

A GLAM tool

How manage a huge mass of data... and still survive!

Ilario Valdelli, Wikimedia CH

The need of GLAM and our need..

To measure an impact!

With data which are reliable, consistent and complete.



The limit of data...

To measure an impact
with data which are
reliable, consistent and complete...

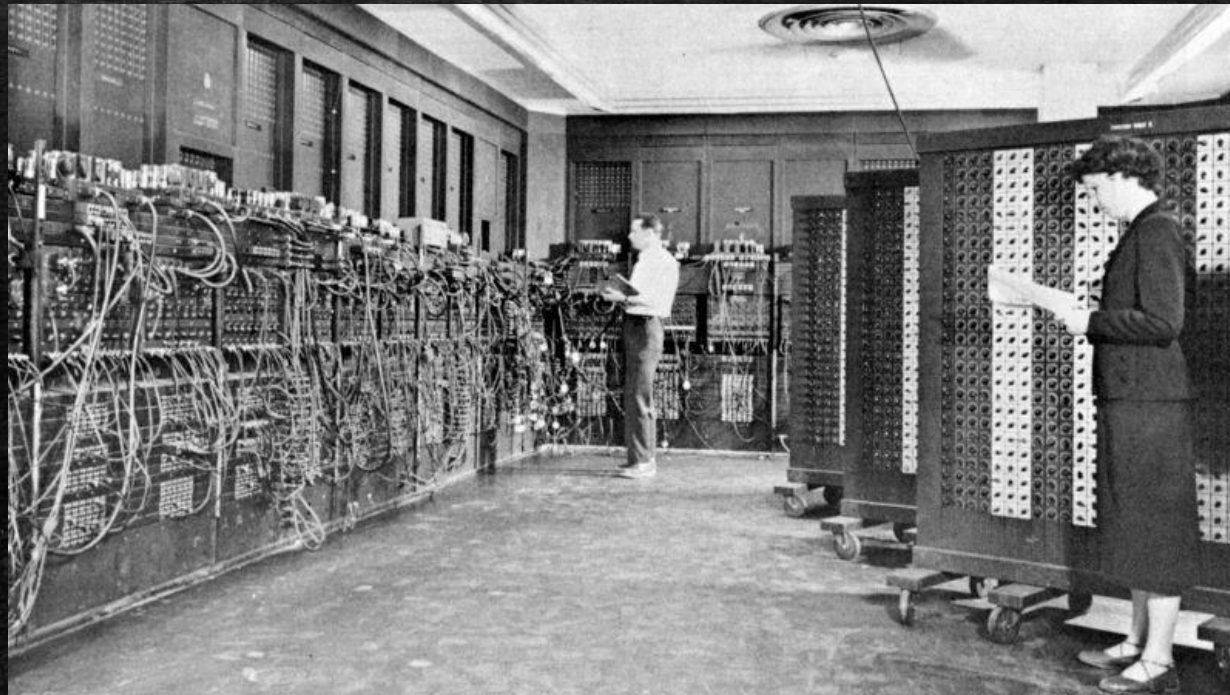
and

unfortunately

without value for an analysis.

Big computers and big data

It's not sufficient to have a super computer and a big set of data...



Big computers and big data

The result may be correct but cannot be evaluated.

42

“The answer to the ultimate question of life, the universe and everything!”

GLAM tool

Managing a request

GLAM network

Wikimedia Switzerland has a growing network of GLAM meeting periodically and including biggest partners:

National Archives

National Library

Library of ETH of Zürich

University of Basel...

GLAM network: their need

In the 2015 the network indicated to have an urgent need to monitor the impact produced in Wikimedia projects and to analyze and define the activities to do to improve this impact!

