

# **Human-Machine Collaboration in Networked Information Systems**

**Claudia Müller-Birn**

Institute of Computer Science, Freie Universität Berlin

March 2, 2017



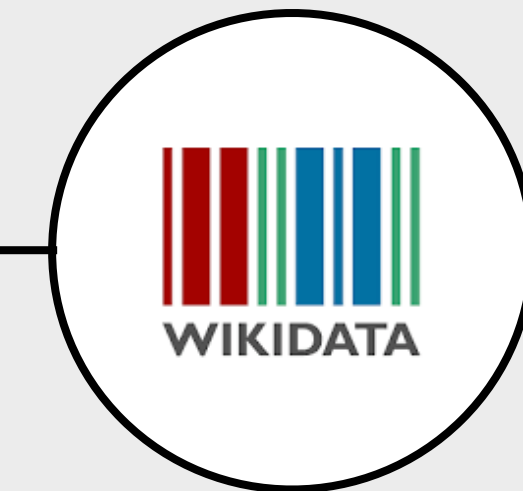
**Social Web**  
that connects  
people and knowledge

**Semantic Web**  
that connects  
machines and knowledge

**Human-Machine  
Collaboration**

# Research Context

Networked information systems are socio-technical systems that integrate **distributed information resources** based on **usable user and programming interfaces** to enable **human-machine collaborations**.

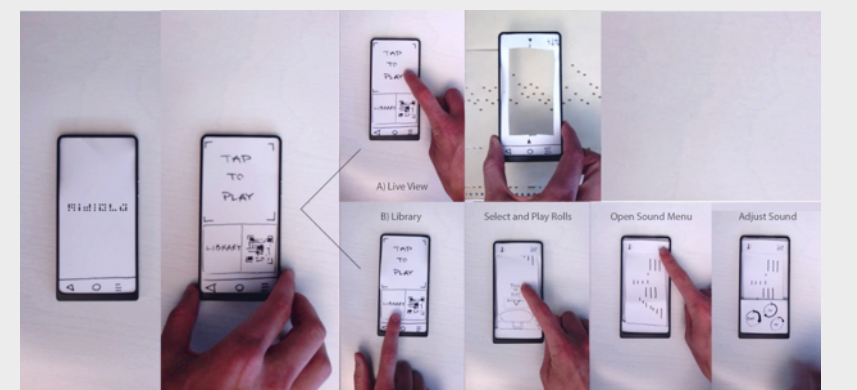
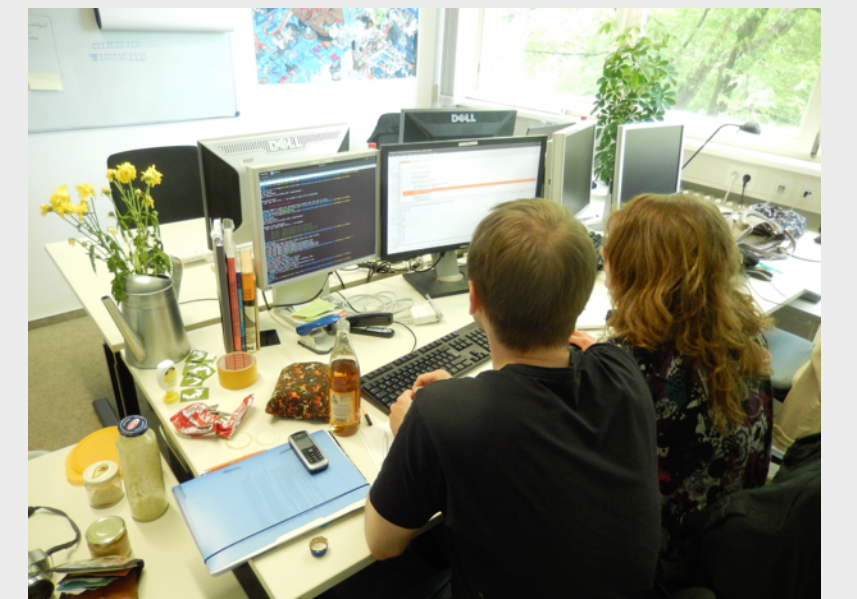
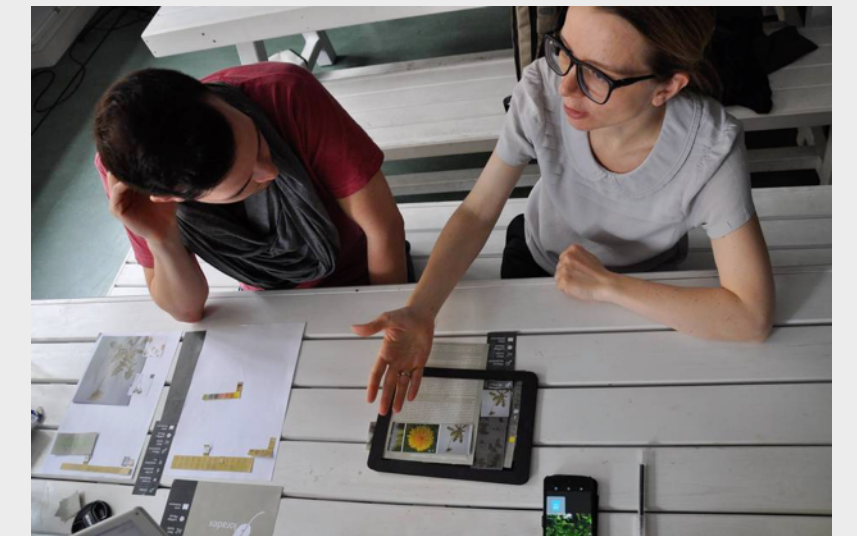




# Research Areas

We focus on designing networked information systems for both, human and algorithmic agents in the following three areas:

- (1) Designing efficient, effective and easy-to-use interfaces.
- (2) Enabling a seamless coordination between knowledge processes.
- (3) Expanding collaboration into the physical world.





# Work Location 1



The "almost" complete HCC:Team (Summer 2016)





# Work Location 2



Bild  
Wissen  
Gestaltung

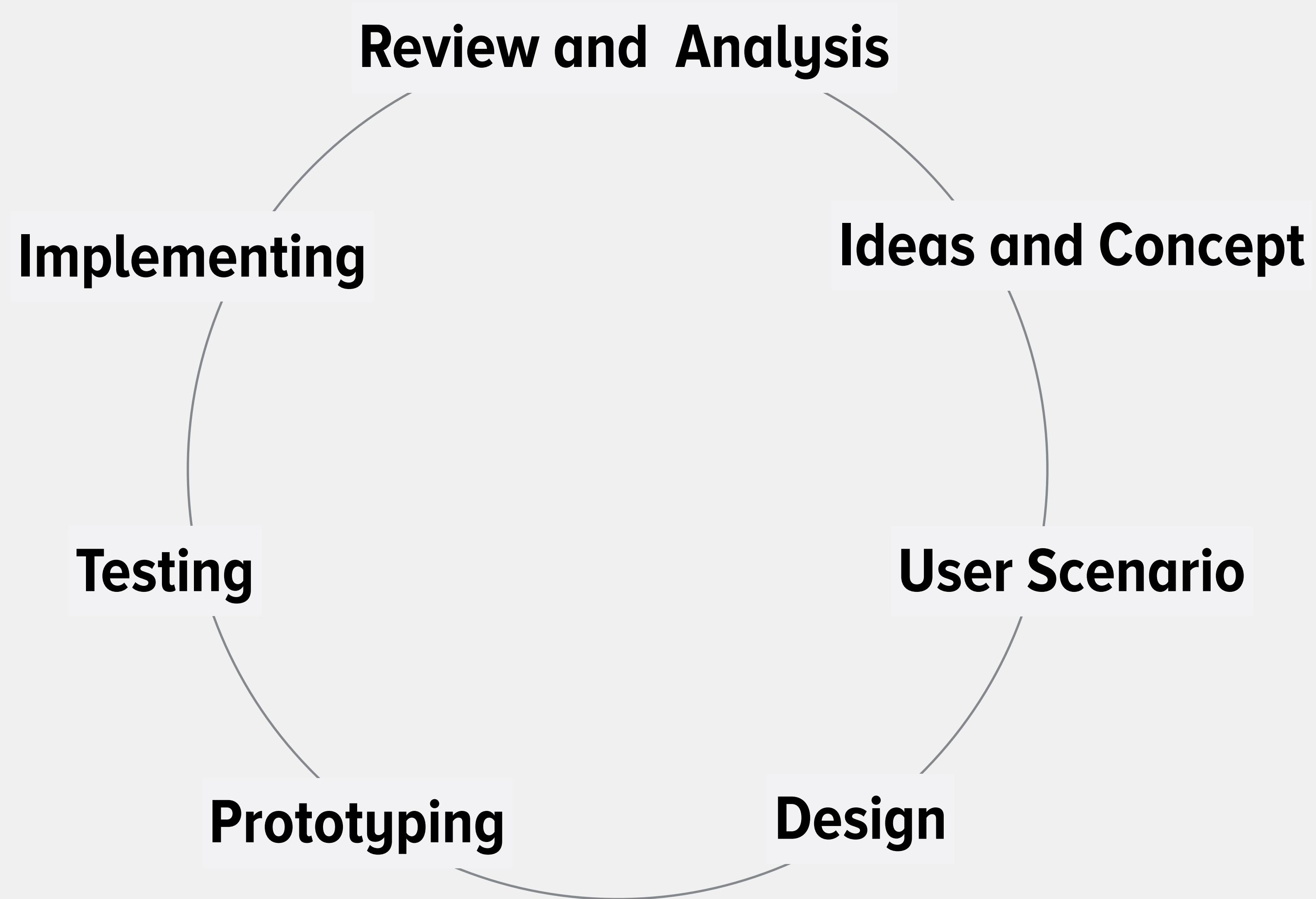
Ein Interdisziplinäres Labor



Gefördert durch die  
**DFG**



# Working Process





**Designing** networked information systems

**Analysing** relationships between humans and machines in networked information systems

**Developing** coupled relationships between humans and machines in networked information systems

**Extending** networked information systems

**Designing** networked information systems

**Analysing** relationships between humans and machines in networked information systems

**Developing** coupled relationships between humans and machines in networked information systems

**Extending** networked information systems

# Challenges in online communities

Everyone is better off if everyone contributes than if no one does



Each individual is even better off if she does not contribute while the others do

