Boost your eHumanities and eHeritage research with Research Infrastructures (PARTHENOS Webinar)

Darja Fišer, Ulrike Wuttke

▶ **To cite this version:**

Darja Fišer, Ulrike Wuttke. Boost your eHumanities and eHeritage research with Research Infrastructures (PARTHENOS Webinar). 3rd cycle. Germany. 2018. cel-01784097

HAL Id: cel-01784097
https://hal.science/cel-01784097

Submitted on 3 May 2018

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution 4.0 International License

PARTHENOS
Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

Boost your eHumanities and eHeritage research with Research Infrastructures

(Darja Fišer & Ulrike Wuttke)

A Webinar of the PARTHENOS eHumanities and eHeritage Series

Pictures clockwise :
https://upload.wikimedia.org/wikipedia/commons/1/17/Headline_of_the_New_York_Times_June-29-1914.jpg CC0
https://pxhere.com/de/photo/1163578 CC0
https://de.wikipedia.org/wiki/Wikipedia:Wien/Archiv/2015#/media/File:20140415_TFM_Edit-a-thon_Wien_0301.jpg CC BY-SA 3.0 at
https://de.wikipedia.org/wiki/Pariser_Manuskripte_(Leonardo_da_Vinci)#/media/File:Paris_Manuscript_F_Fol_18v19r.jpg CC0

Webinar start: **24.04.2018, 13:00 CEST**
Moderation: Ulrike Wuttke
(University of Applied Sciences Potsdam)

FH;P Fachhochschule Potsdam University of Applied Sciences

Join the conversation on Twitter:
#PARTHENOSWebinar
@Parthenos_EU

PARTHENOS
Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

FH;P Fachhochschule Potsdam University of Applied Sciences

PARTHENOS eHumanities and eHeritage Webinar Series

Date and Time: 24.04.2028, 13:00 - 14:00 CEST

Moderation: Ulrike Wuttke
(University of Applied Sciences Potsdam)

Join the conversation on Twitter:
#PARTHENOSWebinar
@Parthenos_EU

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.
Graphic: Otto, the Open Access Otter CC-BY 4.0, Katja Diederichs

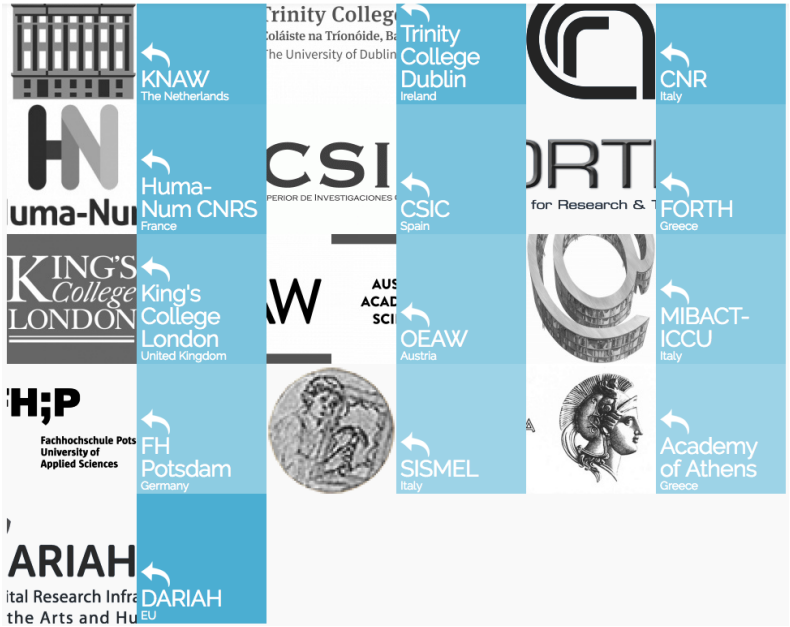
Some Housekeeping…

- As a participant you will be **muted** and not be seen throughout the webinar.
- For **questions** and **remarks** please use the **chat**. For questions, please use the word **question**.
- Your questions are going to be answered by the trainers during / after the presentation.
- If you have **sound problems**, please test your technical settings. You have to click on the **speaker symbol** on top to to be able to hear (it has to be green).
- **Help us improve**: Follow up e-mail with a link to a short **feedback survey**.

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Some words about PARTHENOS…

- **PARTHENOS** stands for: Pooling Activities, Resources and Tools for Heritage eResearch Networking, Optimization and Synergies
- PARTHENOS is a **Horizon 2020** project with the aim to **strengthen the cohesion of Heritage related E-research**
- Running time: 1 May 2015 - 30 April 2019
- PARTHENOS has 16 partners from 9 European countries
- PARTHENOS Coordinator: PIN Scrl (Italy)
- The PARTHENOS webinar series is a **cross PARTHENOS training effort**



Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Which country are you participating from?