First approach

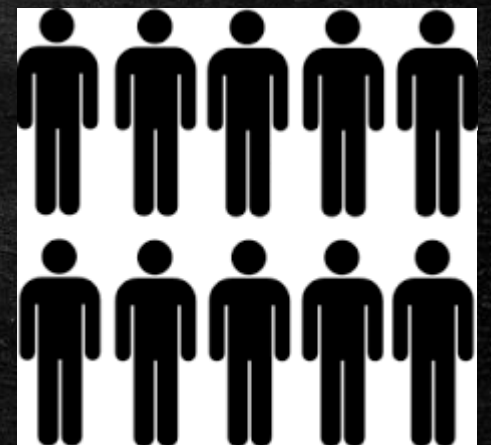
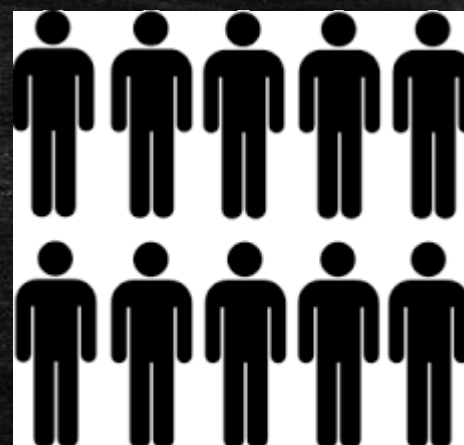
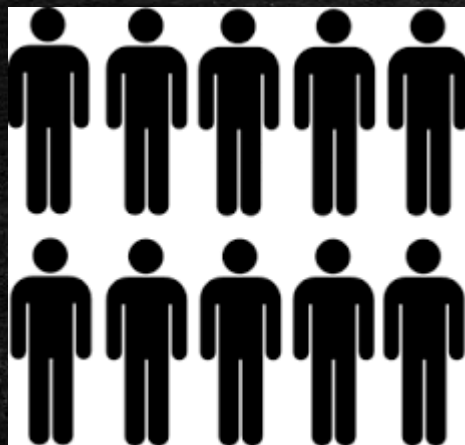
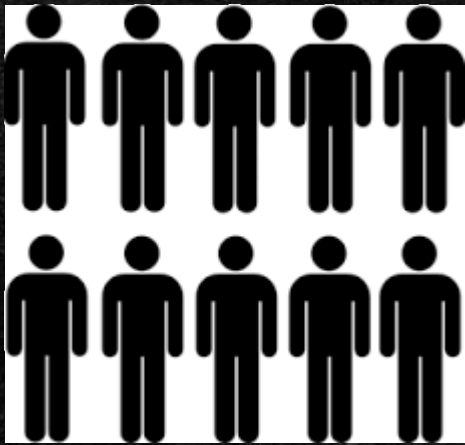
Giving them the list of usual tools used
and developed by the community
about statistics.

Result

Unhappy.

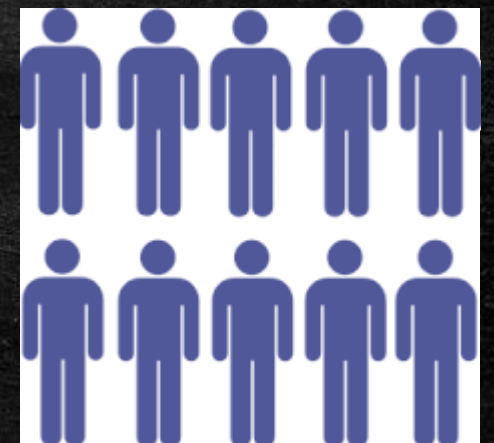
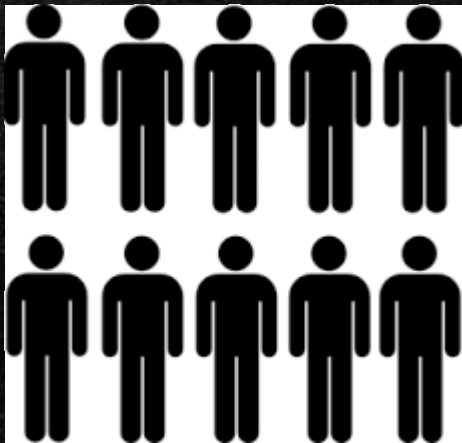
GLAM network: The use of tools

Tools are designed and developed by people who suppose that the final users are an indistinct mass of users.