Critical mass

Social loafing



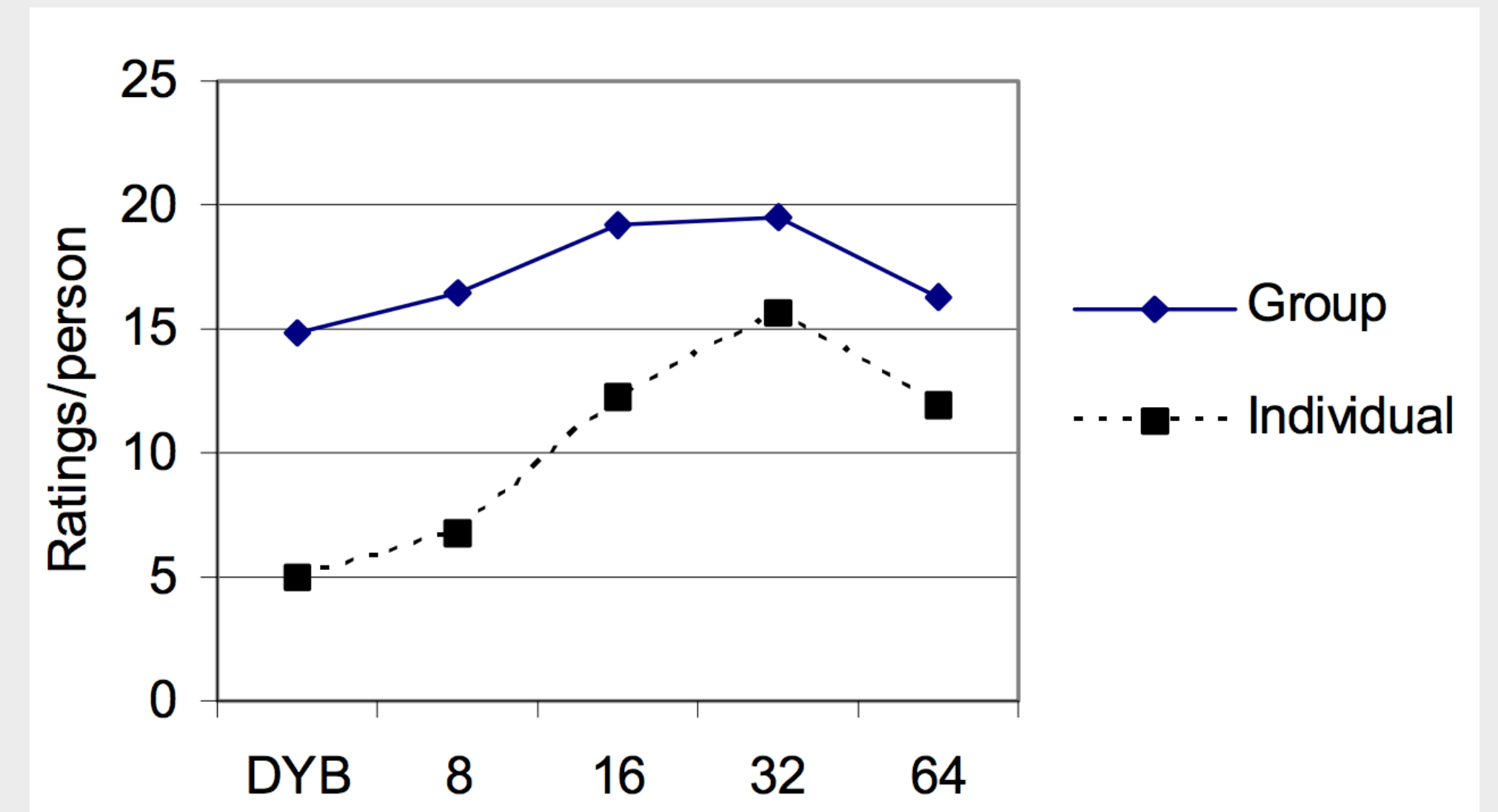
# Feedback: Goal setting

**Context:** MovieLens is a movie recommendation site (<http://www.movielens.org>)

**Hypothesis:** In an online community, specific, numeric goals will motivate greater contributions than non-specific goals

**Approach:** Members were recruited by sending them an email message which contained an invitation to participate in a movie rating campaign

**Result:** Providing community members with specific goals increases their contributions.



# Sociality: Social contact

Context: Wikipedia

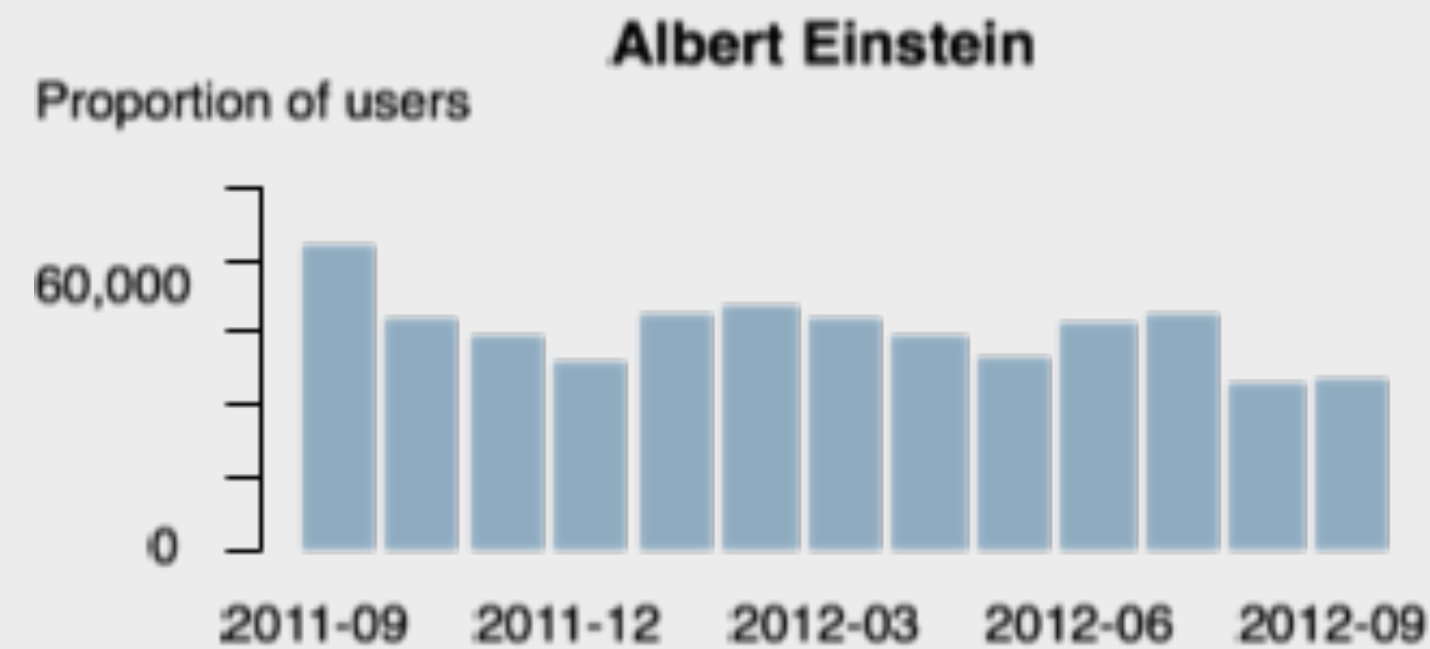
Hypothesis: Social contact improves the retention rate of new editors

Approach: Create an online teahouse for meeting new editors

Result: Social contact with other similar contributors causes members to stay and to contribute more.

