

Challenges in disseminating Scientific and Cultural Resources in the Long Tail

Prof. Dr. Michael Granitzer
University of Passau
26.03., Hamburg, Germany

Agenda

- Our Content Distribution Process today
 - The Long Tail of Information
 - Popular vs. Scholarly
 - The Filter Bubble – Rich get Richer – Popularity ranking
- EEXCESS – personalised just-in-time recommendation of long tail content
 - Federated Recommendation
 - User and Usage Mining
 - Discovery via visualisations
- Summary

Motivation

Availability of Digital Resources

- Vast amounts of scientific resources digital available
- Content distribution processes optimized towards highly popular, commercial media content
 - Niches do play less of a role
 - Scientific content **not main stream**
 - Remains hidden in the **Deep Web**
 - Hard to **discover** unknown items
- Lets have a look in detail, why that is the case

ZBW
Leibniz-Informationszentrum
Wirtschaft
Leibniz Information Centre
for Economics



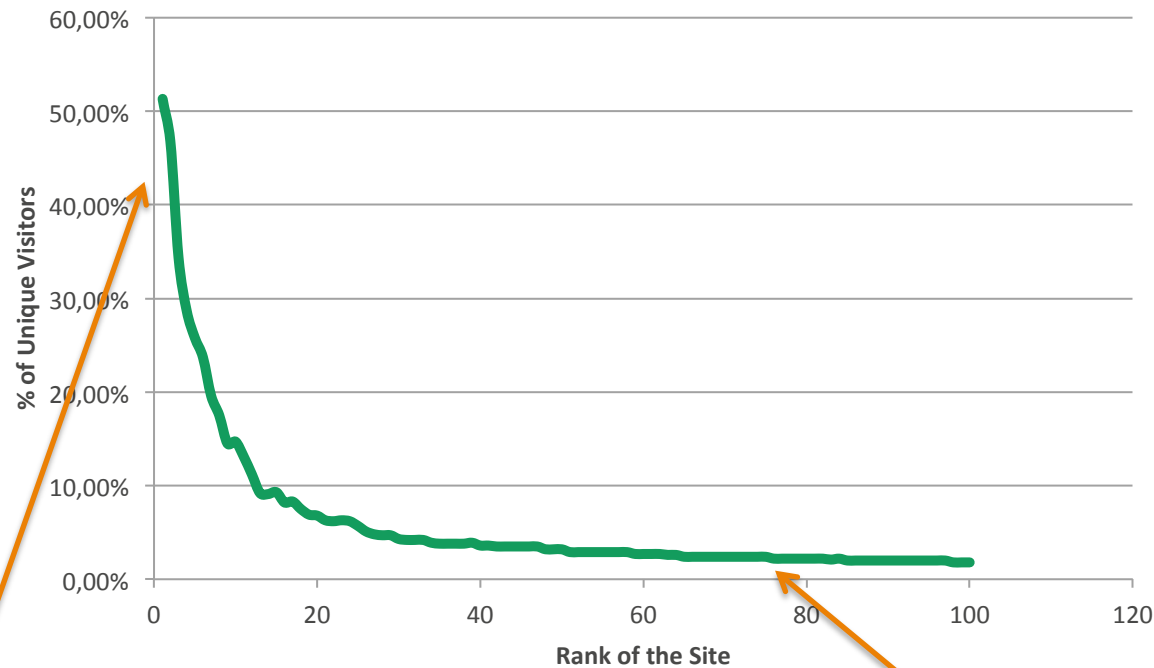
MENDELEY



Challenge

The long tail of websites

Reach of Sites on the Internet



Google

YouTube



Search Engine Optimization
Social Media Marketing etc.



ZBW

MENDELEY

Archäologie und Museum
Bildungs-, Kultur- und Sportdirektion
Kanton Basel-Landschaft

