Highlight
- ...

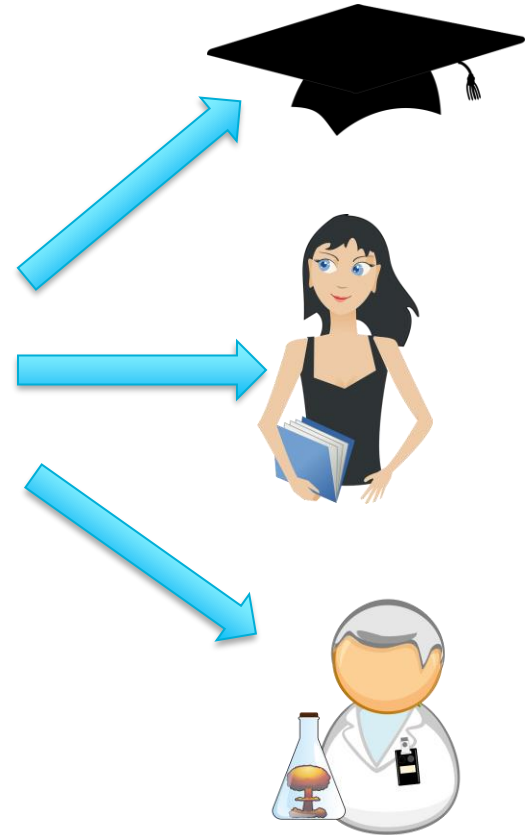
Dashboards

Université de Lorraine - Pioneer times : Designing, installing, testing, strengthening a local ezPAARSE installation



Using ezPAARSE

- The key question : how to merge ezPAARSE output with valid, local data such as : the cursus, the research lab, the user's status
- I'll expose how we did it at the Université de Lorraine, in 2 different ways (one currently active, the other we wish to implement more completely)



The context, the goals

Context : how we became a pilot institution

- University born from a merger in 2012 but we use ezproxy since 2009
- **All accesses** (distant AND local) go through the reverse proxy
- A geographical proximity with the Inist, making collaborations easier

Goals

- Supplement the publisher statistics with data regarding the non-Counter platforms
- Deepen publisher statistics with user profile data
- Use the data and indicators produced as a steering tool (documentary policy, service delivery)

Resources

- 1 librarian, 1 technician (part-time for both)

Brief chronology (UL – Inist – Couperin)

2012

Product vision
Version 0.1

2013

First experiments at U. de Lorraine
First dashboards!