Cartography OpenStreetMap, licensed as CC BY-SA (copyright page).

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

PARTHENOS
Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

FH;P Fachhochschule Potsdam University of Applied Sciences

Boost your eHumanities and eHeritage research with Research Infrastructures

Trainers

Darja Fišer
(University of Ljubljana, Slovenia)

Ulrike Wuttke
(University of
Applied Sciences
Potsdam, Germany)

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Boost your eHumanities and eHeritage research with Research Infrastructures

Darja Fišer & Ulrike Wuttke

Unless otherwise stated this presentation is licensed under a Creative Commons Attribution 4.0 International License.

What we will do in this webinar:

- highlight advantages of developing research questions in an „infrastructural way“
- exemplify aspects of finding, working, and contributing data to Research Infrastructures (RIs) and using Virtual Research Environments (VREs) and tools
- present Citizen Science approaches for Humanities and Heritage research

Poll: What is for you the first step in a research project? How do you start a project?

What is developing research questions in humanities?

- Gesta Danorum
  - written in about 1200 in High Latin by Danish historian Saxo
  - describes in 16 books the period of time from King Dan to Canute VI of Denmark
  - traditionally divided into 2 parts: B1-9 Norse mythology & B10-16 describing Christianity

Research question: “Is it B9 or B10 that represents the transition from the heathen to the Christian period in the Gesta Danorum?”

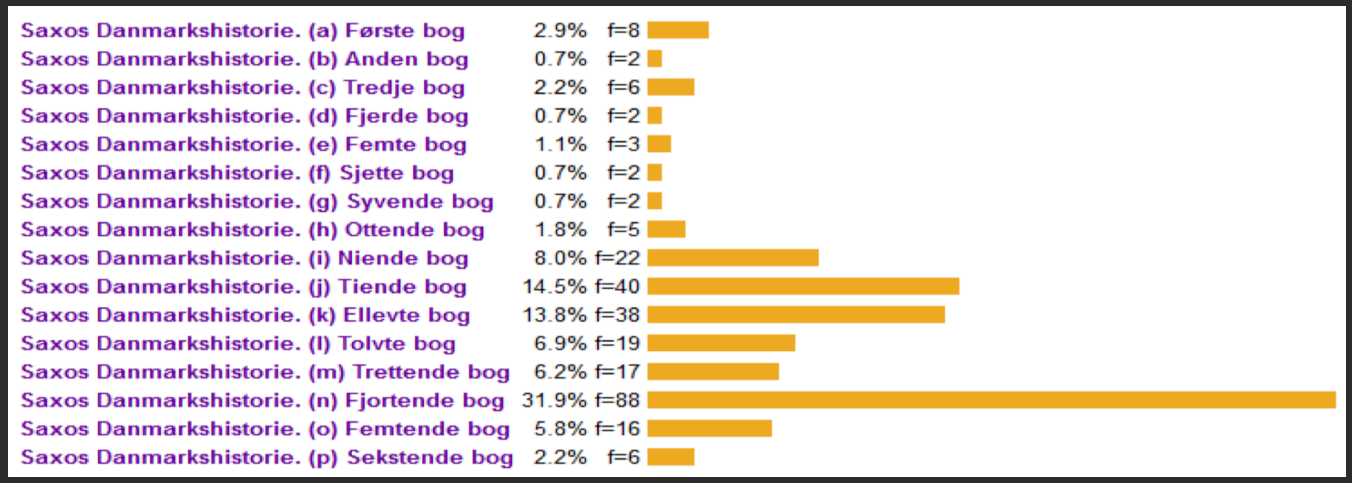
Source: Clavs Povlsen https://www.clarin.eu/showcase/gesta-danorum

What is developing research questions in humanities?

- frequent use of elements from the Christian register of words indicates that the topic described has to do with Christianity in this case the transition from heathen to the Christian period in the Gesta Danorum

helgen (eng: saint -> noun singular)
helgene (eng: saints -> noun plural
kristent (eng: Christian -> adjective neuter singular)
kristen (eng: Christian -> adjective common singular)
kristendom (eng: Christianity -> singular)
synd (eng:sin -> noun singular)

- frequency per thousand of Christian elements for B8 is significantly lower compared to B9&10 (0.25, 2.0, and 2.0)
- the change in terms of Christian language usage lies between B8&9