wissenmedia
verlag
digital
service
international

Collections
Trust

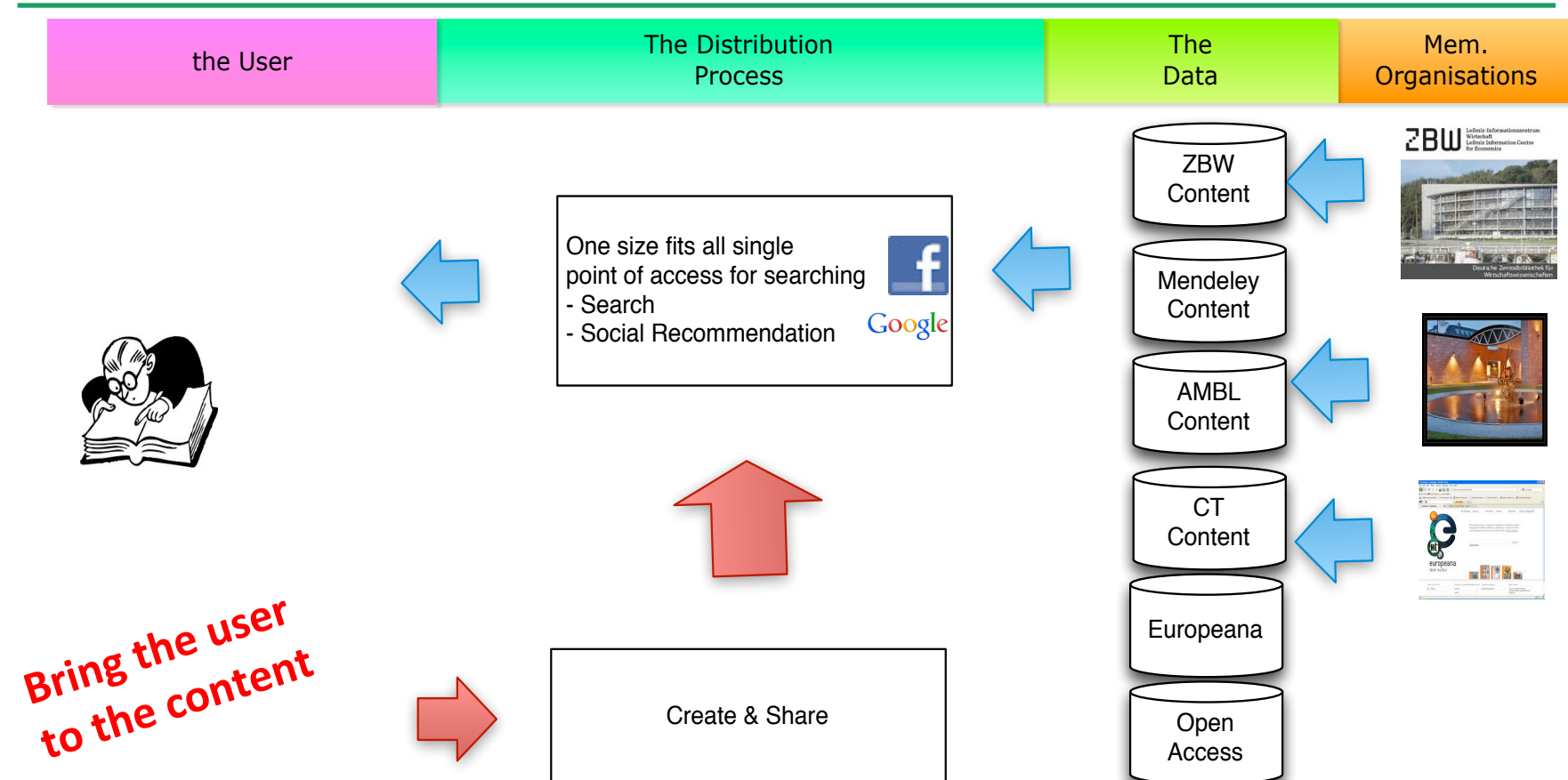
bit
best in training

EEXCESS



Challenge

The content distribution process today

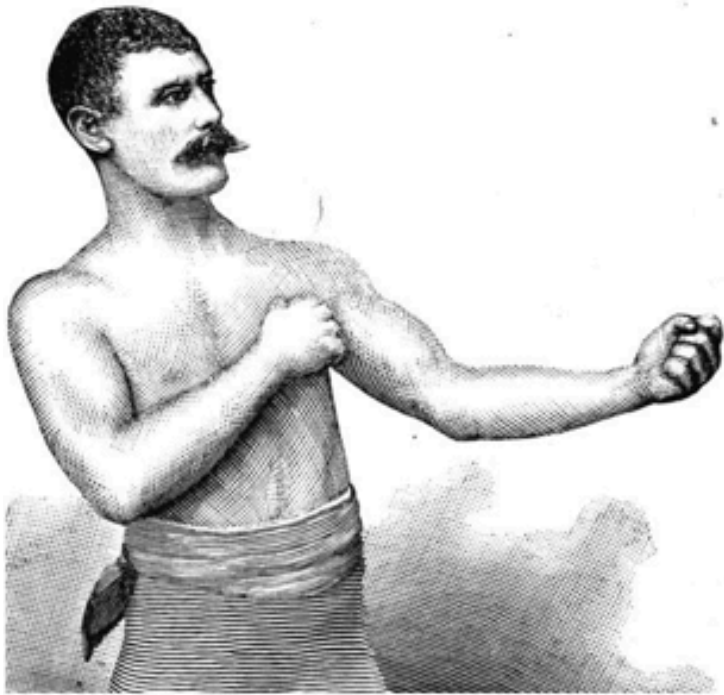


- **Discovering** non-main stream cultural, scientific or educational content is challenging
- **Expertise** of memory organisation remains **untapped**, since not part of the distribution process
- **Users are disconnected** from the experts

Challenge

Scholarly vs. Popular Resources

SCHOLARLY VERSUS POPULAR



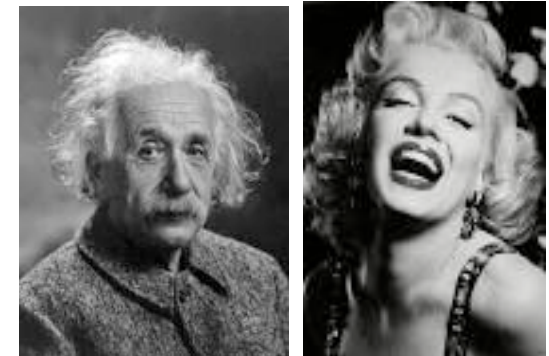
http://southwestern.edu/su_blogs/library/files/2012/02/scholar-v-pop-flierTZ.jpg

Challenge

What is wrong with popularity?

Phenomena

- Rich-get-Richer effects
 - Hard to become rich when starting in the long tail
- Potential Filter Bubble
- Popularity can make rankings unfair



Challenge

Is popularity based ranking the best we can do?

Experiment by Salgankik, Dodds, and Watts

- Download site for 48 obscure music songs
 - User get feedback on the popularity of a song
 - Hidden to user: 8 different copies of the system
 - ➔ 8 very different market shares
 - 9th System: no user feedback
 - ➔ Less variation in market shares of different songs
- ➔ **Popularity reinforce rich-get-richer phenomena**



DISCOVERING SCHOLARLY & CULTURAL CONTENT IN THE LONG TAIL

Project Details

- EU FP7 funded Integrated Project
- 10 Partners
 - R&D Partners
 - Memory Organisations



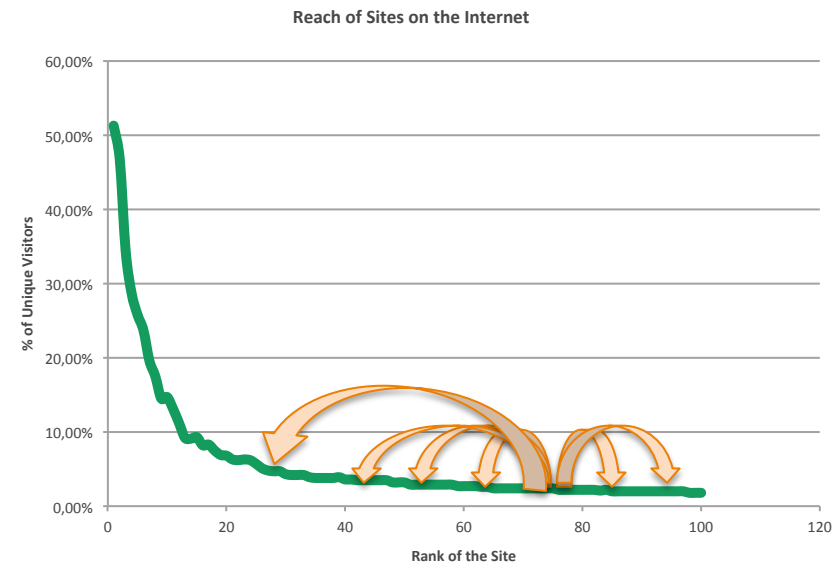
***Vision: Bring the user to the content, not
the content to the user***