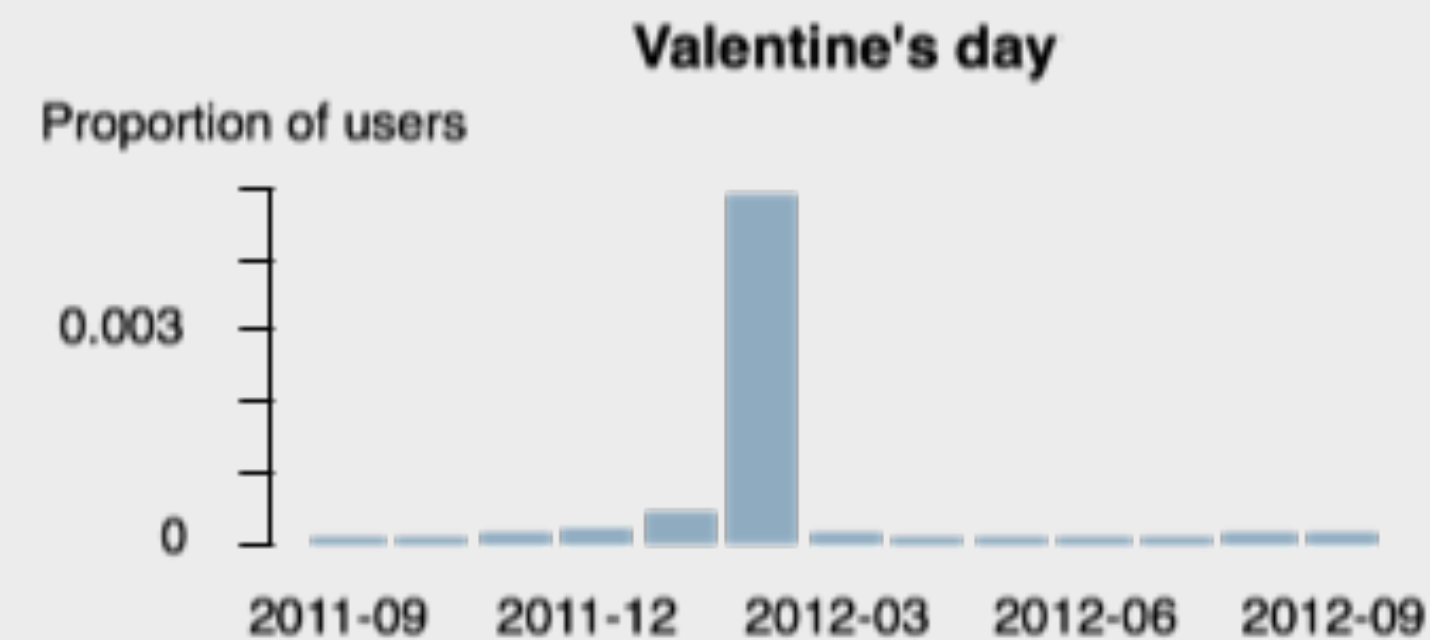
The screenshot shows the Wikipedia:Teahouse website. At the top, it says "Wikipedia:Teahouse" and "From Wikipedia, the free encyclopedia". The main header features a green background with the word "teahouse" in blue and a "TH" logo. Below the header, there are navigation links for "guest intros", "host intros", and "Q&A forum". A "welcome" message is displayed on the right. The main content area is divided into two columns. The left column is titled "Guests!" and contains a section for "Please introduce yourself" with a profile picture of Syed S Haider and a quote: "Share to expand or be a foll". Below this is a link to "welcome this guest". The right column is titled "Recent questions—" and contains a question about a source reposted on a blog. Below the question is a link to "Find the answer here" and a link to "...see another question". At the bottom right, there is a section titled "Meet your hosts" with a profile picture of Nolelover and a quote by Thomas Jefferson: "As Thomas Jefferson (probably) said, 'Whenever you do a thing, act as if all the world were watching.' This quote is extremely apt for Wikipedia, where the work you put in may be seen by far more people than you would think possible."

# Feedback: Make readers' interests visible



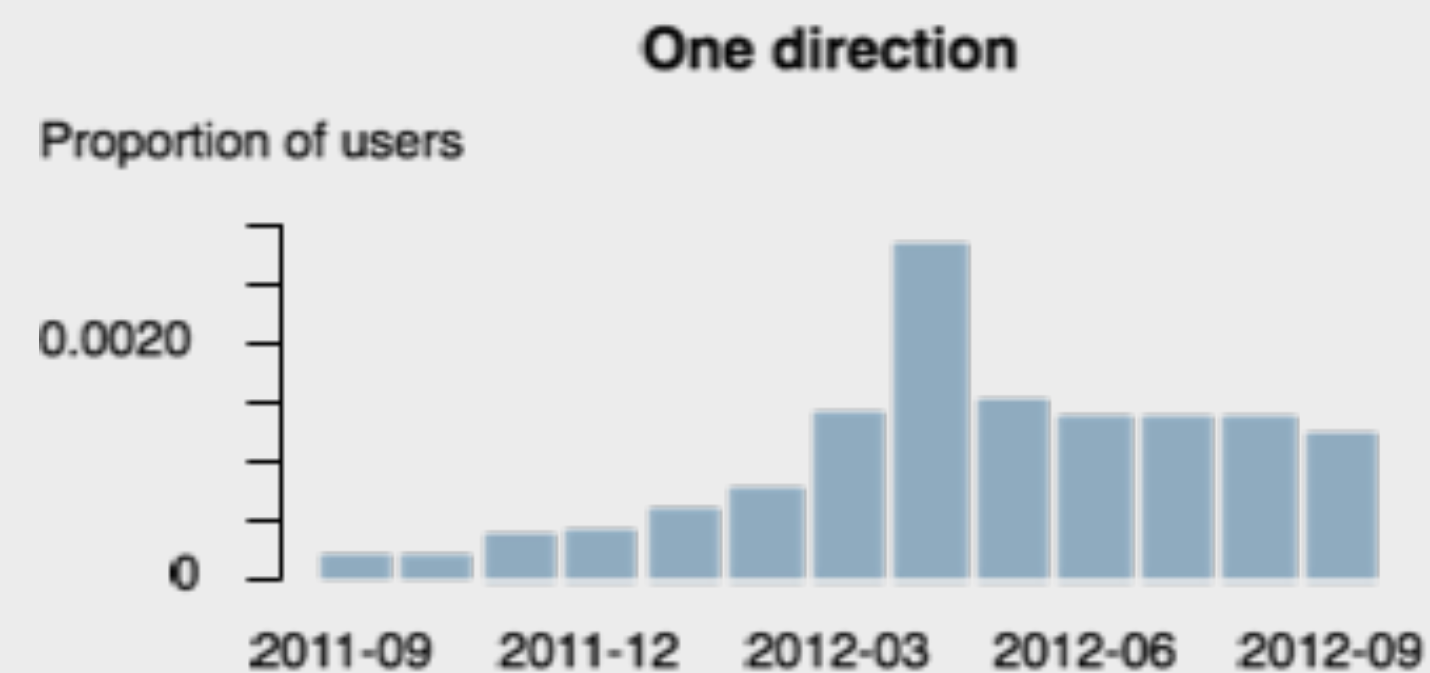
## Constant interest

Regularly accessed articles, sometimes only for fact finding, e.g. Albert Einstein, Facebook, IMDB



## Peak interest

Death of people, game and movie releases, e.g. Whitney Houston, The Hunger Games, 2012 Phenomeno



## Increasing/decreasing interest

Items that became popular/loose popularity during our observation period, e.g. One direction, Instagram