GLAM network: Experiment

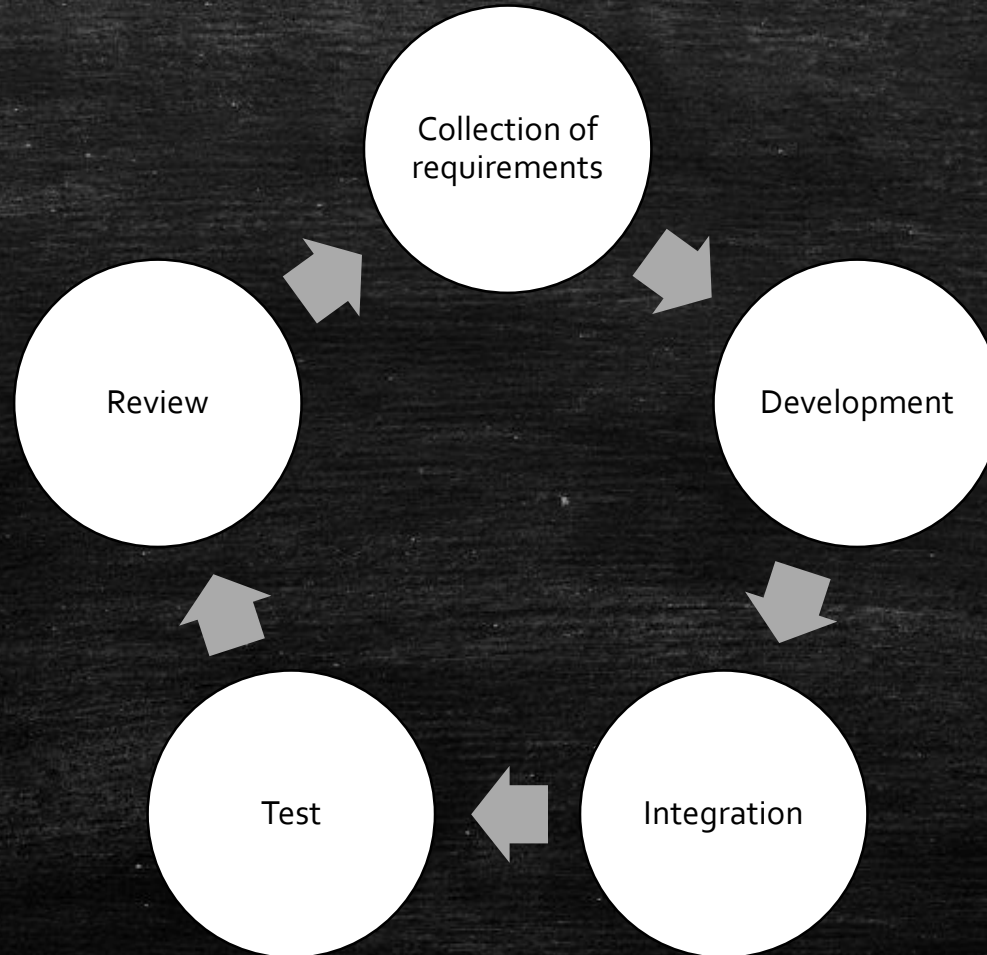
GLAM used several existing tools and discovered to don't be satisfied or to have limited tools.



GLAM tool

The approach to the usual software development

The approach



First problem

- To understand the needs
- To understand the workload
- To understand requirements

Anforderungen an das GLAMWiki-Tool

1 → Anforderungen an das GLAMWiki-Tool



1.1 → Änderungskontrolle

Version	Datum	Beschreibung, Bemerkung	Name oder Rolle
0.1	20.06.2016	Initialversion	Rimi
0.2	20.06.2016	Prüfung	Nepfer
0.3	23.06.2016	Prüfung	Mm

1.2 → Beschreibung

Die Anforderungen wurden in zwei Sektionen eingeteilt:

[Vorbemerkung: Ich würde die Statistikanforderungen eher trennen von den

Verwaltungsanforderungen, um das Fuder nicht zu überladen. Für die NB sind die Statistiken wichtiger als die Verwaltungsinstrumente, letztere können auch später implementiert werden.]



Majoleth Marco BAR

Auch wir sind vorerst vor allem an Statistiken interessiert. Vielleicht die Verwaltungsanforderungen nicht gerade aus dem Dokument entfernen, sondern die Sektionen entsprechend priorisieren.