A Quick Introduction

Idea

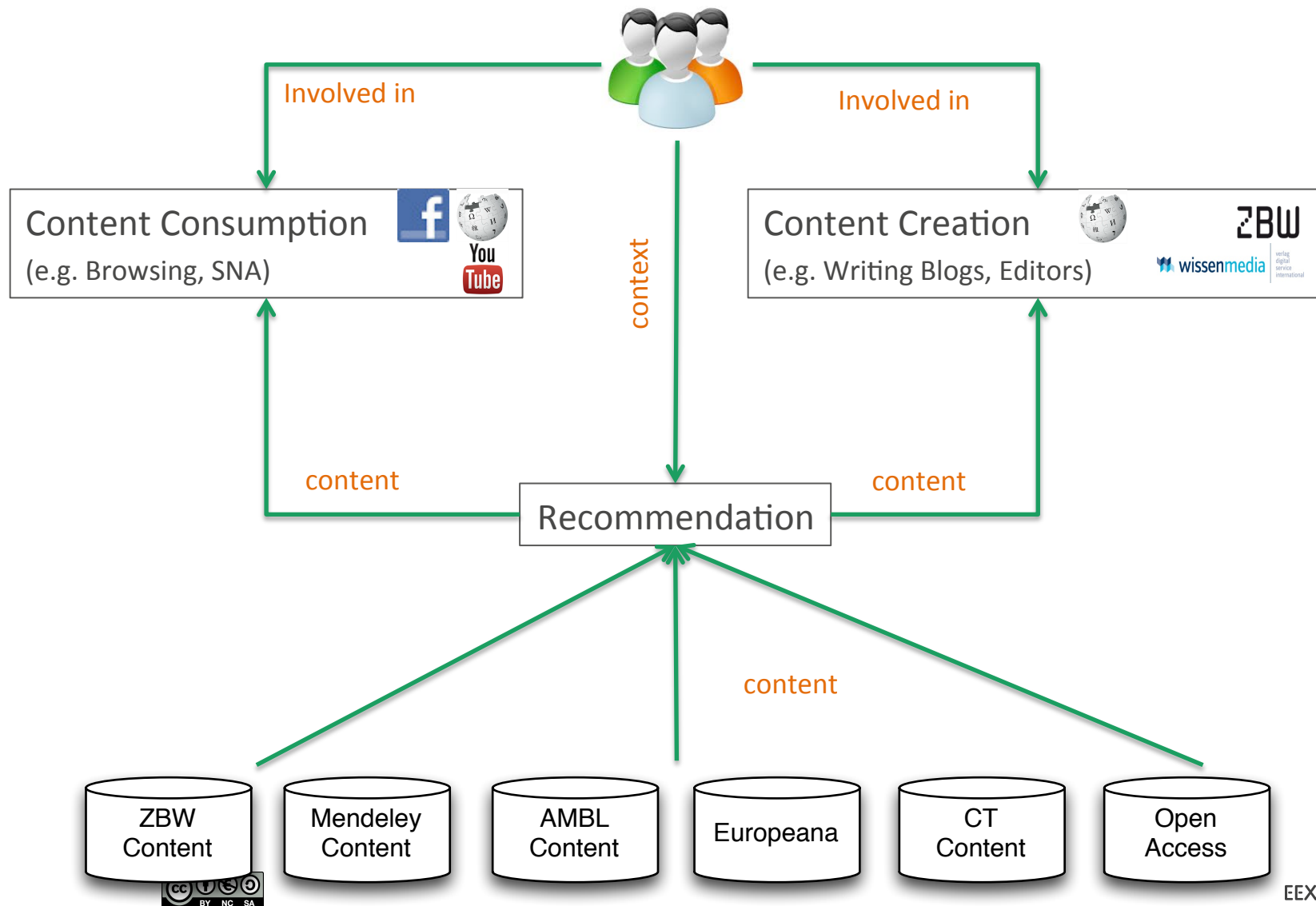
„Bring the content to the user, not the user to the content“

- Inject cultural and scientific content into existing web channels
 - Websites (Wikipedia, etc.)
 - CMS/LMS
 - Social media channels (Twitter, etc.)
 - Support “head-channels” as well as tail-channels
- Gather user and usage feedback such that memory organisations can learn how resources are distributed and used



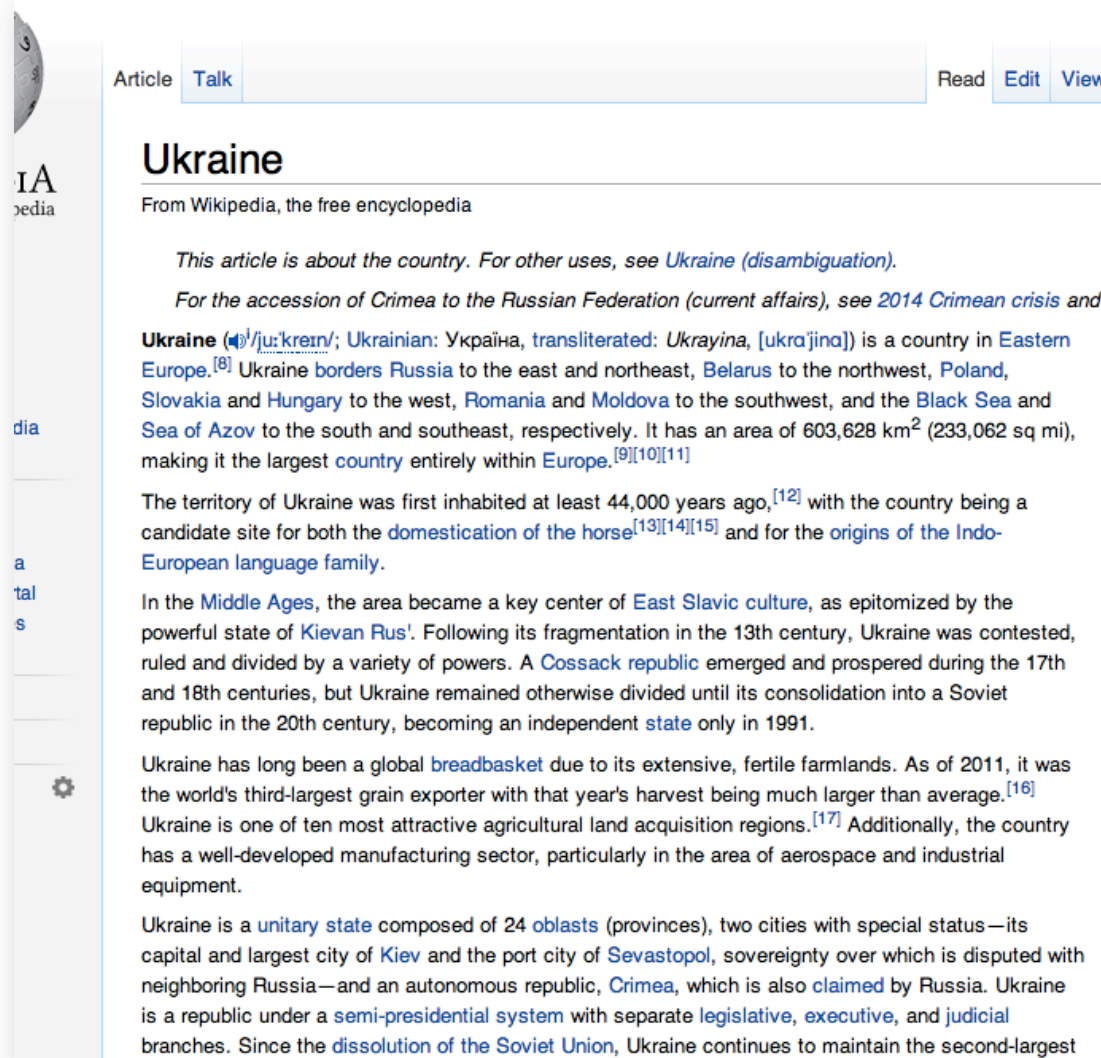
Requires high-quality, context dependent
injections of resources
(aka personalised recommendations)