**Designing** networked information systems

**Analysing** relationships between humans and machines in networked information systems

**Developing** coupled relationships between humans and machines in networked information systems

**Extending** networked information systems

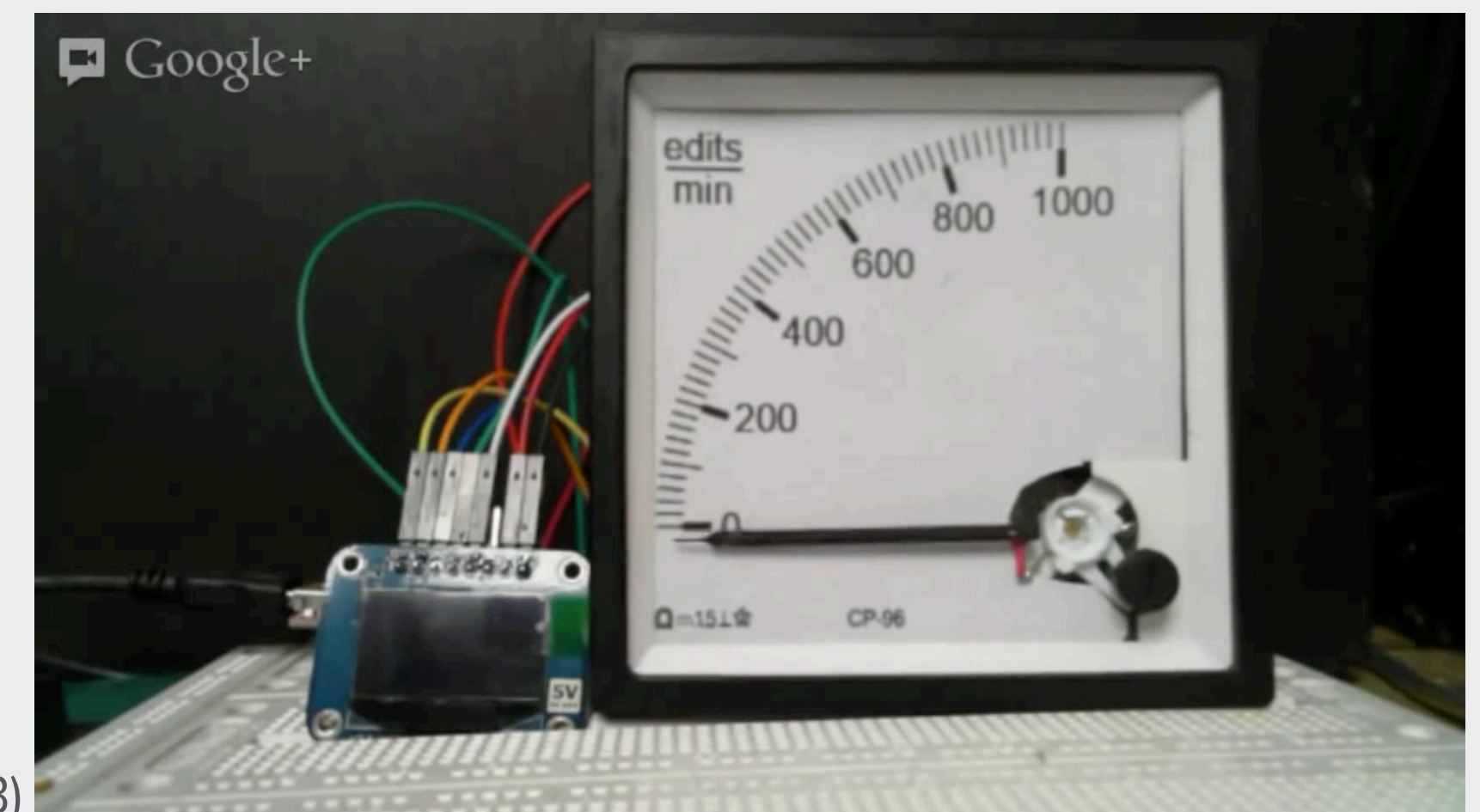
# Analysis of Contribution Patterns



**Research Goal** » Identifying participation patterns for collaborative ontology development

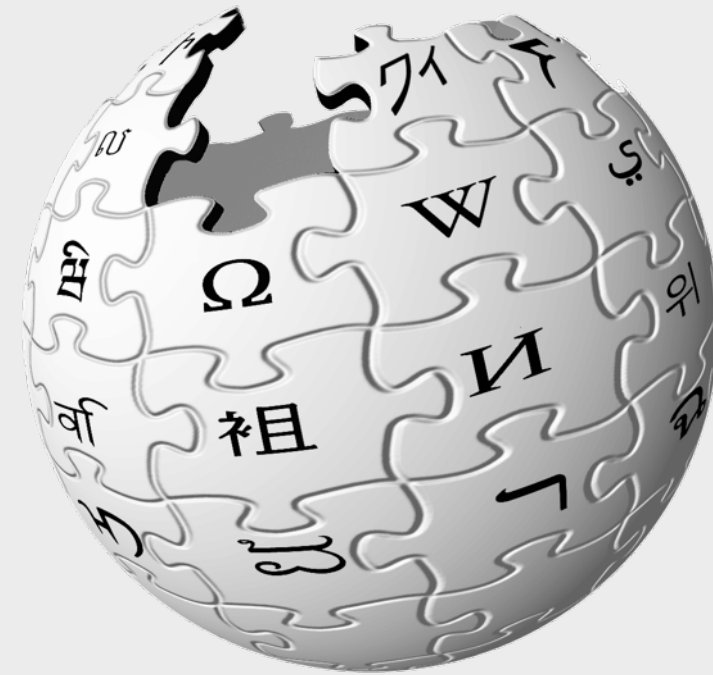
**Funding** » DFG-funded research project in collaboration with GESIS - Leibniz-Institute for Social Science

**Data Collection** » JSON and XML-based data dump (10/2012 to 10/2014)



Wikidatameter by Lukas Benedix (2013)

# “What is the population of Berlin?”



Create account Log in

Location within European Union and Germany  
Coordinates: 52°31′N 13°23′E﻿ / ﻿52.517°N 13.383°E﻿ / 52.517; 13.383

<b>Country</b>	Germany
<b>Government</b>	
<span> </span> • <b>Governing Mayor</b>	Michael Müller (SPD)
<span> </span> • <b>Governing parties</b>	SPD / CDU
<span> </span> • <b>Votes in Bundesrat</b>	4 (of 69)
<b>Area</b>	
<span> </span> • <b>City</b>	891.85 <span> </span> km <sup>2</sup> (344.35 <span> </span> sq <span> </span> mi)
<b>Elevation</b>	34 <span> </span> m (112 <span> </span> ft)
<b>Population</b> (December 2014) <sup>[1]</sup>	
<span> </span> • <b>City</b>	3,562,166
<span> </span> • <b>Density</b>	4,000/km <sup>2</sup> (10,000/sq mi)
<b>Demonym</b>	Berliner
<b>Time zone</b>	GET (UTC+1)
<span> </span> • <b>Summer (DST)</b>	CEST (UTC+2)
<b>Postal code(s)</b>	10115–14199
<b>Area code(s)</b>	030
<b>ISO 3166 code</b>	DE-BE
<b>Vehicle registration</b>	B <sup>[2]</sup>
<b>GDP/ Nominal</b>	€109.2 billion (2013) <sup>[3]</sup>
<b>NUTS Region</b>	DE3
<b>Website</b>	berlin.de <span><span><span></span></span></span>

1.3 17th to 19th centuries  
1.4 20th to 21st centuries

2 Geography

2.1 Topography

2.2 Climate

2.3 Cityscape

2.4 Architecture

3 Demographics

3.1 International communities

3.2 Religion

4 Government

4.1 City state

4.2 Boroughs

4.3 Sister cities

4.4 Capital city

5 Economy

5.1 Companies

5.2 Tourism and conventions

5.3 Creative industries

5.4 Media

6 Infrastructure

6.1 Transport

6.2 Energy

6.3 Health

7 Education

7.1 Higher education

7.2 Research

8 Culture

8.1 Galleries and museums

8.2 Nightlife and festivals

8.3 Performing arts

8.4 Cuisine

8.5 Recreation

8.6 Sports

9 See also

- Acèh
- Адыгэбзэ
- Адыгабзэ
- ★ Afrikaans
- Akan
- Alemannisch
- አማርኛ
- Ænglisc
- АӀсшәа
- العربية
- Aragonés
- አድሬ
- Arpetan
- Asturianu
- Avañe'ê
- Aymar aru
- Azerbaycanca
- تۆرکجه
- Bamanankan
- বাংলা
- Bân-lâm-gú
- ★ Башкортса
- Беларуская
- Беларуская (тарашкевіца)
- Bikol Central
- Bislama
- Български
- Boarisch
- བོད་སྐད་
- Bosanski
- Brezhoneg
- Буряад
- Català
- ЧӀавашла
- Cebuano
- Čeština
- Chavacano de Zamboanga
- Chi-Chewa
- ChiTumbuka
- Corsu
- Cymraeg
- ★ Dansk
- Deutsch
- ★ Deutsch
- Dolnoserbski
- Eesti
- Ελληνικά
- Эрзянь
- Español
- Esperanto

**Population** 3 452 911 hab. (09/2014)<sup>1</sup>

**Einwohner:** 3.466.164 (30. November 2014)<sup>[4]</sup>

**Biztanleria** 3.479.740 bizt. (2011)

ජනගහණය (31 December 2011)<sup>[1]</sup>

• **Total** 3,501,872

**Население** (31 март 2011)<sup>[1]</sup>

• **Град** 3.468.900 жит.