Source: Clavs Povlsen https://www.clarin.eu/showcase/gesta-danorum

Finding, working, and contributing data to RIs

CLARIN Resource Families
www.clarin.eu/resource-families

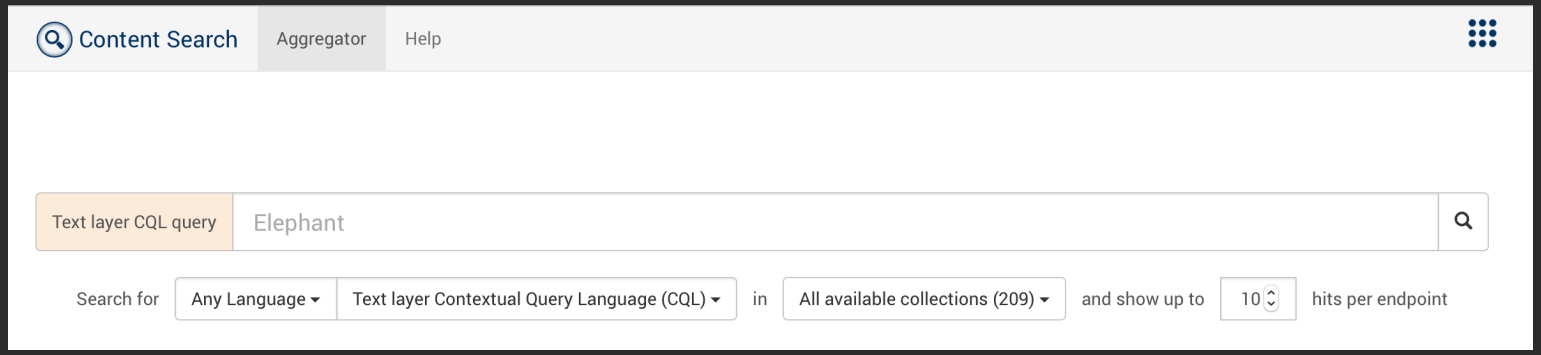
- parliamentary corpora
- newspaper corpora
- parallel corpora
- CMC corpora
- L2 learner corpora

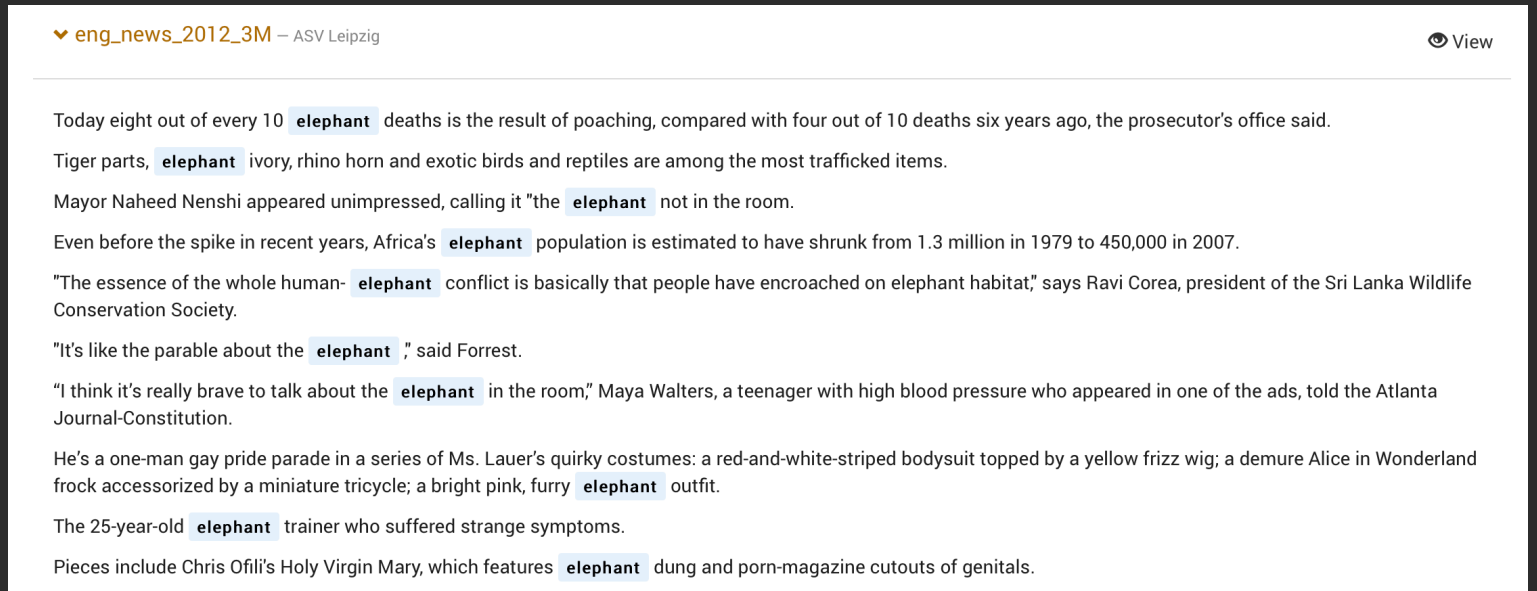
Slovenian parliamentary corpus SlovParl 2.0 **Size:** 3.2 million tokens **Annotation:** tokenised, PoS-tagged, lemmatised **Licence:** CC-BY	Slovenian	The corpus contains Slovenian parliamentary debates from 1990 to 1992. The corpus is available for download from the CLARIN.SI repository and through noSketch Engine. For the relevant publication, see Pančur and Šorn (2016).	Concordancer Download
Riksdag's Open Data **Size:** 1.25 billion tokens **Annotation:** tokenised, lemmatised **Licence:** CC-BY	Swedish	The corpus contains Swedish parliamentary debates from 1971 to 2016. It is annotated with Sparv. The corpus is available for download from Språkbanken (all entries with "Riksdag's Open Data" in the subtitle) and through the concordancer Korp. For the relevant publication, see Borin et al. (2016).	Concordancer Download