Approach Overview



Approach

Example Wikipedia



The screenshot shows the Wikipedia article for Ukraine. At the top, there's a navigation bar with 'Article' and 'Talk' tabs, and buttons for 'Read', 'Edit', and 'View'. The title 'Ukraine' is prominently displayed. Below the title, it says 'From Wikipedia, the free encyclopedia'. A disclaimer follows: 'This article is about the country. For other uses, see [Ukraine \(disambiguation\)](#).' Another line of text mentions 'For the accession of Crimea to the Russian Federation (current affairs), see [2014 Crimean crisis](#) and [2014 Crimean status](#).' The main text begins with 'Ukraine (ⁱˈjuːˈkreɪn/; Ukrainian: Україна, transliterated: *Ukrayina*, [ukrɑˈjinɑ]) is a country in Eastern Europe.^[8] Ukraine borders Russia to the east and northeast, Belarus to the northwest, Poland, Slovakia and Hungary to the west, Romania and Moldova to the southwest, and the Black Sea and Sea of Azov to the south and southeast, respectively. It has an area of 603,628 km² (233,062 sq mi), making it the largest country entirely within Europe.^{[9][10][11]}

The territory of Ukraine was first inhabited at least 44,000 years ago,^[12] with the country being a candidate site for both the domestication of the horse^{[13][14][15]} and for the origins of the Indo-European language family.


In the Middle Ages, the area became a key center of East Slavic culture, as epitomized by the powerful state of Kievan Rus'. Following its fragmentation in the 13th century, Ukraine was contested, ruled and divided by a variety of powers. A Cossack republic emerged and prospered during the 17th and 18th centuries, but Ukraine remained otherwise divided until its consolidation into a Soviet republic in the 20th century, becoming an independent state only in 1991.

Ukraine has long been a global breadbasket due to its extensive, fertile farmlands. As of 2011, it was the world's third-largest grain exporter with that year's harvest being much larger than average.^[16] Ukraine is one of ten most attractive agricultural land acquisition regions.^[17] Additionally, the country has a well-developed manufacturing sector, particularly in the area of aerospace and industrial equipment.

Ukraine is a unitary state composed of 24 oblasts (provinces), two cities with special status—its capital and largest city of Kiev and the port city of Sevastopol, sovereignty over which is disputed with neighboring Russia—and an autonomous republic, Crimea, which is also claimed by Russia. Ukraine is a republic under a semi-presidential system with separate legislative, executive, and judicial branches. Since the dissolution of the Soviet Union, Ukraine continues to maintain the second-largest


Approach

Example Blog Writing

 Add New Post















Jupiter and the Galilean Moons

Permalink: <http://astrojoanne.wordpress.com/?p=2>

 Add Media

Visual

Text

B *I*              

Last night I have discovered Jupiter standing unusually high in the night sky. The explanation for this phenomenon: Jupiter has reached opposition (directly opposite the Sun, as seen from our beloved Earth) the other day. This year's opposition is particularly favourable for us here on the Northern hemisphere, with Jupiter climbing high in the Southern sky at midnight.

Even the four largest moons of Jupiter, called Galilean Moons (discovered by Galileo Galilei in 1610) were to be seen. The derive their names from the lovers of Zeus in the Greek mythology: Io, Europa, Ganymede and Callisto.

Path: p

Recommended Links [Apply all](#)

W Galilean moons

W Jupiter

W Earth

W Sun

W night sky


Publish

Save Draft


Preview


Status: **Draft** [Edit](#)


Visibility: **Public** [Edit](#)


 Publish immediately [Edit](#)


Publicize: Not Connected [Show](#)