**Población** (2013)

• **Total** 3,421,829 hab.<sup>1</sup>

**Abitanti** 3 562 166<sup>[1]</sup> (2014)

**Халык саны** 3 431 420 <sup>[1]</sup> кеше

**Population** (December 2014)<sup>[1]</sup>

• **City** 3,562,166

人口 (2013年8月31日) <sup>[1]</sup>

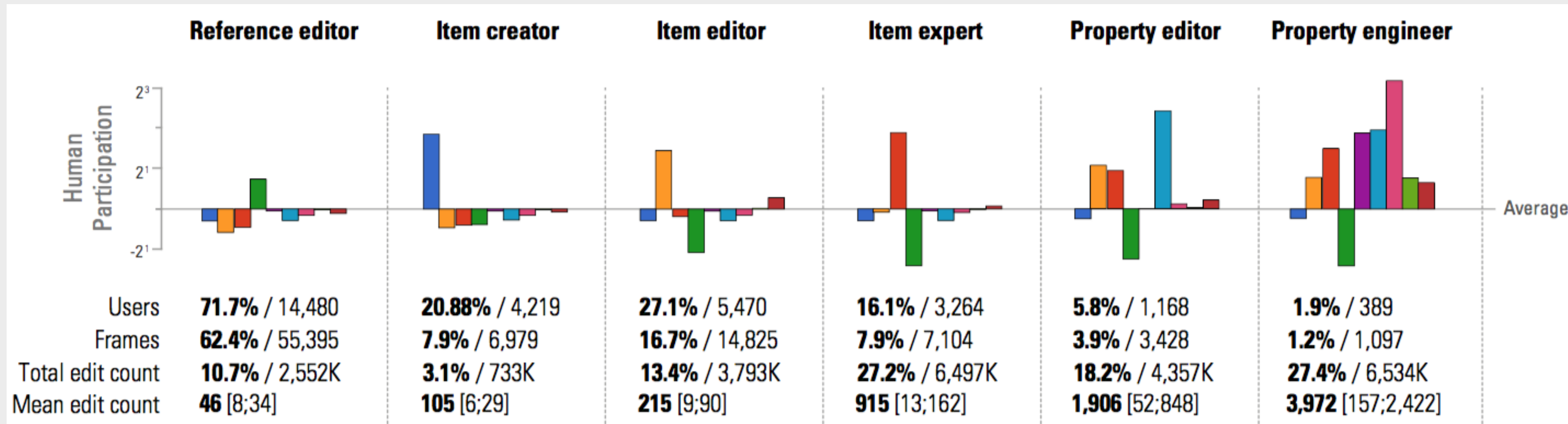
• **总计** 3,401,147





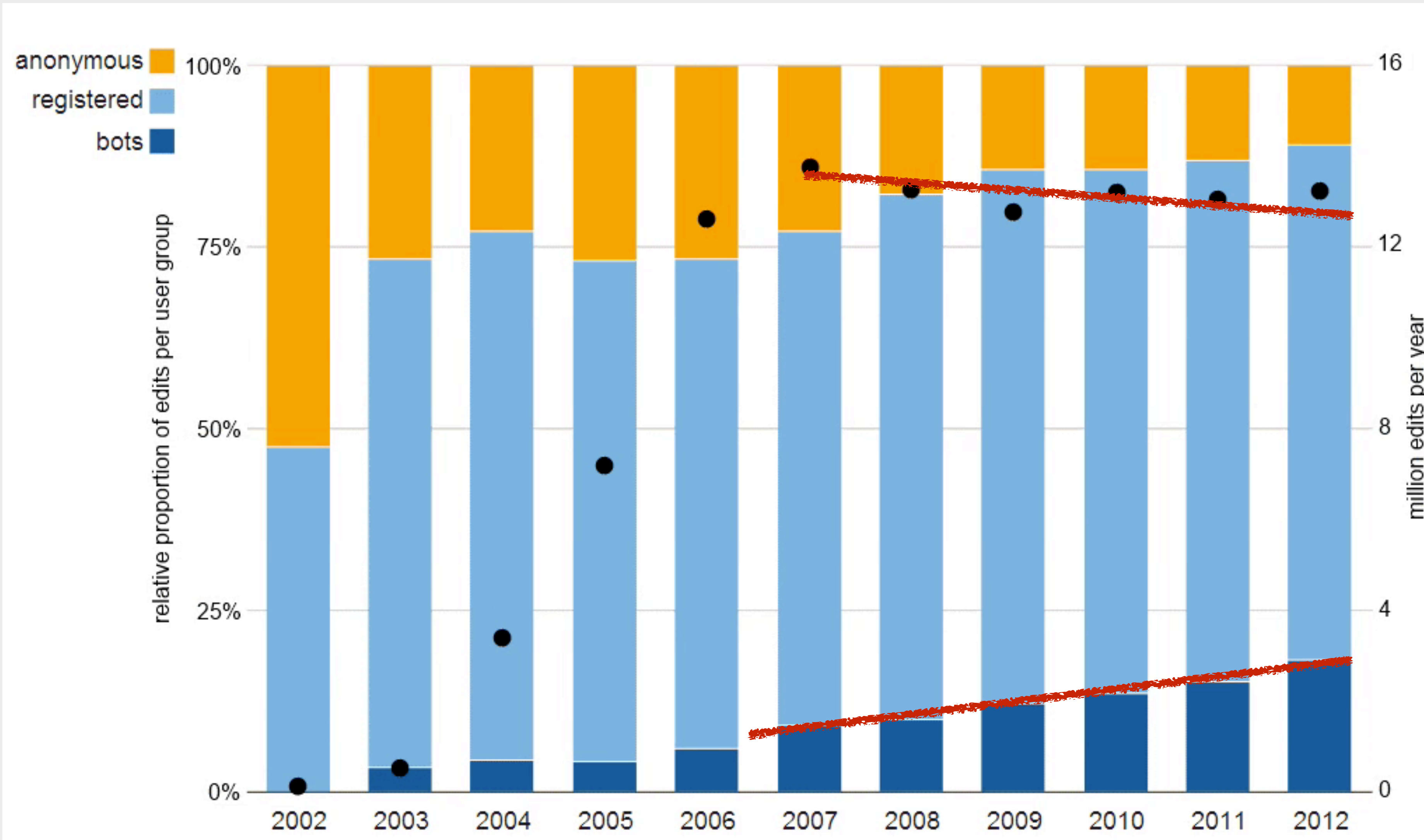
project launch in 2012 | over 460 million edits | 25 million items

# Human Contribution Patterns



- Created items
- Item terms
- Item statements
- Set sitelinks
- Created properties
- Property terms
- Discussions
- Protected pages
- Reverts