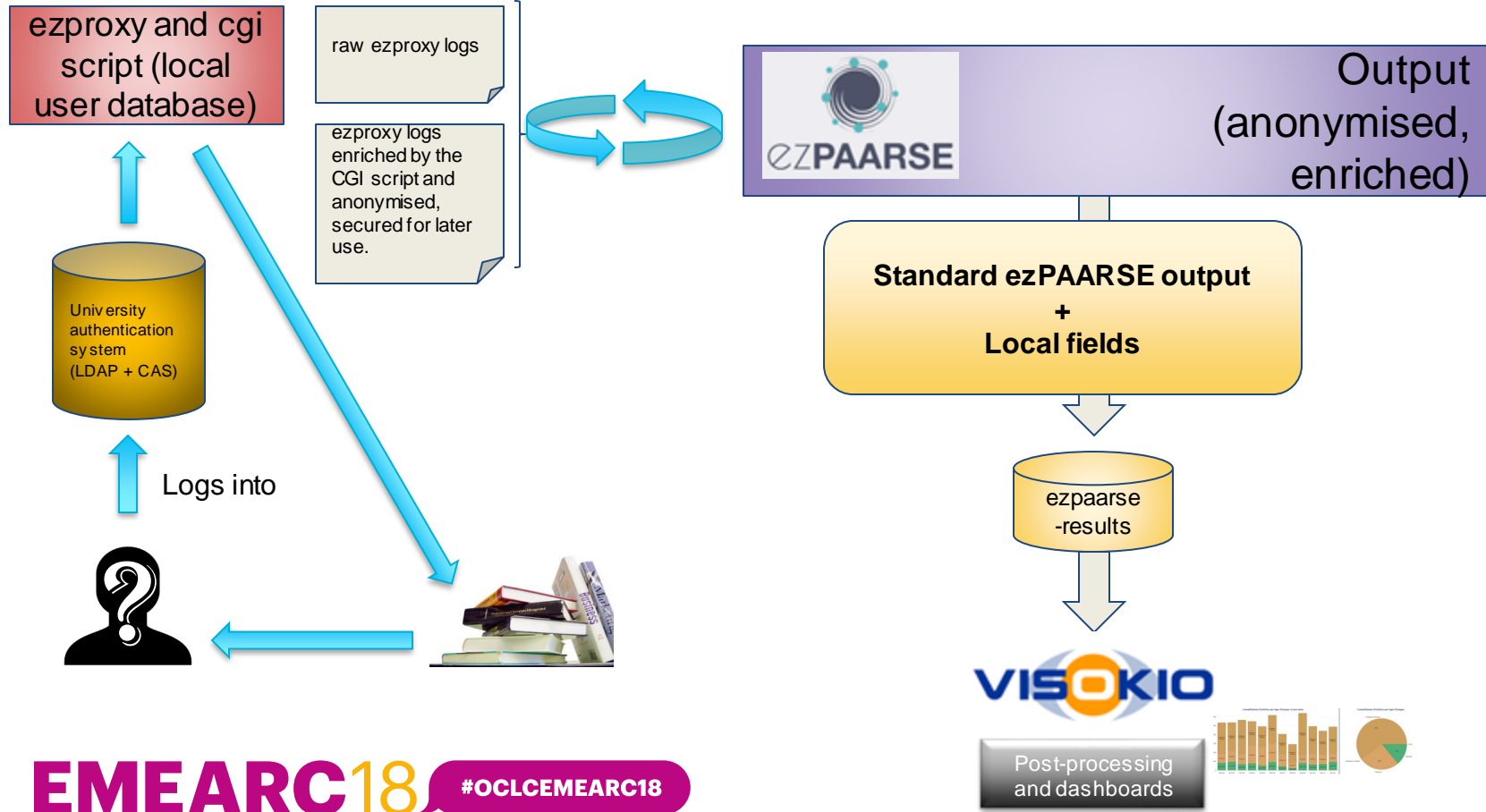
2014

Consolidation of the UL installation
Results files accessed through a web interface

2015-2017

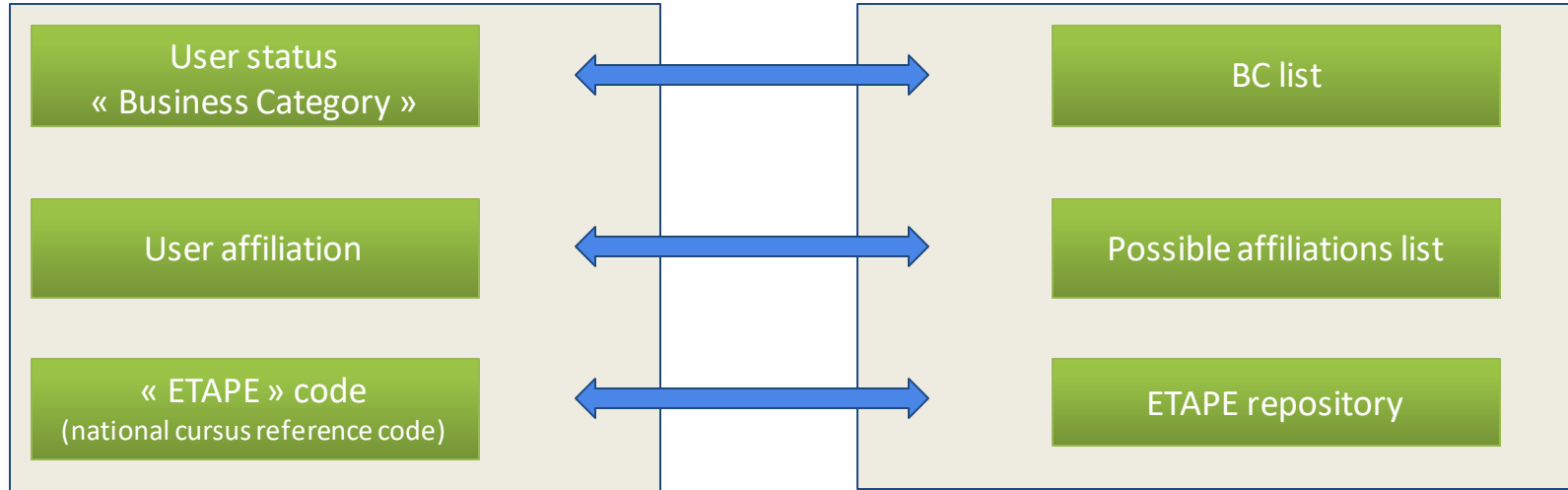
AGIMUS project
New way to characterize the events.

The main installation framework since 2014



Select local relevant fields

Available data...