RecommenderSystems 

Jupiter 

The Galilean Moons 

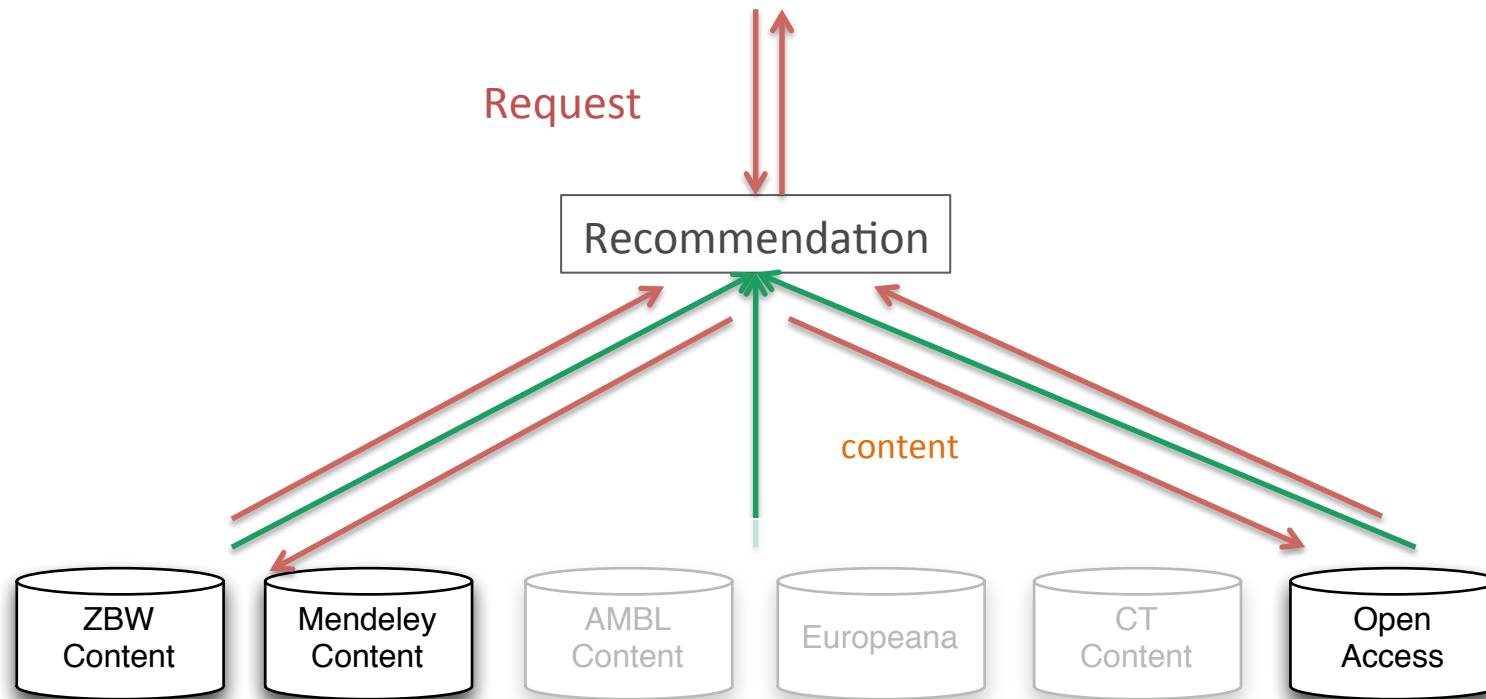






Approach

Federated Recommendation

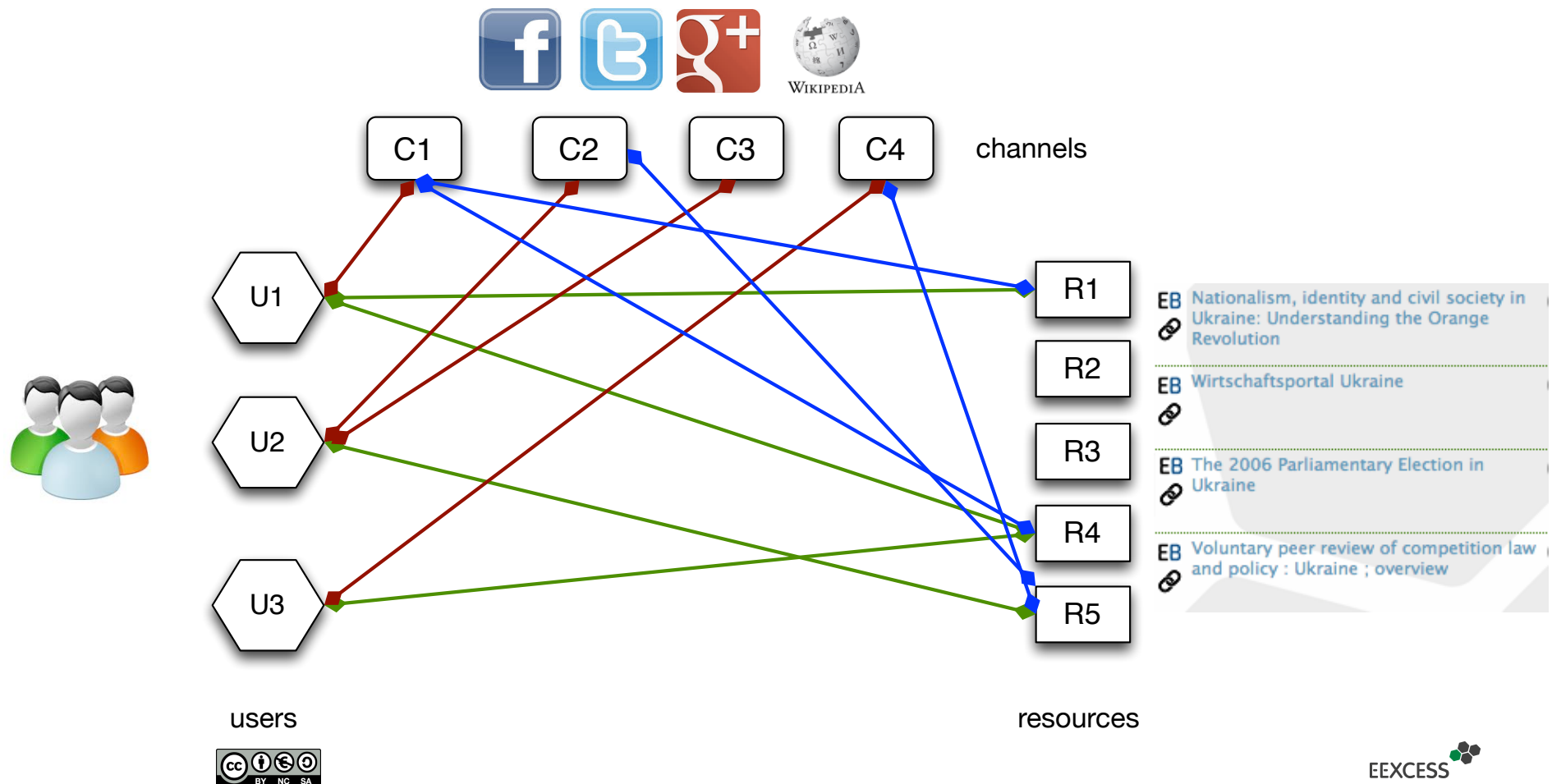


1. Select appropriate source
2. Forward request
3. Gather and merge results
4. Return results

Approach

User and Usage Mining

- Obtaining knowledge over user and usage while retaining privacy



Approach

Intelligent User Interfaces

- Beyond 10 results through visualisation
 - Recommendation Results
 - Relationships among items
 - Metadata
- Empower user to discovery interesting resources in the long tail content