Group of functions

- Monitoring of the categorisation and map of the content
- Use in Wikimedia projects
- Visits and download in Wikimedia projects
- Extension of the focus outside Wikimedia projects
- Tools to facilitate the corrective actions
- Starting to operate in Wikidata!

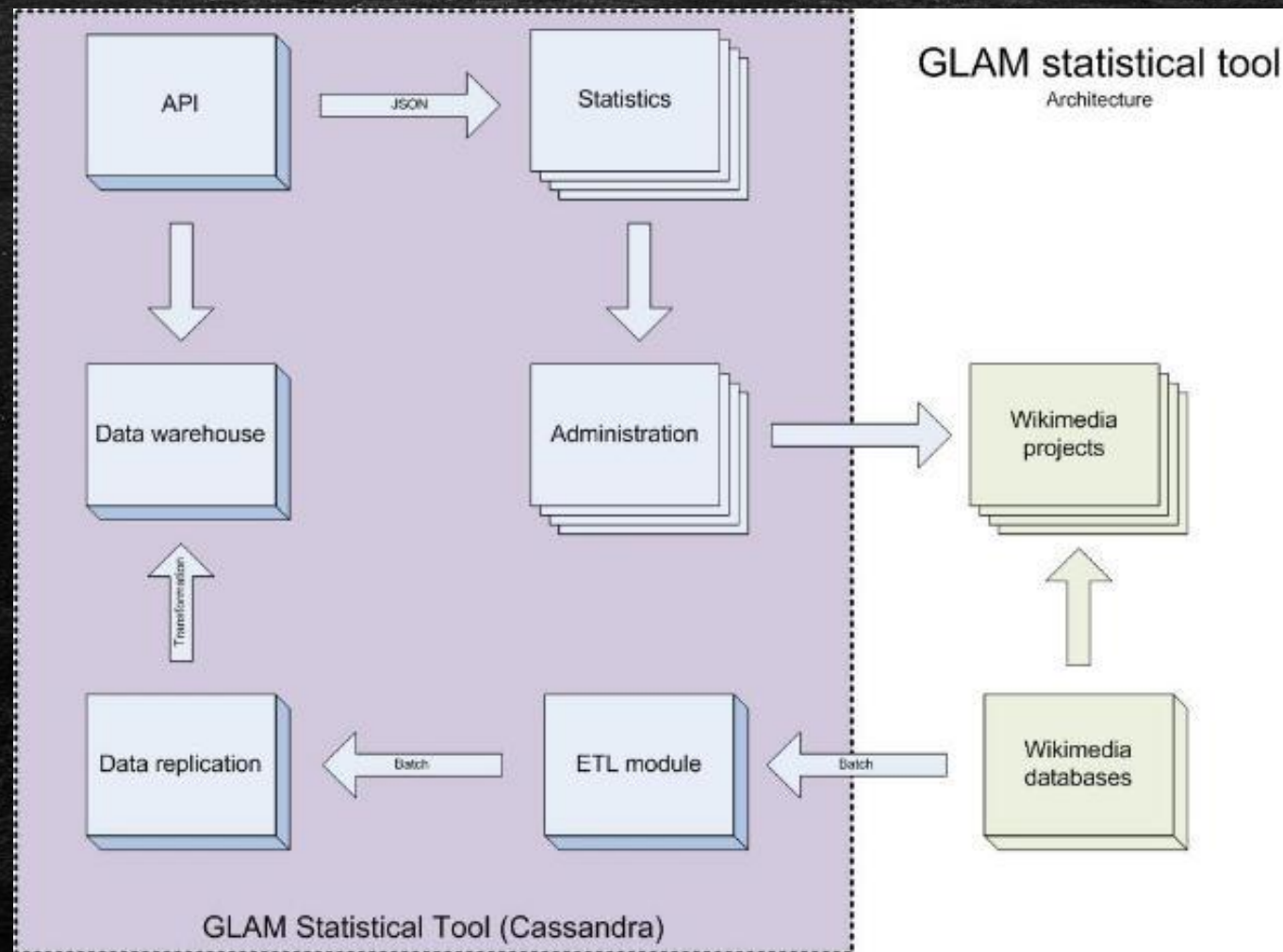
Timeline

- July 2016 – Collection of functional requirements
- August 2016 – Technical definition of requirements
- September 2016 – Definition of the architecture
- October 2016 – Definition of the teams
- February 2017 – Closure of the first phase

GLAM tool

Technical aspects

Architecture 3.0



Modular and distributed architecture

- No desktop application
- Scalable architecture
- Modular architecture
- Distributed architecture
- Web services and API

In February the final result

Two different teams

Working in two separate countries

With two separate documentations

The first team more expert of Wikipedia.