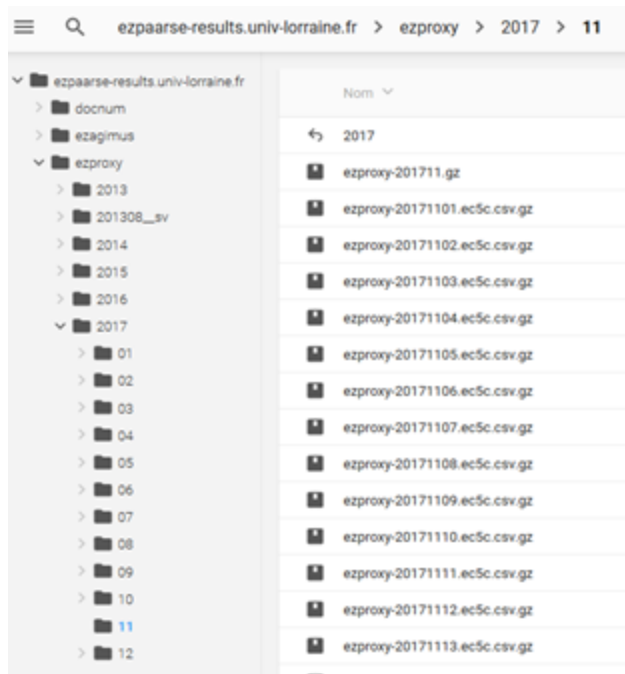
... are translated with the help of static lists during the post-processing phase (Visokio).

→ Early contacts with IT teams have allowed us to obtain a regular access to the BC and affiliation lists (and their updates)

Standard VS local fields

Standard ezPAARSE fields	date	
	login	
	platform	
	platform_name	
	publisher_name	
	rtype	
	mime	
	print_identifier	
	online_identifier	
	title_id	
	doi	
	publication_title	
	unitid	
	domain	
	datetime	
	doi-publication-title	
	doi-publication-date-year	
	doi-publisher	
	doi-type	
	doi-ISSN	
	doi-subject	
	doi-license-content-version	
	doi-license-URL	
	geop-country	
	geop-latitude	
	geop-longitude	
	host	
session		
url		
status		
size		
User related fields	user	
	BCPrin	
	ABC	
	A#Prin	
	AA#	
	cas	

Access to the output files



A server internally accessible with a nice web-based interface allows to retrieve daily ezPAARSE outputs, and monthly concatenations

1 line is 1 consultation event

TOC (Table of contents) et ABS (tracts) excluded :

2014 : More than 1 700 000 ECs

2015 : More than 2 000 000 ECs

2016 : More than 3 000 000 ECs

2017 : More than 4 500 000 ECs

...

Size of output of an active month once compressed :

12 Mb in 2014

50 Mb in 2017.

Post-processing(1) : Research and academic units

Academic departments and « Collegiums »

Arts, lettres et langues (ALL)
Droit, économie, gestion (DEG)
Lorraine – INP (écoles d'ingénieurs)
Lorraine Management Innovation (LMI)
Interface
Santé
Sciences et technologies
Sciences humaines et sociales (SHS)
Technologie

Research : 62 possible research units in 10 « Poles »

A2F : Agronomie, agroalimentaire, forêt
BMS : Biologie, médecine, santé
CPM : Chimie et physique moléculaires
M4 : Matière, matériaux, métallurgie, mécanique
TELL : Temps, espaces, lettres, langues
AM2I : Automatique, mathématiques, informatique et leurs interactions
CLCS : Connaissance, langage, communication, sociétés
EMPP : Énergie, mécanique, procédés, produits
OTELo : Observatoire Terre et environnement de Lorraine
SJPEG : Sciences Juridiques, Politiques, Économiques et de Gestion



There are more than 1000 possible affiliations

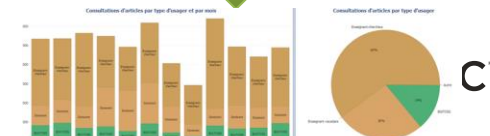
Each patron can have up to 4 affiliations, and very often 2



For every CE we take the first affiliation of each kind : Teaching, Research, Administration.

We add extra information regarding each of those affiliations (the « real » name, the group in the organigram...)

The « teaching » affiliations are attached with their respective collegiums and the « research » with their poles.



Post-processing(2) : Business categories

Catégorie	Intitulé log	BC_Nomcourt	BC_Regroupement
<u>AT</u>	<u>CatAT</u>	Agent en attente	Autre
<u>CB</u>	<u>CatCB</u>	Contractuel BIATOSS	BIATOSS
<u>CE</u>	<u>CatCE</u>	Contractuel enseignant/chercheur	Enseignant-chercheur
<u>DC</u>	<u>CatDC</u>	Doctorants inscrits	Doctorant
<u>ED</u>	<u>CatED</u>	Etudiants doctorants qui sont en même temps personnels	Doctorant
<u>EH</u>	<u>CatEH</u>	Etudiant inscrit en doctorat	Doctorant
<u>EN</u>	<u>CatEN</u>	Étudiant non inscrit à une formation diplômante	<u>Etudiant</u>
<u>ET</u>	<u>CatET</u>	Etudiant inscrit à une formation diplômante hors doctorat	<u>Etudiant</u>
<u>FB</u>	<u>CatFB</u>	Fonctionnaire BIATOSS	BIATOSS
<u>FE</u>	<u>CatFE</u>	Fonctionnaire enseignant/chercheur	Enseignant-chercheur
<u>HB</u>	<u>CatHB</u>	Hébergé non-enseignant	BIATOSS
<u>HC</u>	<u>CatHC</u>	Hébergé contractuel	BIATOSS
<u>HE</u>	<u>CatHE</u>	Hébergé enseignant/chercheur	Enseignant-chercheur
<u>HM</u>	<u>CatHM</u>	Professeur ou chercheur émérite	Enseignant-chercheur
<u>IT</u>	<u>CatIT</u>	Invité temporaire	Autre
<u>PE</u>	<u>CatPE</u>	Personnalité extérieure	Autre
<u>RE</u>	<u>CatRE</u>	Retraité	Autre
<u>ST</u>	<u>CatST</u>	Stagiaire	Autre
<u>VA</u>	<u>CatVA</u>	Vacataire administratif ou technique	BIATOSS
<u>VE</u>	<u>CatVE</u>	Vacataire d'enseignement	Enseignant vacataire