# Contributions on German Wikipedia

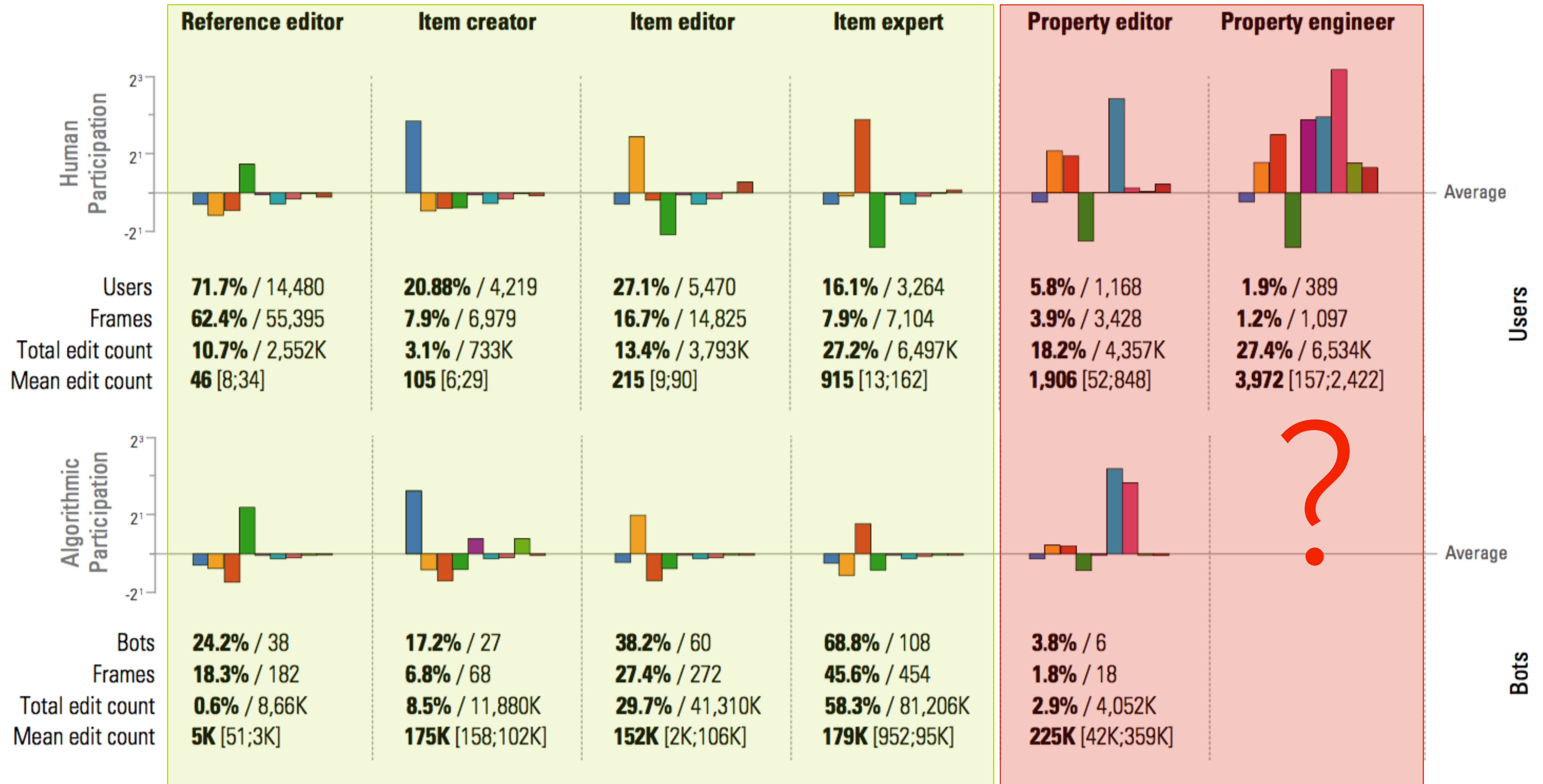


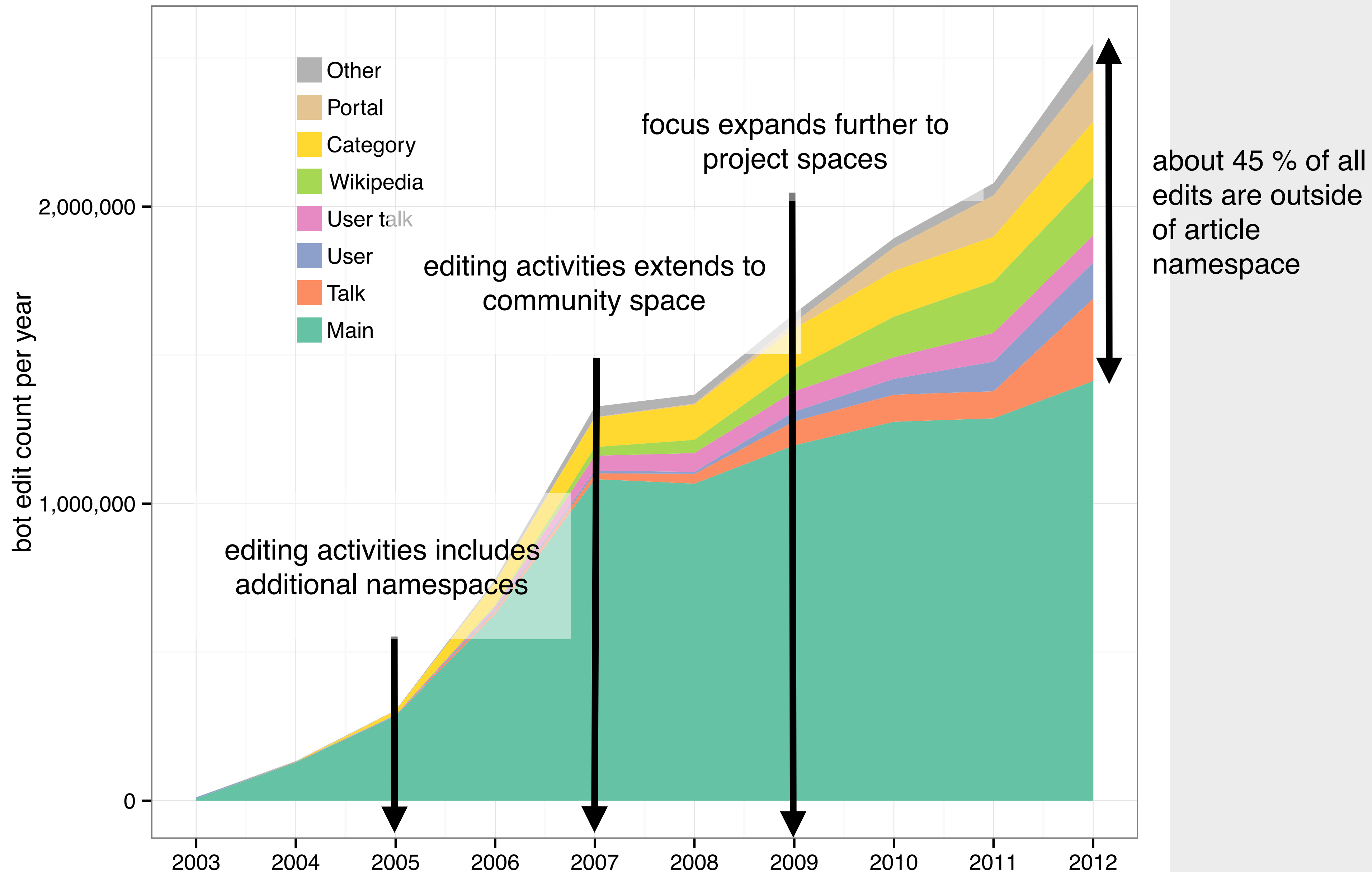




# Human and Algorithmic Contribution Patterns

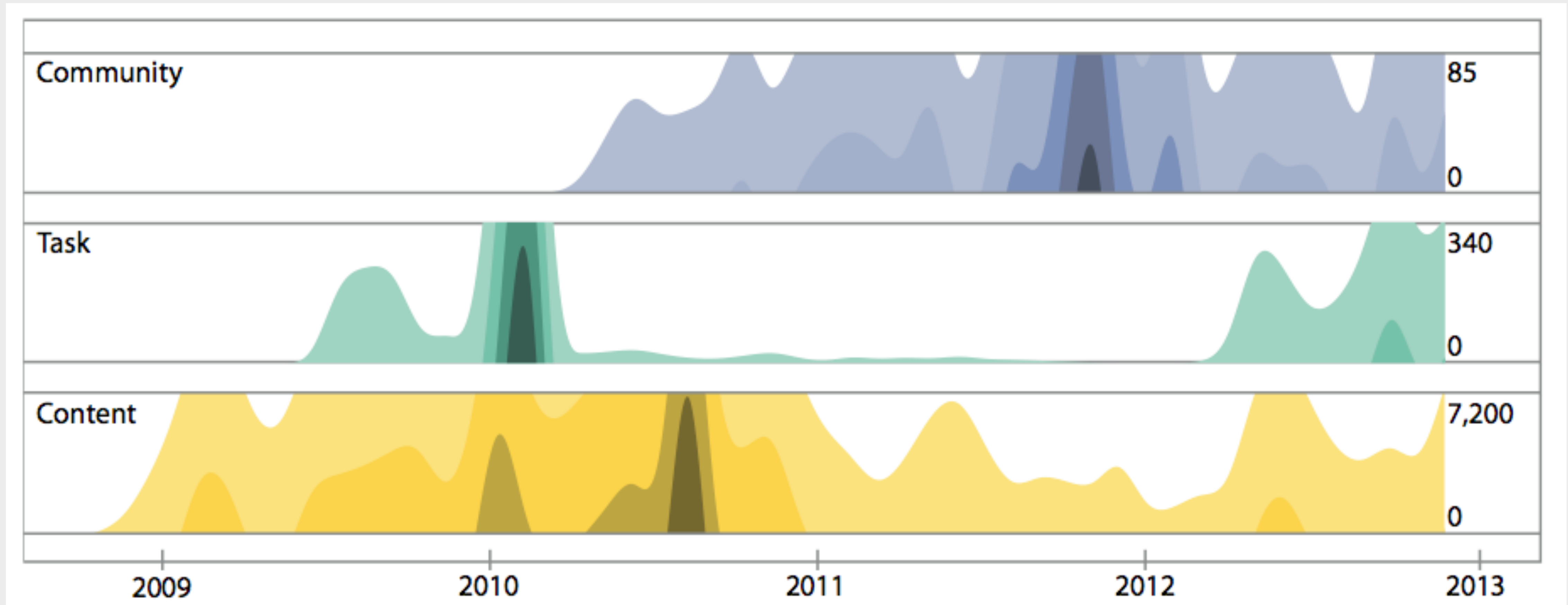
- Created items
- Item terms
- Item statements
- Set sitelinks
- Created properties
- Property terms
- Discussions
- Protected pages
- Reverts