Approach

Intelligent User Interfaces

- Facetted browsing of large result lists for discovery purpose

Women in Workforce Search Query Results (100) : Filter

<p>2011-12² 1996² 1988² 1999² 2006² 2011-10³ 1998² 2012-02³ 2011⁸ 2005³ 1978⁴ 2012¹¹ 2008⁸ 1982¹ 2003⁵ DATE 2013⁵ 2007⁴ 2009¹² 1983⁴ 2004² 2010⁹ 1981² 1984² 2010-05⁴ 1997² 1973² 1990² 2009-09-07²</p> <p>usbk⁷ repec base⁹ SOURCE econis⁵⁵</p>	<p>National Advis..² Sloan School o..² Economics Rese..³ Amsterdams Ins..⁵ eSocialSciences⁵</p> <p>INSTITUTION unknown⁷² Directorate fo..³ Segawa, Yoshik..² Federal Reserv..² Department of ..² Department of ..² Australien⁸ Arbeitskräfte⁹ Germany³ United States¹¹ Women¹⁴ Gender¹¹ Women workers²⁷ SUBJECT Weibliche Arbe..²⁷ unknown¹⁷ USA¹² Frauen¹⁰ Frauenarbeit¹¹ Workforce⁹</p>	<p>The journal of..² The Bangladesh..² Renewing inter..³ Handbook of re..³ Japanese and A..³</p> <p>ISPARTOF unknown⁵⁷ Handbook on wo..³ Women and work..² International ..² The journal of..²</p> <p>article⁴⁸ other TYPE book⁵³</p>	<p>Williams, Mela..⁵ Tijdens, Kea⁵ Klaveren, Maar..⁵ Wagner, Joachim⁷ PERSON Pfeifer, Chris..⁷ Martin, Nuria ..⁵</p> <p>Collection of ..² Amtsdruckschri..³ Government doc..³ Sammelwerk² Article in book¹¹ Bibliographie² Graue Literatur¹⁶ Aufsatz im Buch¹¹ Article in jou..¹⁶ Aufsatz in Zei..¹⁶ Fallstudie² Case study² TYPE_GENRE unknown⁵³ Non-commercial..¹⁶ Thesis³ Arbeitspapier¹⁵ Working Paper¹⁵ Congress Report³</p>
---	---	---	---

Women in the workforce
article - Broderick, Elizabeth

Myths in the media : how the news media portray women in the workforce
article - Dunn-Jensen, Linda M., Stroh, Linda K.

Myths in the media : how the news media portray women in the workforce
article - Dunn-Jensen, Linda M., Stroh, Linda K.

Rural women workforce in India
book - Delhi : B. R. Publ. Corp. - Nath, Madhuri

Globalization and women in the Japanese workforce
book - London [u.a.] : RoutledgeCurzon - Bishop, Beverley

Technological change, organizational change and skill requirements : impacts on women in the workforce
article - Betcherman, Gordon, Lauzon, Darren, Leckie, Norman

CC BY NC SA

Approach

Intelligent User Interfaces

- Query Summary

The screenshot displays a query summary interface. The main area features a network graph with nodes and edges. A central node is labeled "world war". Below it, two nodes are labeled "WoWa" and "WoWa WW". The "WoWa" node is connected to "world war" and has a red bookmark icon. The "WoWa WW" node is also connected to "world war". Below these nodes are two yellow labels: "women warfare" and "women warfare world war".

On the right side, there is a details panel with the following sections:

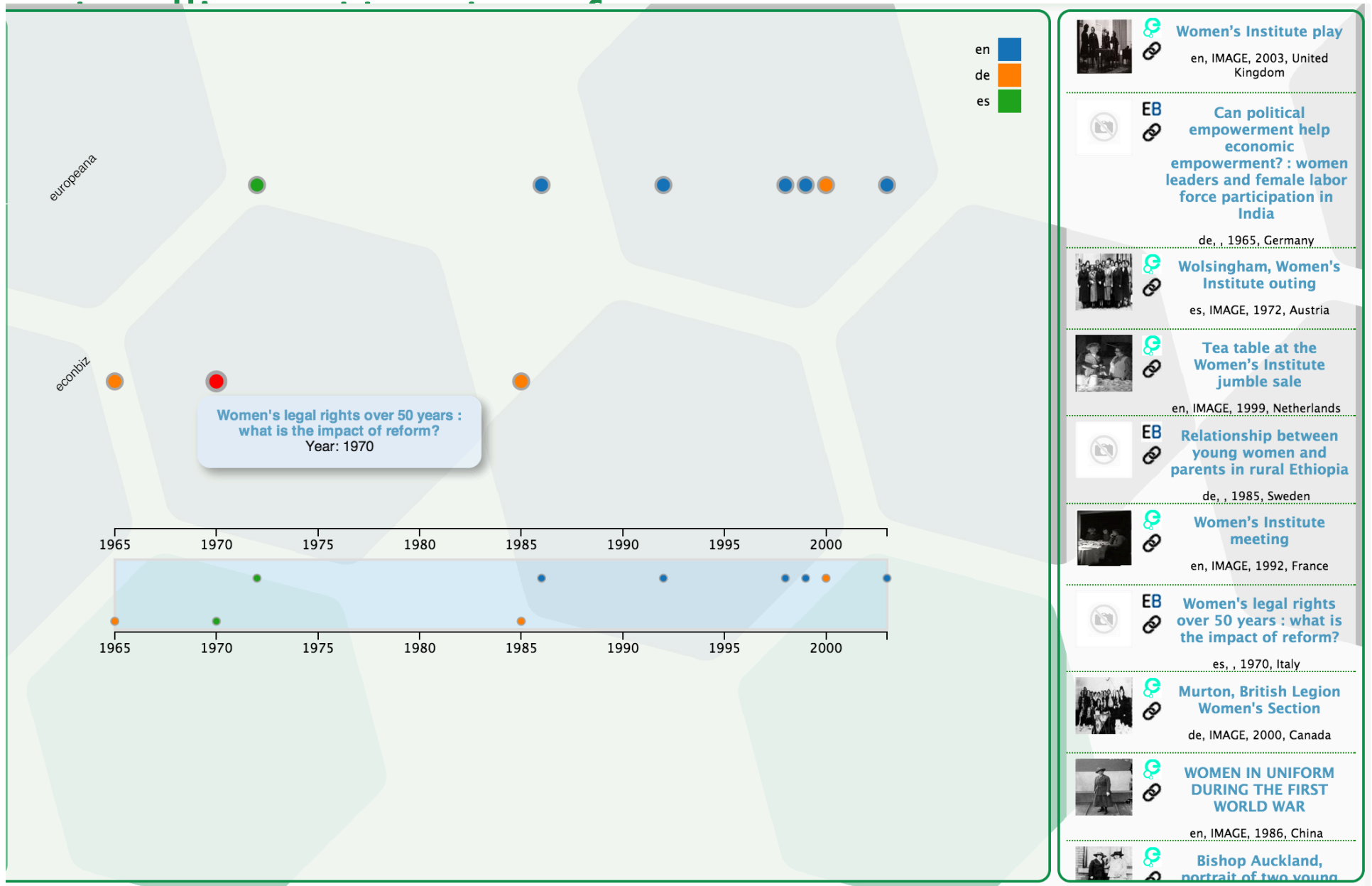
- details** (selected) and **export metadata**
- clear data** button
- title:**
- link:**
- image:**
- id:**
- facets:**
- language:**
- partner:**
- provider:**
- type:**
- year:**

At the bottom, there is a status bar with the following information:

- redraw** button
- Results per keywords:** 5
- Last** 10
- th keywords of the searchgraph.**
- Number of keywords in searchhistory:** 13.