Finding, working, and contributing data to RIs

Federated Content Search

https://spraakbanken.gu.se/ws/fcs/2.0/aggregator/

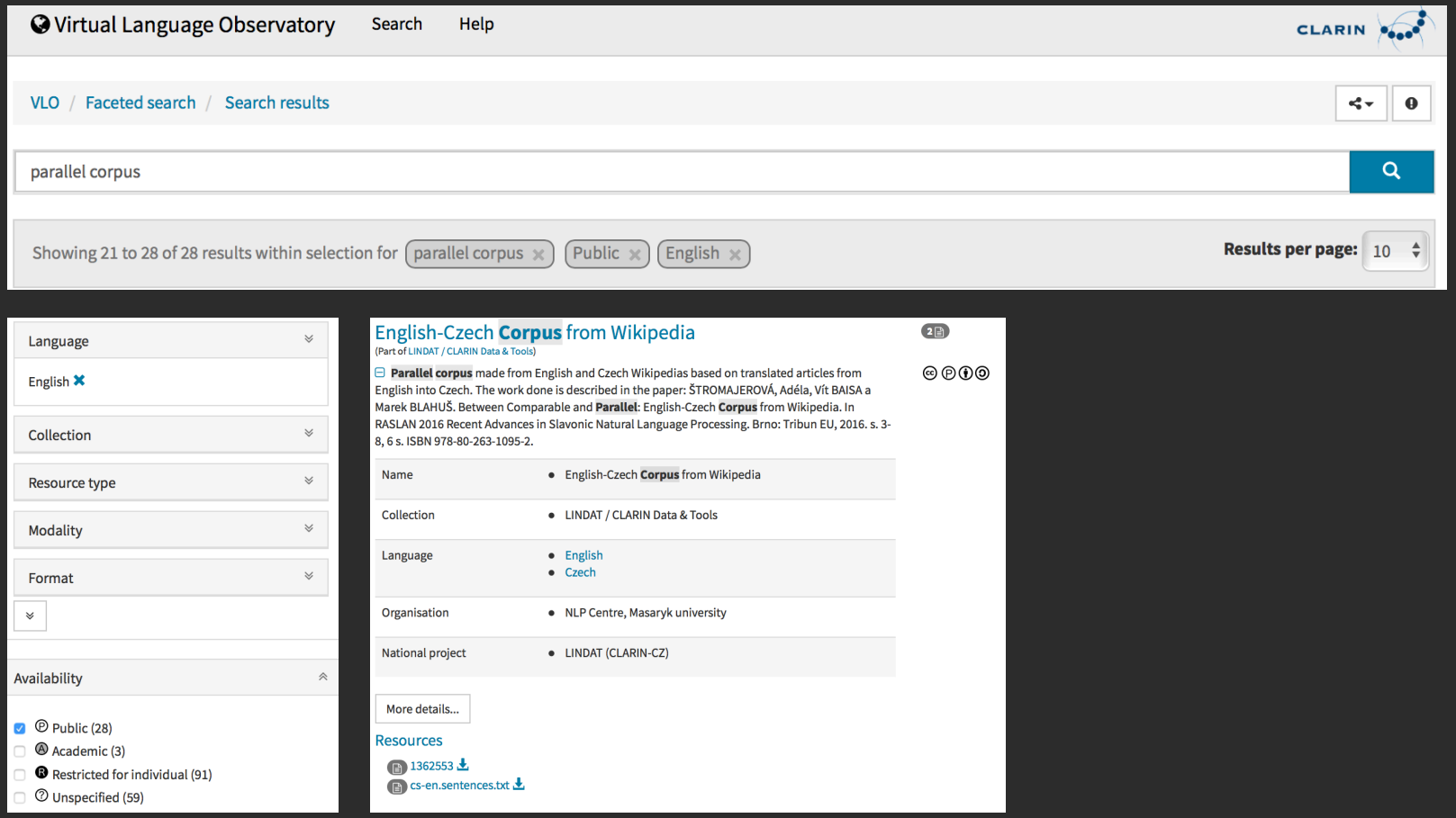




Finding, working, and contributing data to RIs

Virtual Language Observatory

https://vlo.clarin.eu/



Using VREs and tools

CLARIN Language Switchboard

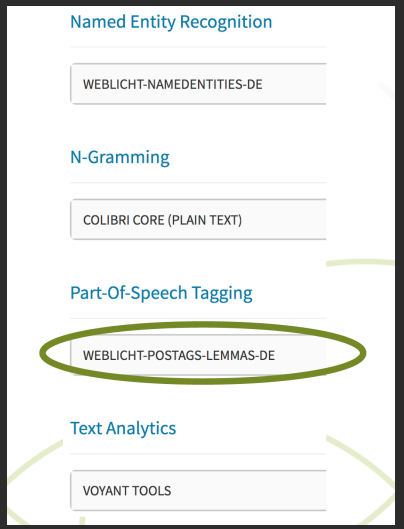
https://switchboard.clarin.eu

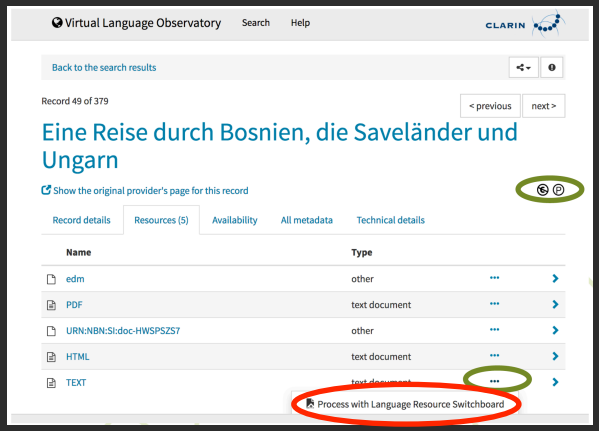
- digitized by Slovenian National University Library
- about Bosnia
- written in German

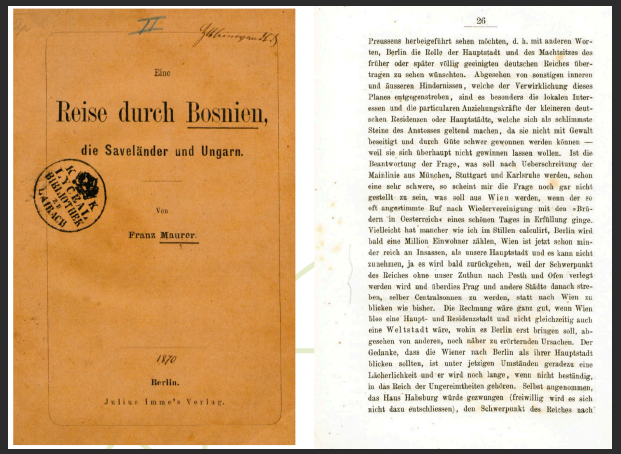
[8719 tokens]

tokens	POStags	lemmas	namedEntities
den	ART	d	
Schwerpunkt	NN	Schwerpunkt	
des	ART	d	
Reiches	NN	Reich	
nach	APPR	nach	
Ungarn	NE	Ungarn	B-GPE
zu	PTKZU	zu	
verlegen	VVINF	verlegen	

[etc, up to 144.000 tokens]







WebAnno

Where to find/store data/tools/expertise/training:

European RIs for SSH & CH

RIs set up under the auspices of ESFRI, each based on national consortia of universities, libraries, museums, archives etc.:

cessda Consortium of European Social Sciences Data Archives | CLARIN | DARIAH-EU | E-RIHS EUROPEAN RESEARCH INFRASTRUCTURE FOR HERITAGE SCIENCE | European Social Survey | SHARE SURVEY OF HEALTH, AGEING AND RETIREMENT IN EUROPE

In addition a number of past or ongoing EC supported Infrastructure Projects, such as