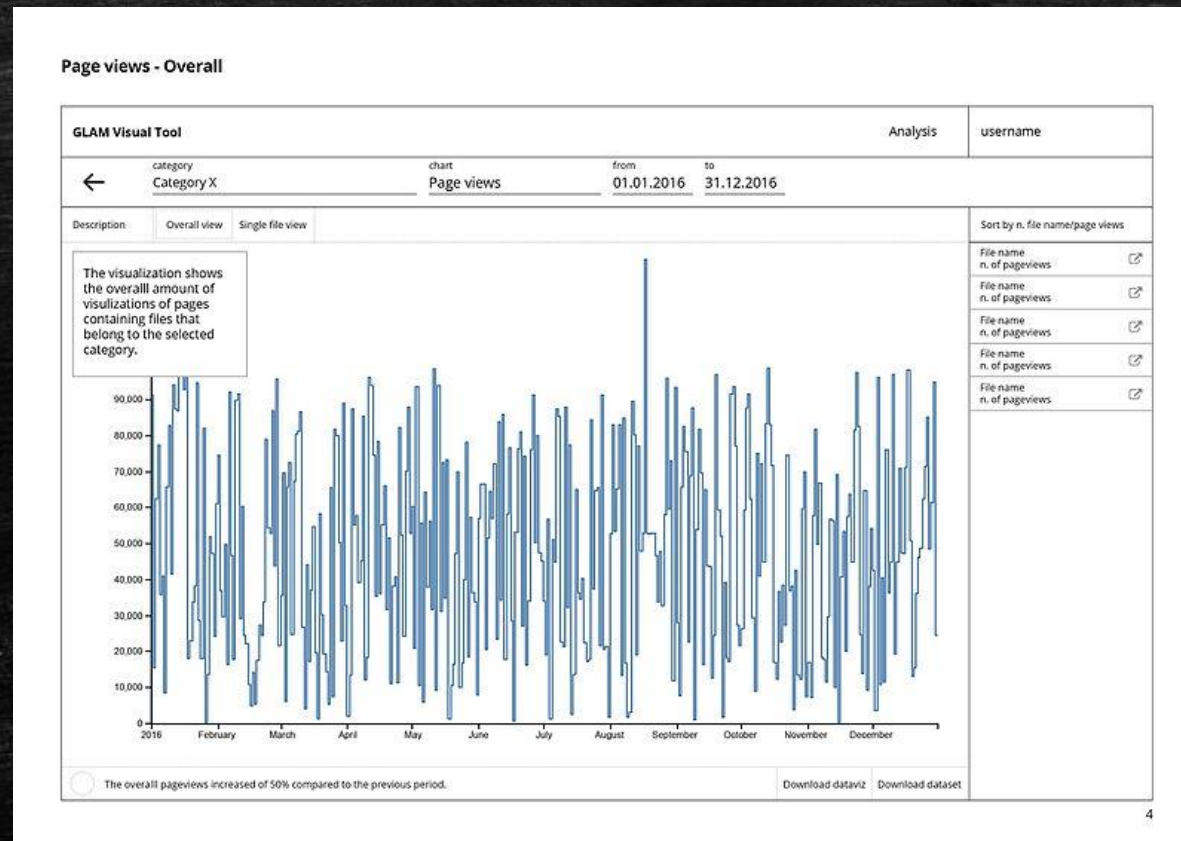
The second team more expert of software development.

The results

- A first team closer to the Wikimedia world started to rethink and to rebuild requirements and to refactor them.
- The second team had a very pragmatic approach to the problem implementing all requirements.

The front end

- https://meta.wikimedia.org/wiki/Research:GLAM_visual_tool



The back end

- <http://cassandra.synapta.io/docs>

Thank you