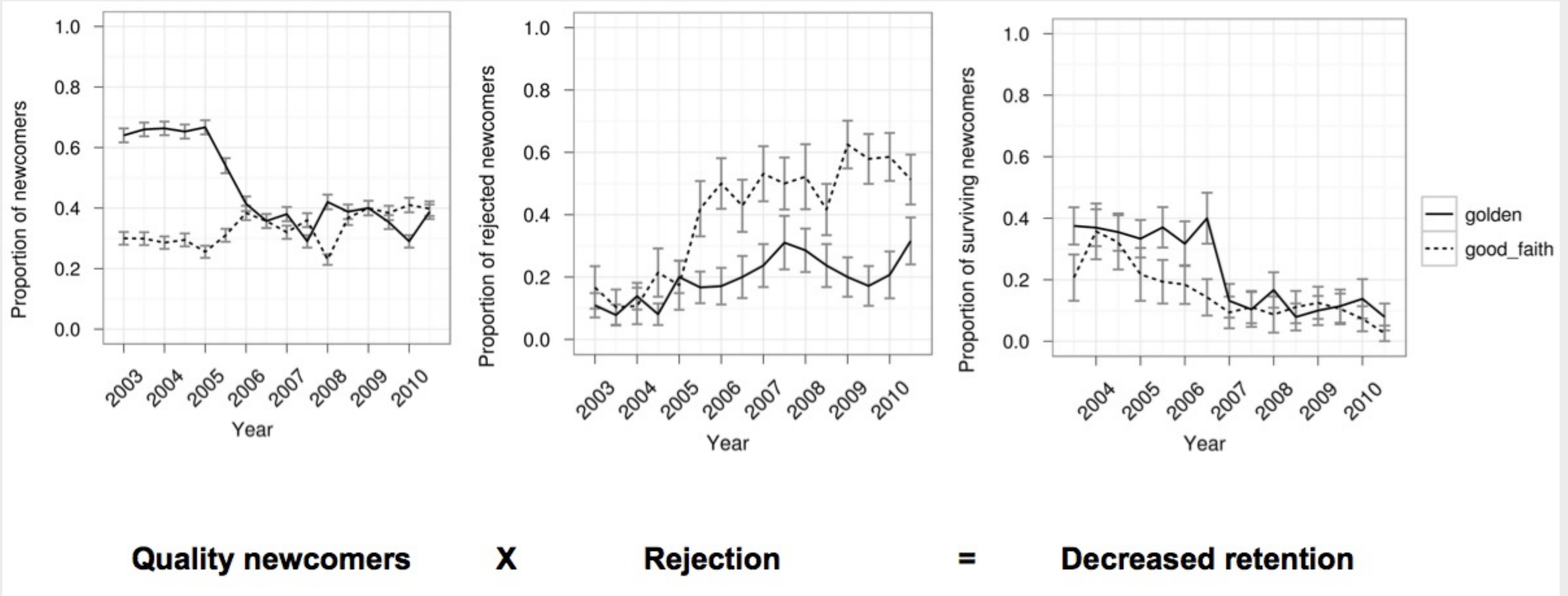


### Changing task focus of a bot over time



**Algorithmically translated rules are more scalable, but they are less likely to handle exceptions.**

# Algorithmic Governance - how bots influence user retention



**Designing** networked information systems

**Analysing** relationships between humans and machines in networked information systems

**Developing** coupled relationships between humans and machines in networked information systems

**Extending** networked information systems

# Integrating Interdisciplinary Research Data

**Research Context** » Research Data in the area of biological research

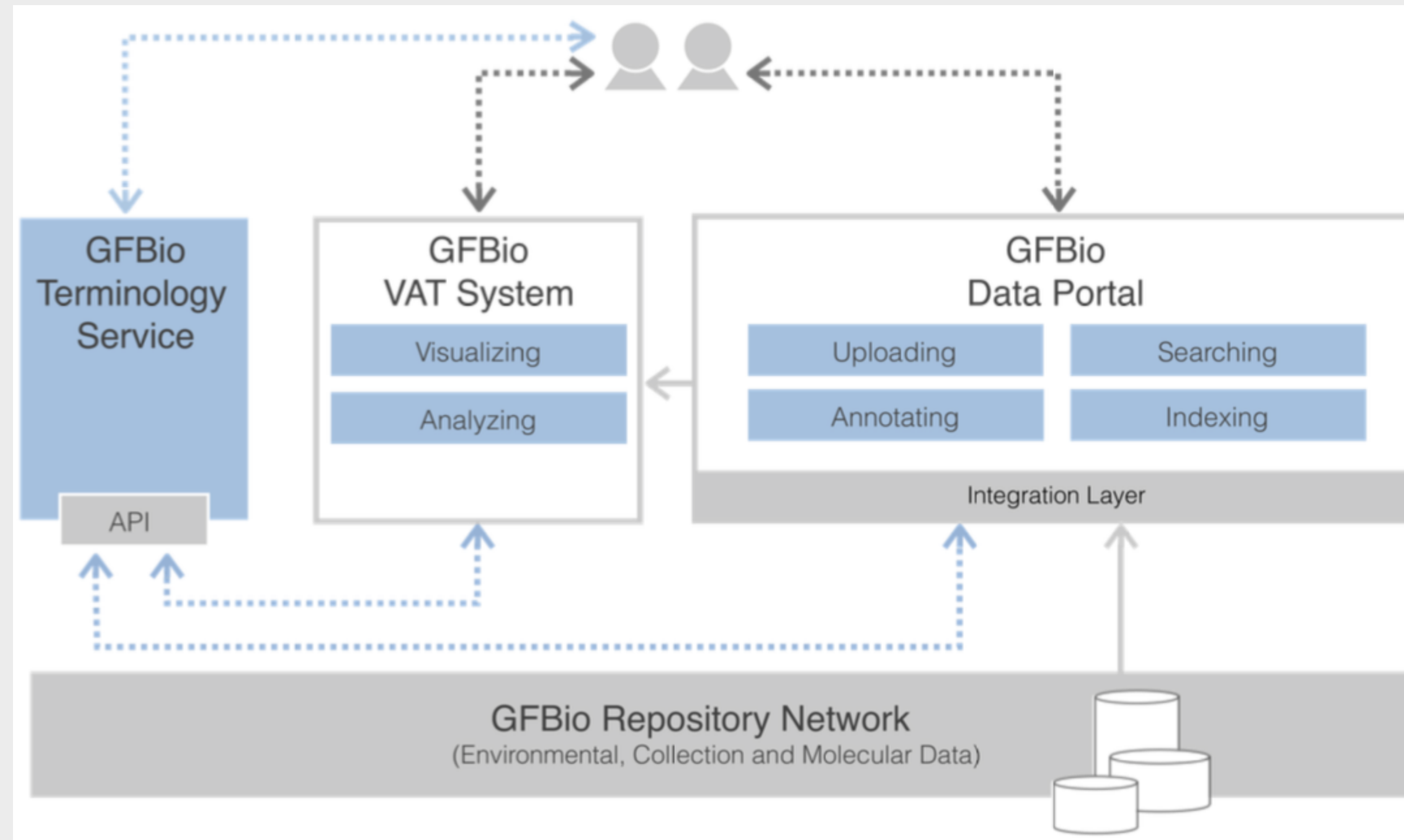
**Research Goal** » Allow experts to discover research data beyond disciplinary terminologies

**Context** » Research project with 20 partners; in close collaboration with Botanical Garden and Botanical Museum (BGBM)





# German Federation for Biological Data (GFBio)

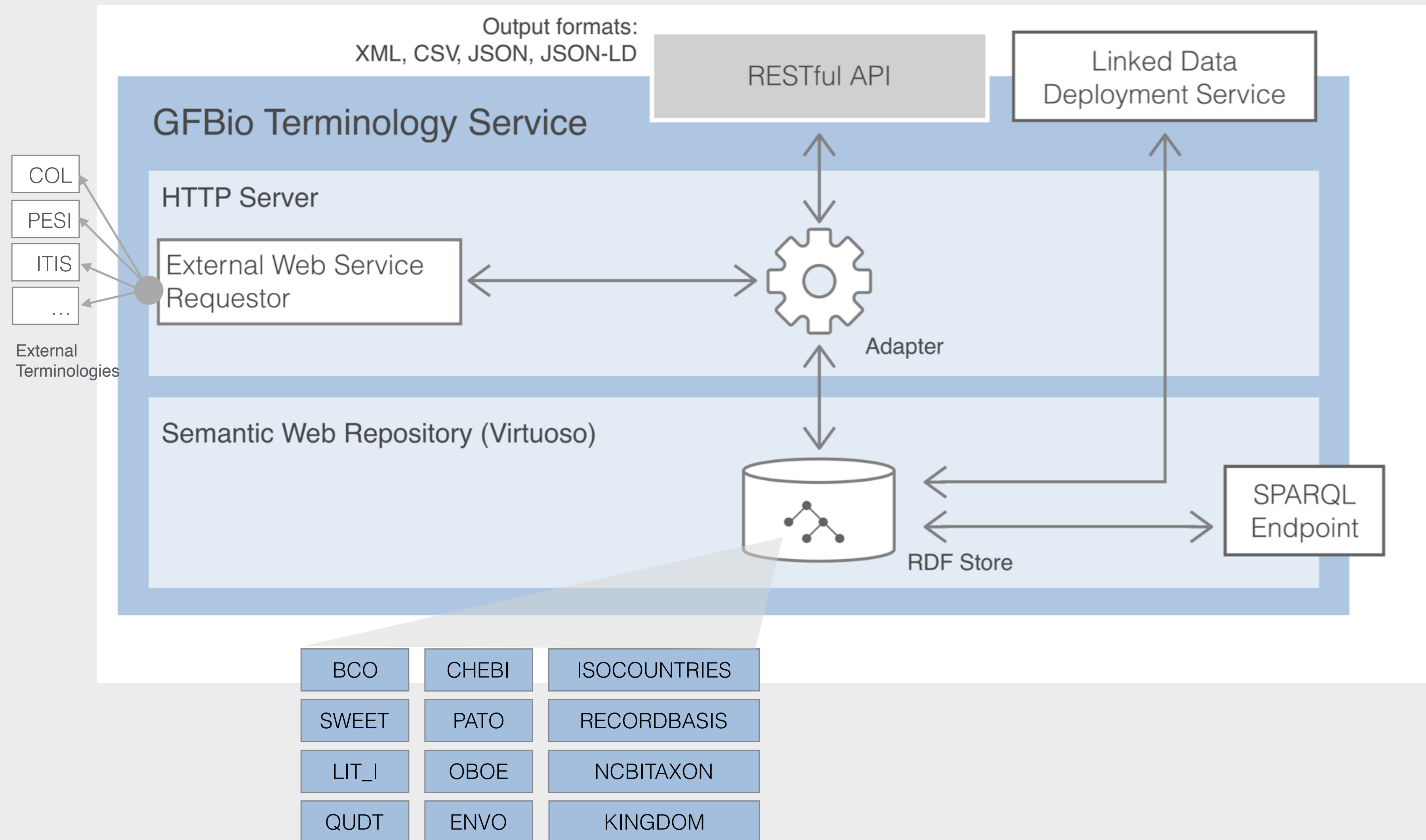




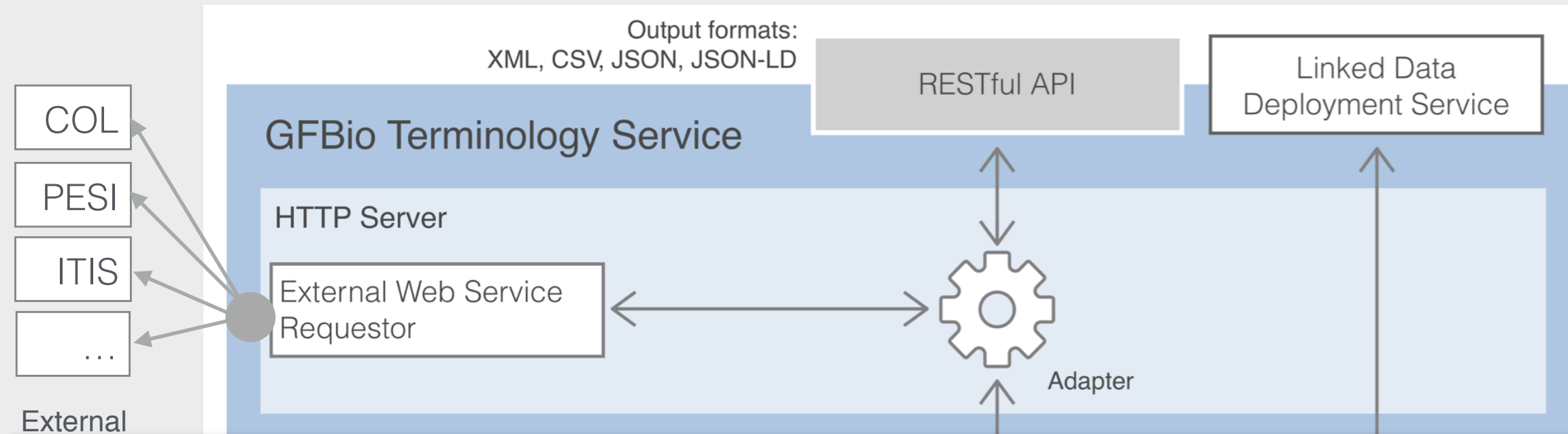
The screenshot displays the VAT web application interface. At the top, there is a navigation bar with tabs for 'Start', 'Operators', 'Project', and 'Debug'. Below this, a control panel includes a 'Layer' section with buttons for 'LINEAGE', 'EXPORT', 'RENAME', 'SYMBOLGY', and 'REMOVE'. A 'Zoom' section features 'LAYER' and 'MAP' search icons, along with '+' and '-' zoom controls. The 'Add Data' section offers options for 'ENVIRONMENTAL', 'ABCD ARCHIVES', 'CSV DATA/IMPORT', 'GFBIO BASKETS', and 'SP. DISTRIBUTION'. A 'Reference Time' section allows selection of year, month, and day (set to 1990, 9, 6) and hour, minute, and second (set to 20, 0, 0). The main area is a world map with data points, primarily concentrated in Africa and South America. A left sidebar shows layer management with 'Loxodonta af...' and 'GBIF - Loxod...' layers, each containing 'Point' and 'Vector' data types. At the bottom, there are tabs for 'Data Table' and 'Citation'.