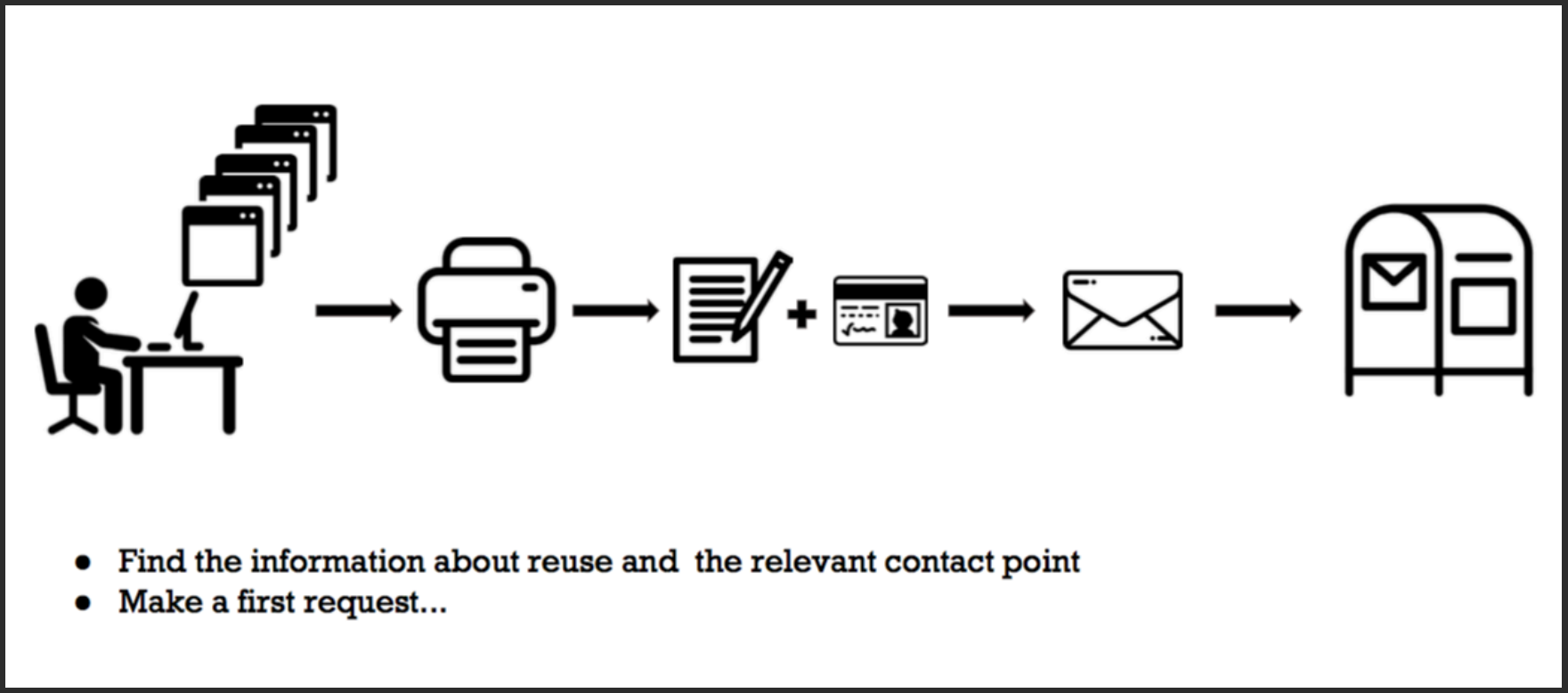
EHRI EUROPEAN HOLOCAUST RESEARCH INFRASTRUCTURE | ARIADNE www.ariadne-infrastructure.eu | CENDARI COLLABORATIVE EUROPEAN DIGITAL & ARCHIVAL RESOURCE INFRASTRUCTURE | CHARISMA | PARTHENOS | DCH RP DIGITAL CULTURAL HERITAGE ROADMAP FOR PRESERVATION

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Source: Slide from Stefan Schmunk, & Steven Krauwer. (2018, March). Slides from "e-Humanities and e-Heritage Research Infrastructures: Beyond tools" (PARTHENOS eHumanities and eHeritage Webinar, Thursday, 22.02.2018, 11:00 – 12:00 A.M. CET). Zenodo. http://doi.org/10.5281/zenodo.1203335.

The Cultural Heritage Data Reuse Charter

- scenario: Researcher wants to use digital cultural heritage assets (=data) hosted by CHIs/datacentres/infrastructures for research
- at the moment, this is mostly far from easy to achieve
- users wish to re-use and share data = straightforward conditions of re-use