At the bottom left, there is a Creative Commons BY-NC-SA license logo. At the bottom right, there is a "RECESS" logo.

Approach



Approach

Evaluation in 3 Testbeds

Educational Support

- Cultural/scientific resources injected to LMS
- Pupils, Teachers

Scholarly Communication

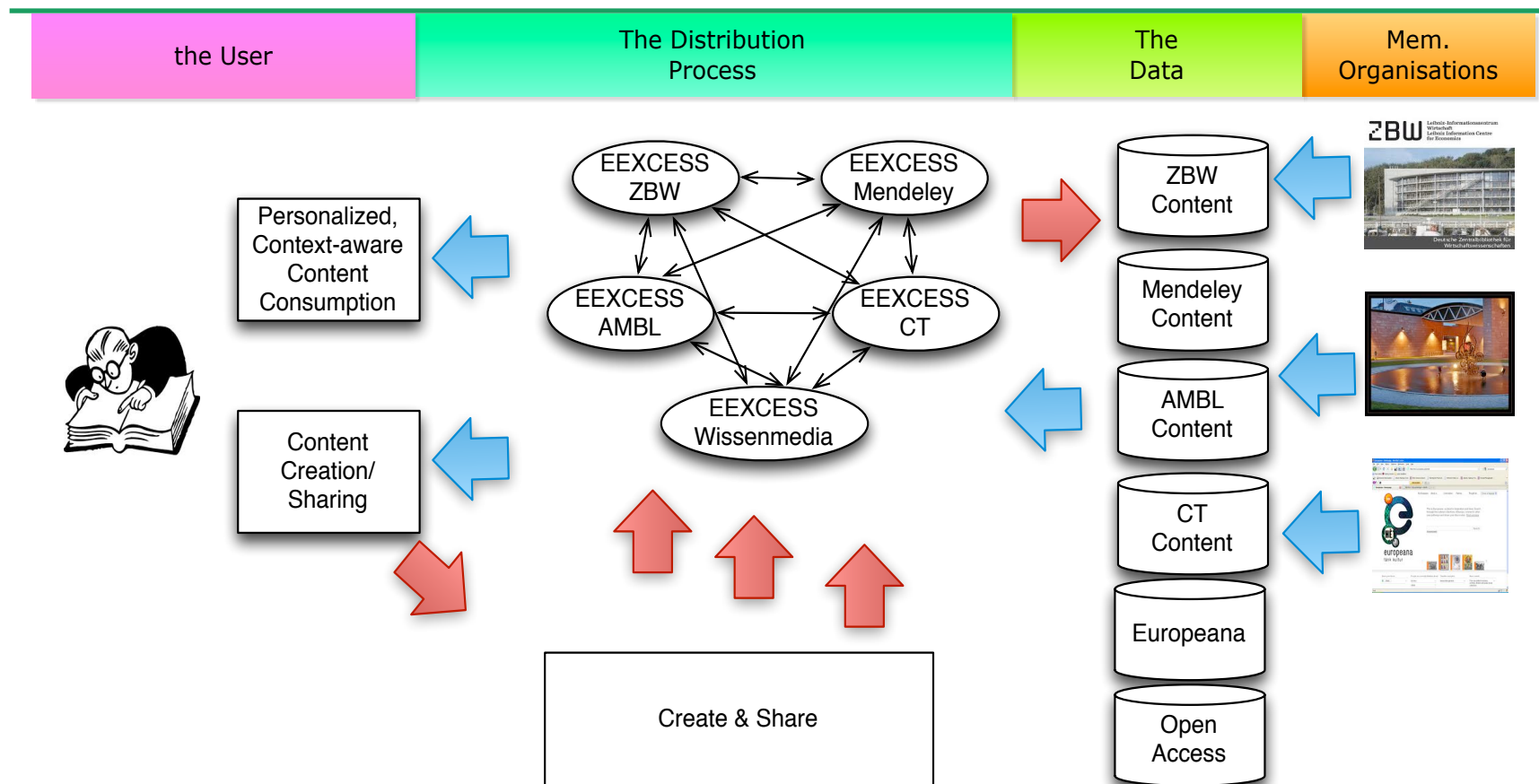
- Interconnecting cultural and scientific resource
- Students, Lecturers, Researchers

General Public Education

- Disseminate cultural/scientific content to the general public
- Regionally interested users, culturally interested users, media consumers