# The Terminology Service



# Terminology Matching



Abbreviation	Name	# classes
NCBITAXON	National Center for Biotechnology Information (NCBI) Organismal Classification	1.507.562
CHEBI	Chemical Entities of Biological Interest Ontology	103.755
ENVO	Environment Ontology	6.190

## Matching Results

Compared ontologies:

<http://purl.obolibrary.org/obo/envo.owl>

<http://terminologies.gfbio.org/terms/cafe>

### Matching by lexical analysis of the labels

(10 possible matches found)

- Nodes [http://purl.obolibrary.org/obo/PATO\\_0000165](http://purl.obolibrary.org/obo/PATO_0000165) (*time*) and <http://terminologies.gfbio.org/terms/cafe#TimeContext> (*Time Context*) share part of the label
- Nodes [http://purl.obolibrary.org/obo/PATO\\_0000165](http://purl.obolibrary.org/obo/PATO_0000165) (*time*) and <http://terminologies.gfbio.org/terms/cafe#TimeRange> (*Time Range*) share part of the label
- Nodes [http://purl.obolibrary.org/obo/PATO\\_0000165](http://purl.obolibrary.org/obo/PATO_0000165) (*time*) and <http://terminologies.gfbio.org/terms/cafe#Time> (*Time*) have the same label
- Nodes [http://purl.obolibrary.org/obo/PATO\\_0000165](http://purl.obolibrary.org/obo/PATO_0000165) (*time*) and <http://terminologies.gfbio.org/terms/cafe#time> (*Time*) have the same label
- Nodes [http://purl.obolibrary.org/obo/PATO\\_0000165](http://purl.obolibrary.org/obo/PATO_0000165) (*time*) and <http://terminologies.gfbio.org/terms/cafe#Timezone> (*Timezone*) share part of the label
- Nodes [http://purl.obolibrary.org/obo/ENVO\\_01000820](http://purl.obolibrary.org/obo/ENVO_01000820) (*pedosphere*) and <http://terminologies.gfbio.org/terms/cafe#Pedosphere> (*Pedosphere*) have the same label
- Nodes [http://purl.obolibrary.org/obo/BFO\\_0000015](http://purl.obolibrary.org/obo/BFO_0000015) (*process*) and <http://terminologies.gfbio.org/terms/cafe#Process> (*Process*) have the same label
- Nodes [http://purl.obolibrary.org/obo/CHEBI\\_26020](http://purl.obolibrary.org/obo/CHEBI_26020) (*phosphate*) and <http://terminologies.gfbio.org/terms/cafe#Phosphates> (*Phosphates*) share part of the label
- Nodes [http://purl.obolibrary.org/obo/ENVO\\_01000620](http://purl.obolibrary.org/obo/ENVO_01000620) (*mesocosm*) and <http://terminologies.gfbio.org/terms/cafe#Mesocosm> (*Mesocosm*) have the same label
- Nodes [http://purl.obolibrary.org/obo/ENVO\\_01000621](http://purl.obolibrary.org/obo/ENVO_01000621) (*microcosm*) and <http://terminologies.gfbio.org/terms/cafe#Microcosm> (*Microcosm*) have the same label

Save Results

Back

Environment Technology	4.549
ions	5.307
	2.603
ogy for igneous rocks	1.870
y	538
ype	204
	127
cordBasis	14
ngdoms	9



**Designing** networked information systems

**Analysing** relationships between humans and machines in networked information systems

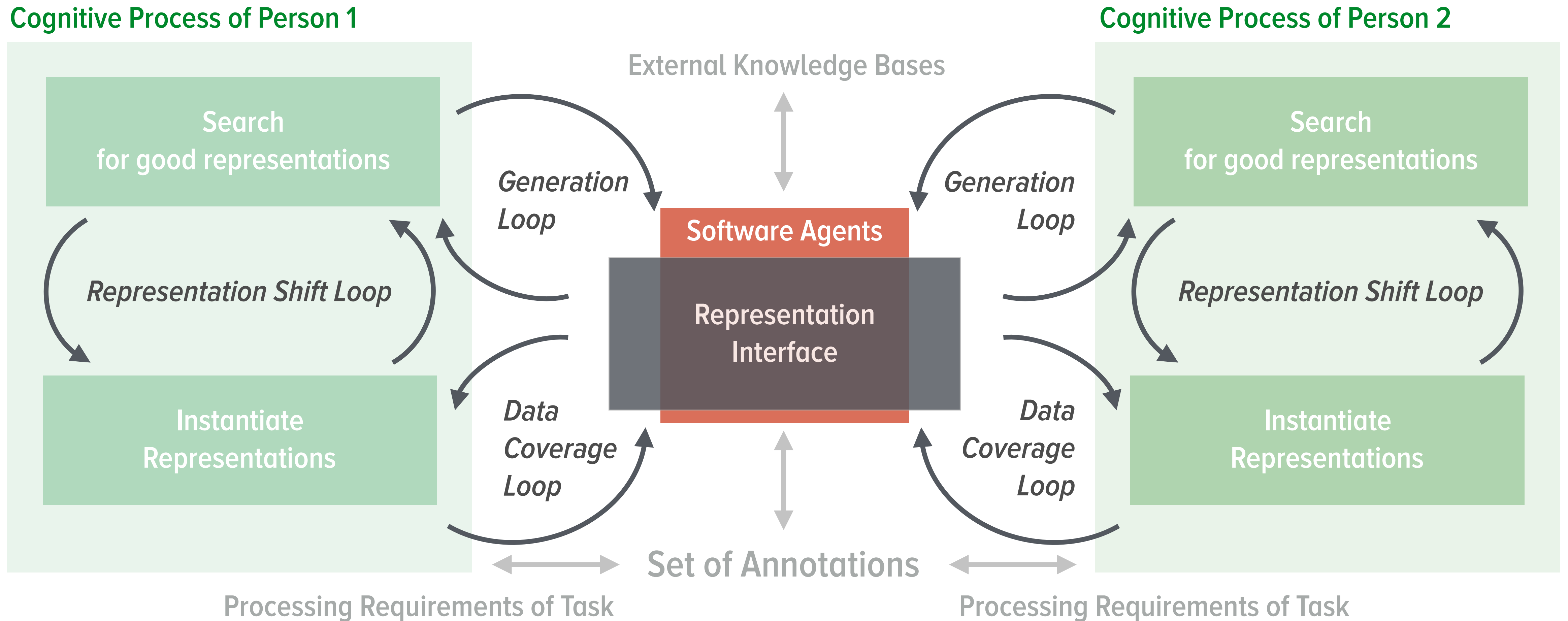
**Developing** coupled relationships between humans and machines in networked information systems

**Extending** networked information systems

# Collaborative, geographically distributed research processes

- Research Context** » Many collaborative research practices evolve around physical artefacts, e.g. in archeology, and are geographically distributed
- Research Goal** » Integrating physical and digital activities more smoothly and allow for computer-supported collaborative sensemaking by networked information systems
- Context** » Research project *Hybrid Knowledge Interaction* of the Cluster of Excellence *Image, Knowledge Gestaltung. An Interdisciplinary Laboratory*

# Computer-Supported Collaborative Sensemaking





# Conclusion & Future Work

- » The presented qualitative knowledge-based workflows build upon Licklider's vision of a “man-machine symbiosis” and Engelbart’s “augmenting human intellect” proposal
- » The technical and human components of a web information system are equally important, whereby the design of the interfaces and the human ability to use them should be coordinated.
- » The algorithmic component should be implemented in a way that it allows users to get an understanding of their functioning

The image displays a grid of 34 presentation slides, numbered 1 through 34. The slides are arranged in a 5x7 grid, with the last cell empty. The central slide, number 34, is titled "Discussion" and is highlighted with a yellow border. The other slides contain various content including diagrams, charts, and text related to human-machine collaboration and research.

# Discussion