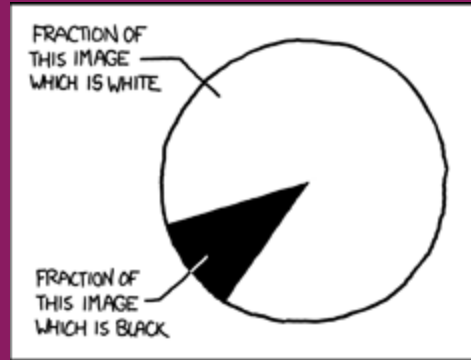


Participation to AnalogIST

The screenshot displays the AnalogIST web application interface. The top navigation bar is red and contains the AnalogIST logo, version information (1.9.7), and a hamburger menu. Below the navigation bar is a sidebar with social media icons and a user profile section showing 'Non connecté'. The main content area is titled 'Plateformes (366)' and features a search bar and a 'Statut' filter. A pagination control shows page 1 of 19. The main list contains two entries:

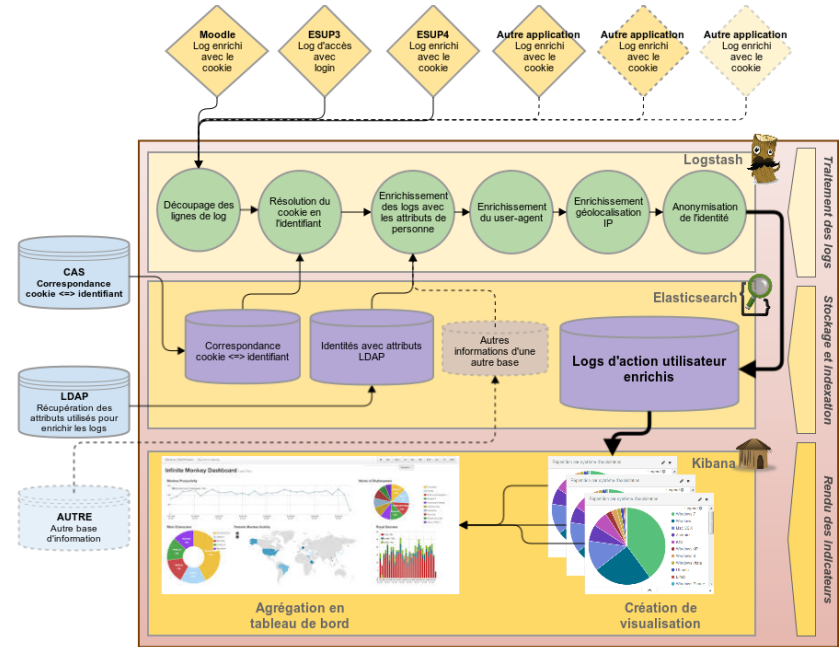
Plateforme	Statut
Acta Sanctorum [AS]	3 - Analyse des URLs terminée 7 analyses
20th Century German History Online [dgo]	12 - Plateforme au "congélateur" 0 analyses

Université de Lorraine - Interactive dashboards : ezMEASURE, AGIMUS (2016 and beyond)