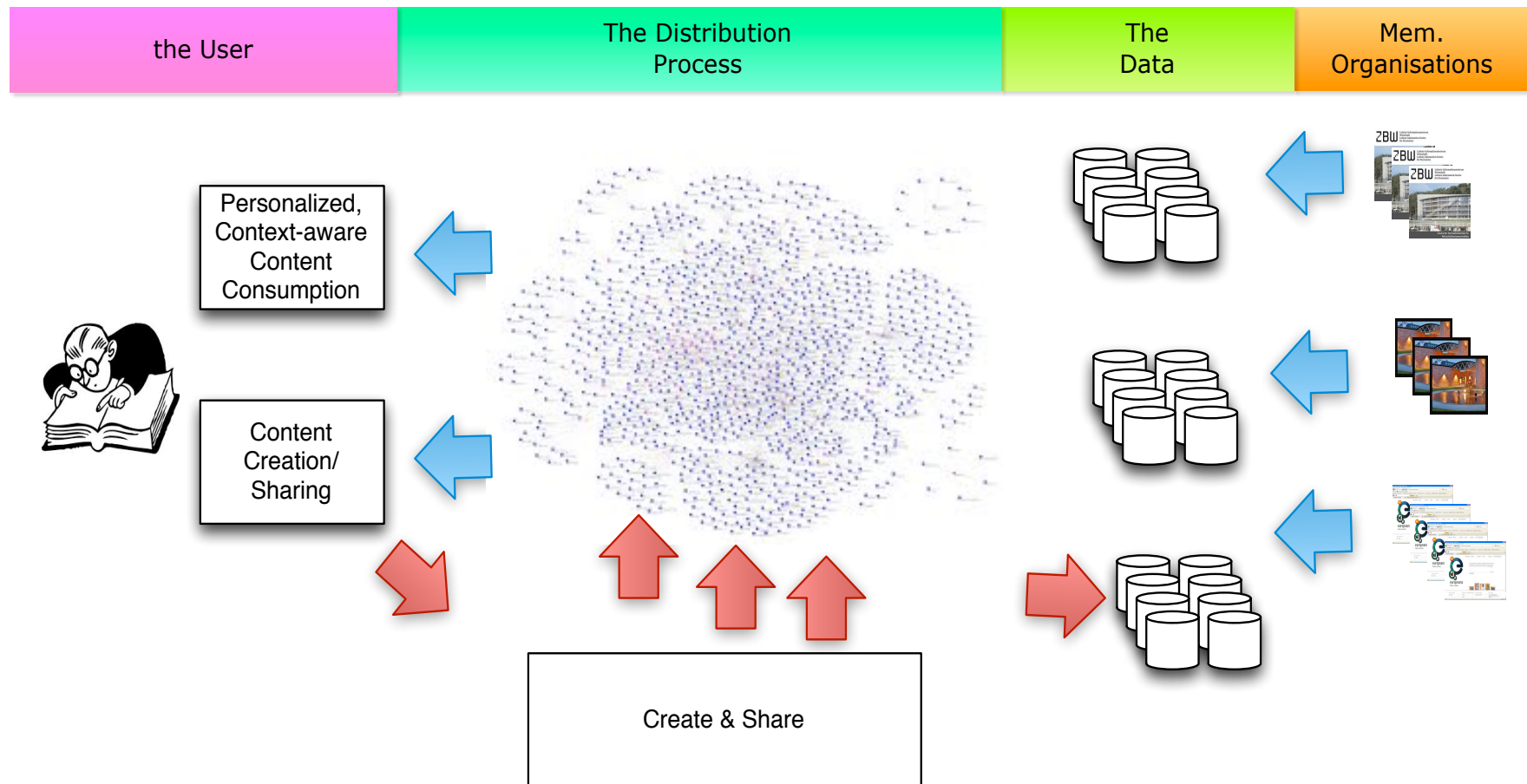
Outcome

New content distribution processes



- **Empower** memory organisations to shape their content distribution process and bring in their expertise
- **Enrich** existing content consumption and creation processes (e.g. blogging)
- **Personalized** recommendation of relevant content based on user and usage analysis
- **Transparent and trusted** distribution processes

Future Vision: Scale beyond EEXCESS



- **Adoption** by a large number of different memory organisations
- **Enriching** the Web with cultural, educational and scientific content

Summary

Popular vs. Scholarly

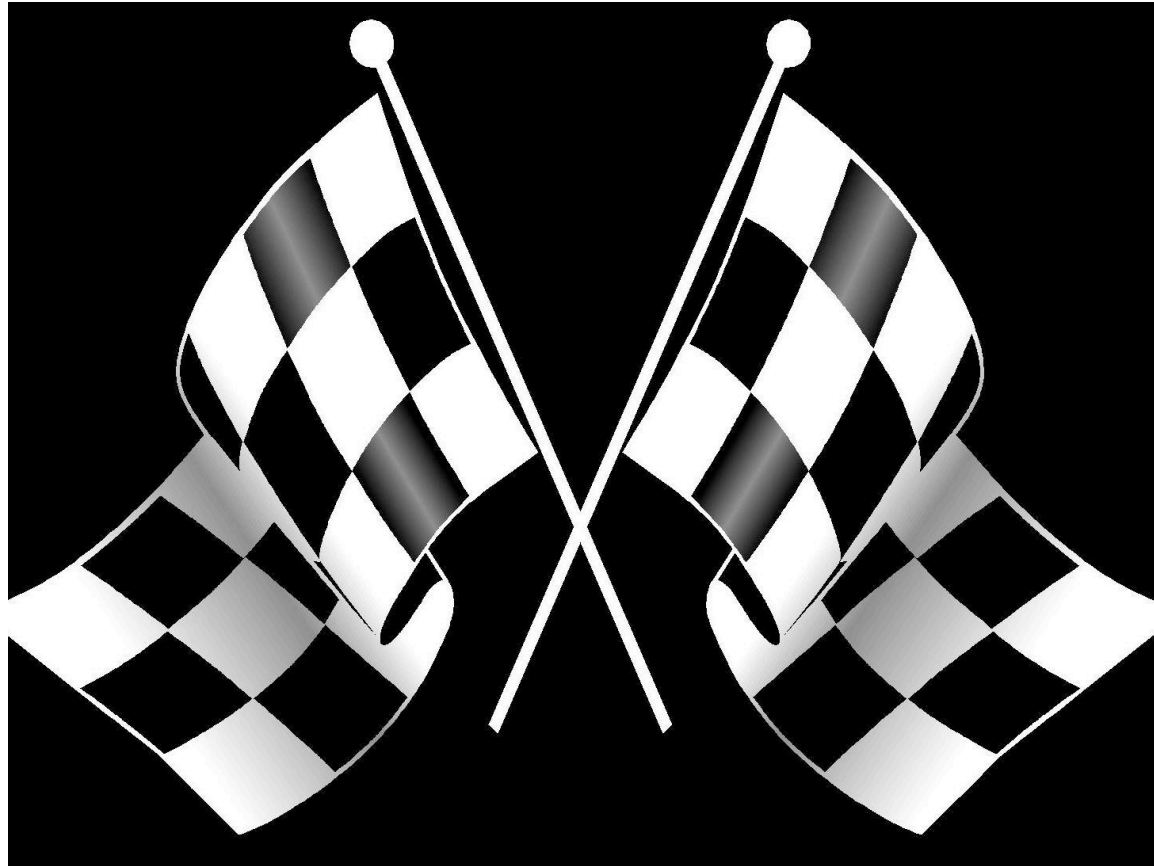
- Long tail content matters
- Challenge in overcoming the content distribution process today

Bring the content to the user

- Federated Recommendation
- User and usage mining while preserving privacy
- Intelligent user interfaces to foster discovery

Connecting users and experts (again)

If you are interested to integrate your portal, do not hesitate to contact us.



In theory there is no difference between theory and practice. In practice there is. (Yogi Berra)