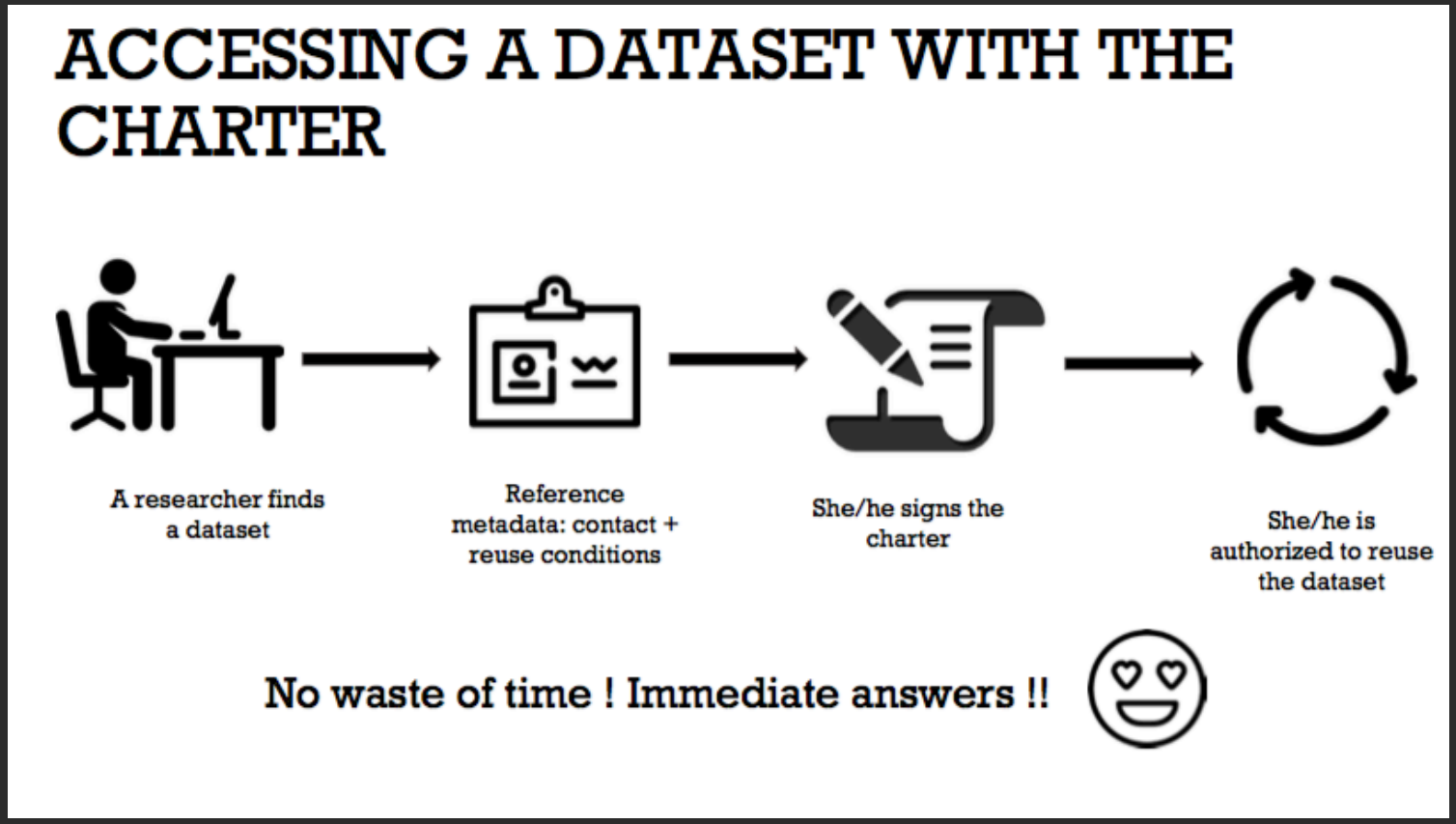
Picture from slides by Laurent Romary 'Access to Cultural Heritage Data: A Challenge for Digital Humanities', 09.04.2018, https://www.ifla.org/files/assets/academic-and-research-libraries/conferences/laurent_romary_data_reuse_charter.pdf

The Cultural Heritage Data Reuse Charter

- aim: Opening up Cultural Heritage data where possible and improve the interplay between scientific research and CHIs
- prerequisite: a non-restrictive, easy to understand and applicable general framework that provides incentives and best practices to all stakeholders about access, use, and reuse
- the **Cultural Heritage Data Reuse Charter** as a common online environment is currently under development with several supra-national infrastructures and projects involved (including CLARIN, DARIAH, PARTHENOS etc.)

Picture: ‘1413 Charter WillemVI aanleg Kostverloren HH Rijnland CHRT-0025.jpg’ Licence: Public Domain, source: https://commons.wikimedia.org/wiki/Category:Charters?uselang=de#/media/File:1413_Charter_WillemVI_aanleg_Kostverloren_HH_Rijnland_CHRT-0025.jpg

The Cultural Heritage Data Reuse Charter: Vision



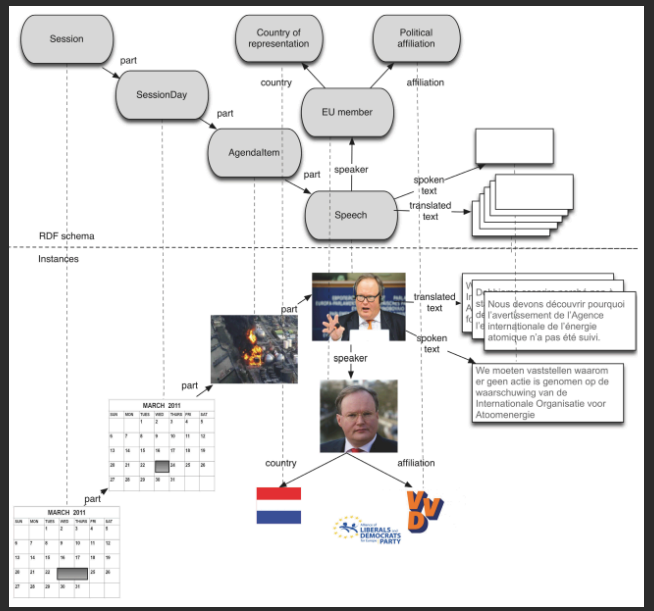
- go to the Charter mission statement
- read more in the PARTHENOS Training Suite about Managing Cultural Heritage Assets and the Cultural Heritage Reuse Charter