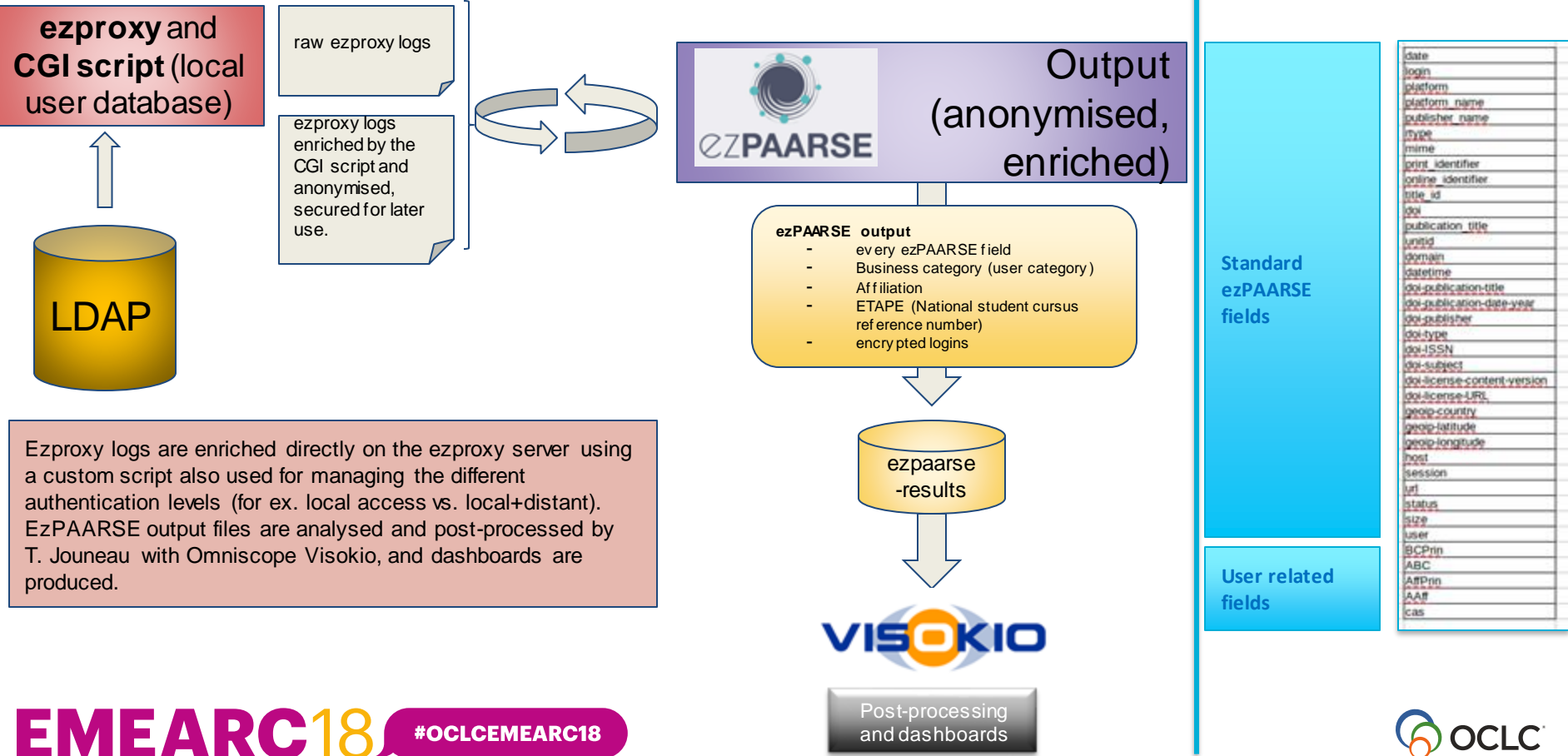


ezPAARSE and AGIMUS

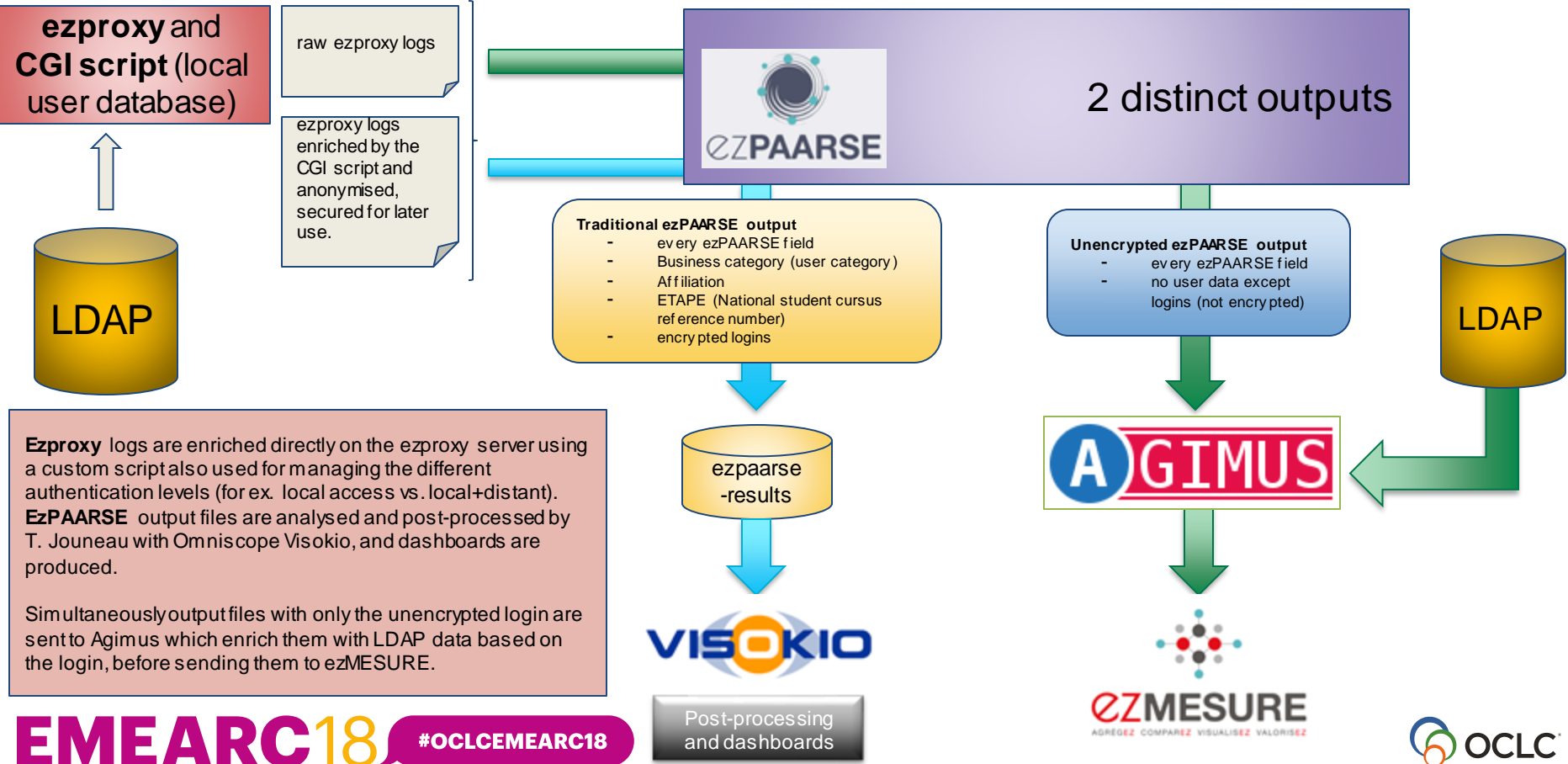
Agimus is a national project used by IT teams in some universities. It works very similarly to ezPAARSE but with every other electronic service (Intranet, wi-fi accesses, Moodles...)



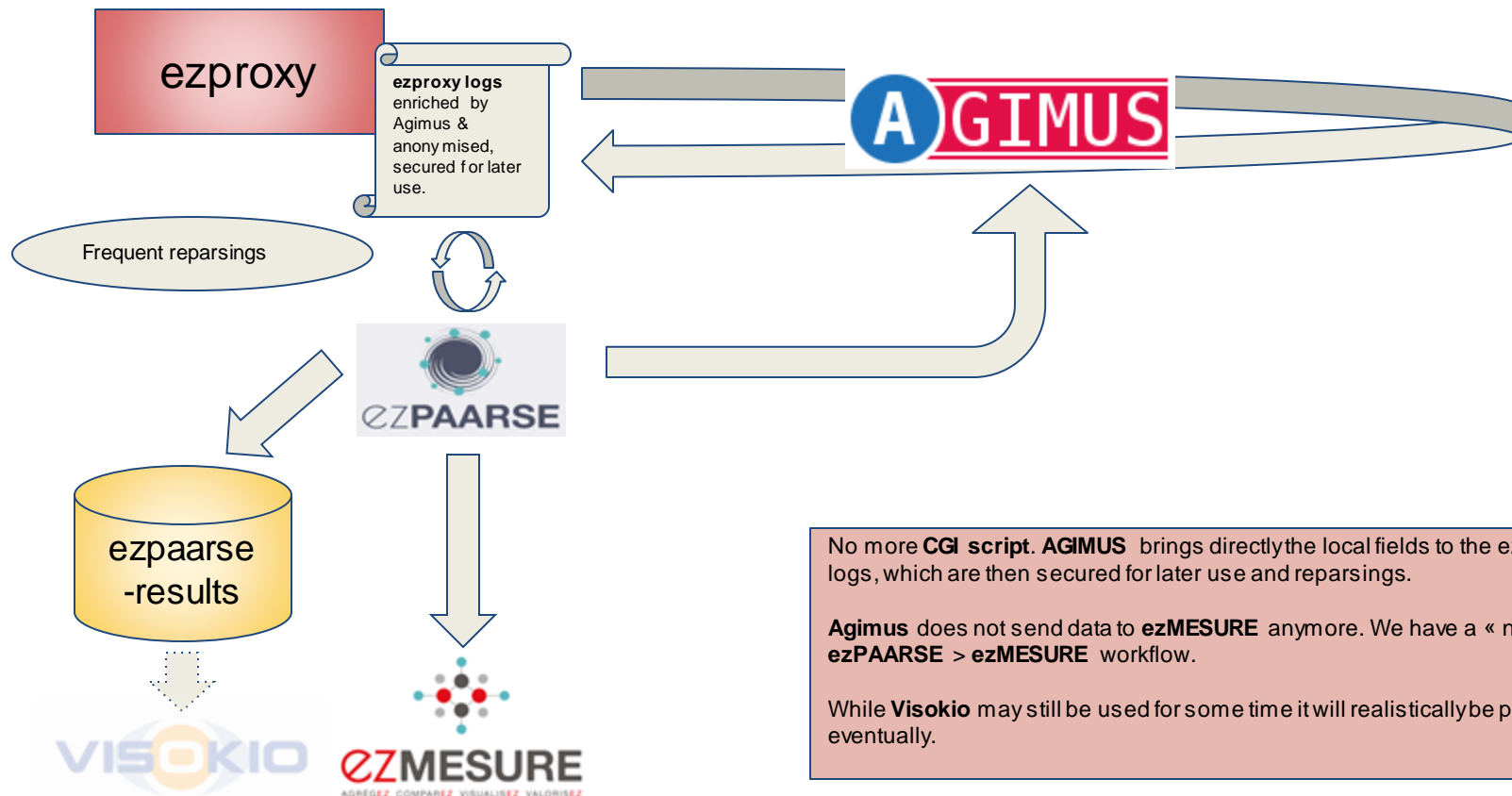
Past (2014-2017) : The main installation framework (quick reminder)



Present (2017) : Workflow ezProxy → ezPAARSE → Agimus → ezMESURE



Future (2018?) : Redesigning the workflow

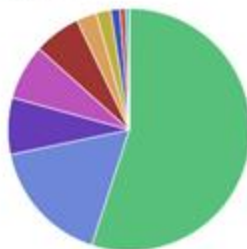


Ex. 3 (ezMESURE) : platform profiles

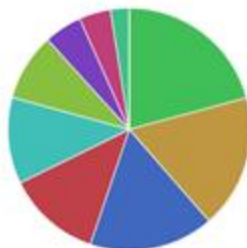
TJ - univ-lorraine-ezagimus - table-plateforme

Plate-forme *	ECs	Utilisateurs uniques
American Chemical Society	15,496	886
American Institute of Physics	2,202	374
American Mathematical Society	9,004	87
American Physical Society	6,176	281
Annual Reviews	136	44
Association for Computing Machinery	2,341	158
Brepols	1,308	20
Brill	1,068	72
CAIRN	87,728	4,782
Cambridge University Press	1,395	379
Cyberlibris	2,141	413
Daloz	88,113	2,553
Daloz Revues	597	88
Dawsonera	1,090	278
De Gruyter	325	123
Docnum	794	336
Doctrinal Plus	7,264	518
Early English Books Online	49	6
EBSCO Discovery Service	532,259	18,834
Editions Francis Lefebvre	1,169	100
EDP Sciences	889	85
Elnet	9,525	192

TJ - univ-lorraine-ezagimus - camembert-BC



TJ - univ-lorraine-ezagimus - camembert-collegiums



TJ - univ-lorraine-ezagimus - camembert-poles



GC - univ-lorraine-ezagimus: Utilisateurs ...

- Etudiant
- Enseignant / Chercheur
- Doctorant
- Contractuel E - C
- Personnel BIATSS
- Personnel Hébergé E ...
- Personnel hébergé
- Vacataire d'enseigne...
- Retraité
- Autres
- Personnel partenaire
- Auditeur libre
- Etudiant conventionné

25,046

Utilisateurs uniques

GC - univ-lorraine-ezagimus : compteur...

1,525,664

événements de consultation

TJ - univ-lorraine-ezagimus - camembert...

- COLL DEG
- COLL SHS
- COLL L-INP
- COLL SCIENCES TECH
- COLL SANTE
- COLL TECHNOLOGIE
- COLL ALL
- COLL L-LMI
- COLL INTERFACE
- LOLF FORMATION



- Master
- Licence
- Doctorat
- INCONNU

- ECOLES DOCTORALES
- POLE SPEG
- POLE M4
- POLE EMPP
- POLE AMZI
- POLE CPM
- POLE OTELO
- POLE CLCS
- POLE AZF

Ex.5 (ezMESURE) : Academic departments

Tj - univ-lorraine-ezagimus - table-collegium

1. Sélectionner un collegium...	Count
COLL L-INP	15.780
COLL SHS	15.553
COLL DEG	14.636
COLL SANTE	10.122
COLL SCIENCES TECH	10.072
COLL TECHNOLOGIE	6.539
COLL ALL	4.684
COLL L-LMI	3.092
COLL INTERFACE	2.599

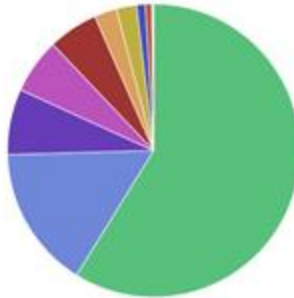
Tj - univ-lorraine-ezagimus - table-composante

2. ... puis une composante	Count
LFR SHS - NANCY	11.223
DSEG NANCY	10.404
FST	7.108
MEDECINE	5.150
LFR SHS - METZ	4.330
DEA METZ	3.728
LFR ALL NANCY	3.286
FNAC	2.684

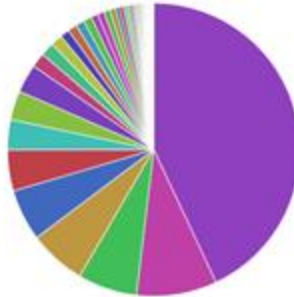
Tj - univ-lorraine-ezagimus - table-BC

Nombre de consultations par type de population	Count
Étudiant	59.695
Enseignant / Chercheur	15.927

Tj - univ-lorraine-ezagimus - camembert-BC



Tj - univ-lorraine-ezagimus - camembert-platform



Tj - univ-lorraine-ezagimus - camembert-type

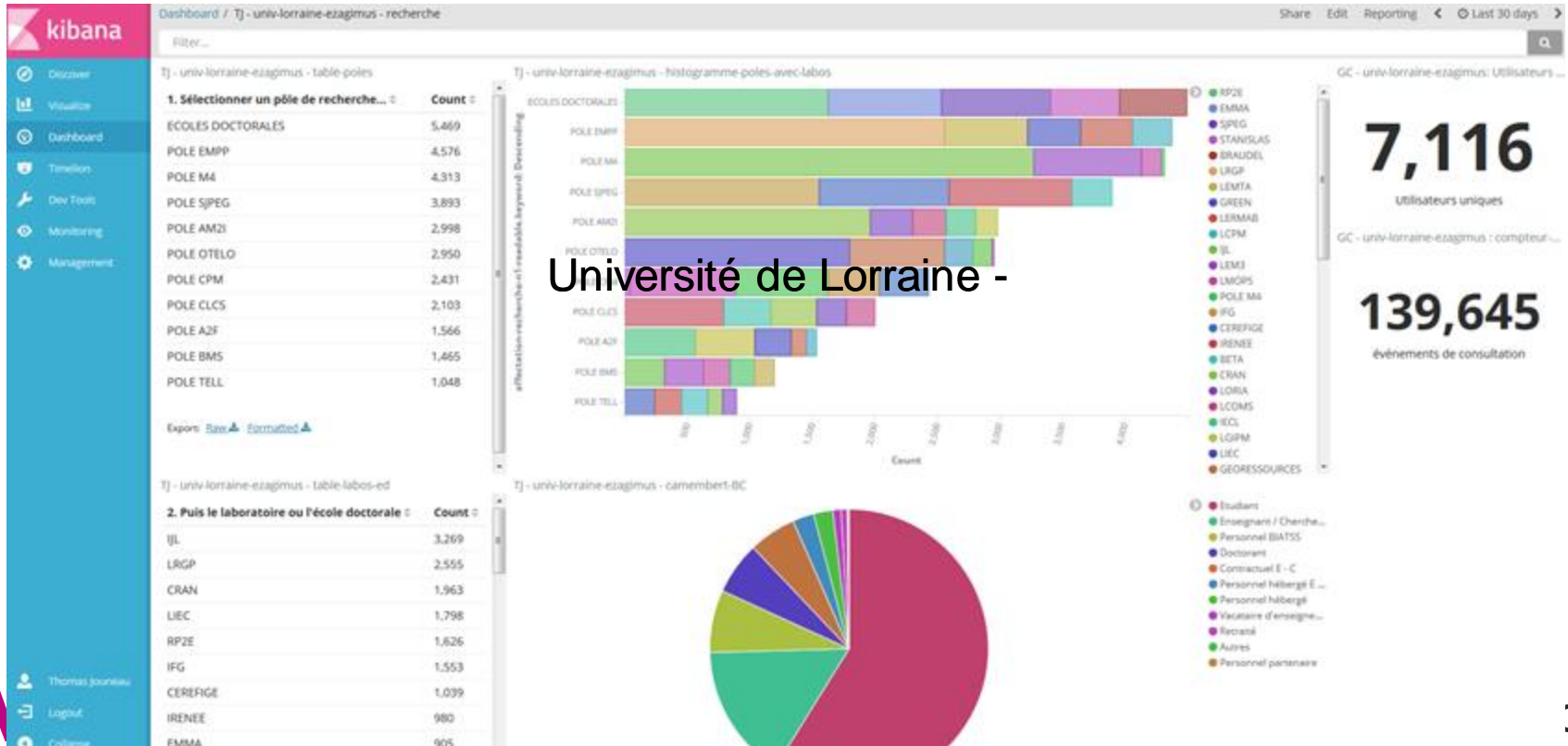
GC - univ-lorraine-ezagimus : Utilisateurs



GC - univ-lorraine-ezagimus : compteurs



Ex.7 (ezMESURE) : Rsearch labs and poles



Thank you

Dominique Lechaudel

INIST-CNRS

dominique.lechaudel@inist.fr

Thomas Jouneau

UNIVERSITÉ DE LORRAINE

thomas.jouneau@univ-lorraine.fr