Picture from slides by Laurent Romary 'Access to Cultural Heritage Data: A Challenge for Digital Humanities', 09.04.2018, https://www.ifla.org/files/assets/academic-and-research-libraries/conferences/laurent_romary_data_reuse_charter.pdf

Talk of Europe (2014-2015)

- Erasmus University Rotterdam, University Amsterdam, DANS, Netherlands Institute for Sound and Vision
- goal: facilitate richer research of the European Parliament by turning the EP proceedings into Linked Open Data and organizing creative camps to spark collaborative development on the new database

- Linked Open Data
  - semantic: *Give me politicians who held a speech on 23 March 2011.*
  - linked: *Give me politicians from the Committee on Foreign Affairs who held a speech on 23 March 2011.*
  - open: *Give me politicians from the Committee on Foreign Affairs who held a speech on the topic of human rights 23 March 2011.*



Source: http://www.talkofeurope.eu/2014/01/about/

A Digital Database of Manuscripts and Intercultural Dialogue in Post-Conquest England

British Library, Additional MS 35283, fol. 75v, CC0.

- this Europeana Research Grant project draws on Europeana data to take a new approach to the status of French writing following the Norman Conquest of England in 1066
- aim: create and analyse an EDM-compliant digital database of all 958 manuscripts containing French literature from medieval England

"Digitisation of medieval manuscripts has been so important to medievalists in general. There are still many reasons to work with a manuscript in person, but among other benefits, the digitisation of medieval manuscripts allows us to visually compare manuscripts from different collections, to see new details on manuscripts and to access manuscripts remotely. The kind of project I'm working on now, which involves ***checking linguistic data of a large number of manuscripts, wouldn't be feasible without digitised manuscript collections****."*
Krista Murchison (Leiden University)

https://pro.europeana.eu/post/medieval-manuscript-databases-krista-murchison-europeana-research-grants-winner

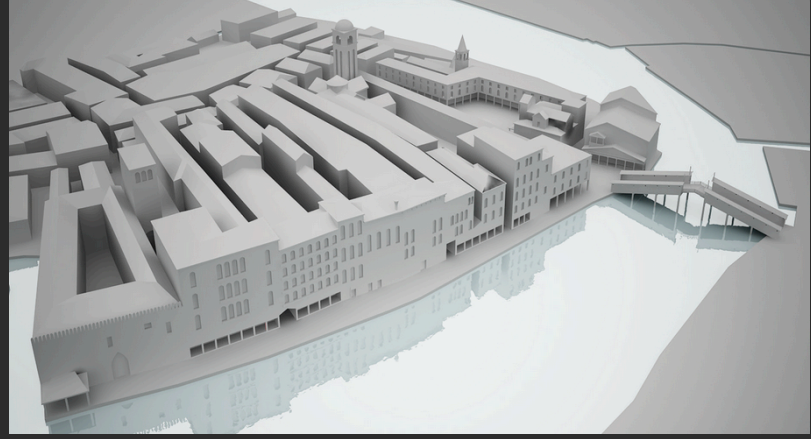
Sources: https://pro.europeana.eu/post/these-are-the-winners-of-the-europeana-research-grants-programme-2017, https://pro.europeana.eu/post/medieval-manuscript-databases-krista-murchison-europeana-research-grants-winner

The JANES Project 2014-2017

- University of Ljubljana, Jožef Stefan Institute
- goal: facilitate research of communication practices in social media by compiling a corpus of Slovenian UGC and developing tools and resources for processing Slovenian UGC
  - Results
    - Richly annotated corpora: user metadata (type, sex, region), text metadata (date, language, standardness level, sentiment), linguistic annotation (normalization, PoS-tagging, lemmatization)
    - Manually annotated corpora: for training tools (normalization, tagging, parsing), for linguistic analysis (abbreviation strategies, comma placement, code-switching, dialects, politicians)

Source: http://nl.ijs.si/janes/english/

Venice Time Machine



3D reconstruction of Rialto neighborhood in 1500 ab. based on the documents of Venetian archives, © DHLAB / EPFL, AMstudio, Venice

- aim: capture 1,000 years of records of the Republic of Venice in dynamic digital form
- machine-learning technologies will enable the researchers to search and cross-reference the information
- for automated text recognition it cooperates with H2020 READ-Transcribus
- challenge: organize transparent processes linking historical records and computer-based visualizations
- first outcomes: development of the Rialto from AD 950
- many possibilities for e.g. language research, personal networks, economic history, medical history etc.
- Vision: link to similar time machines in Europe's historic centres of culture and commerce (e.g. Amsterdam, Paris etc.)

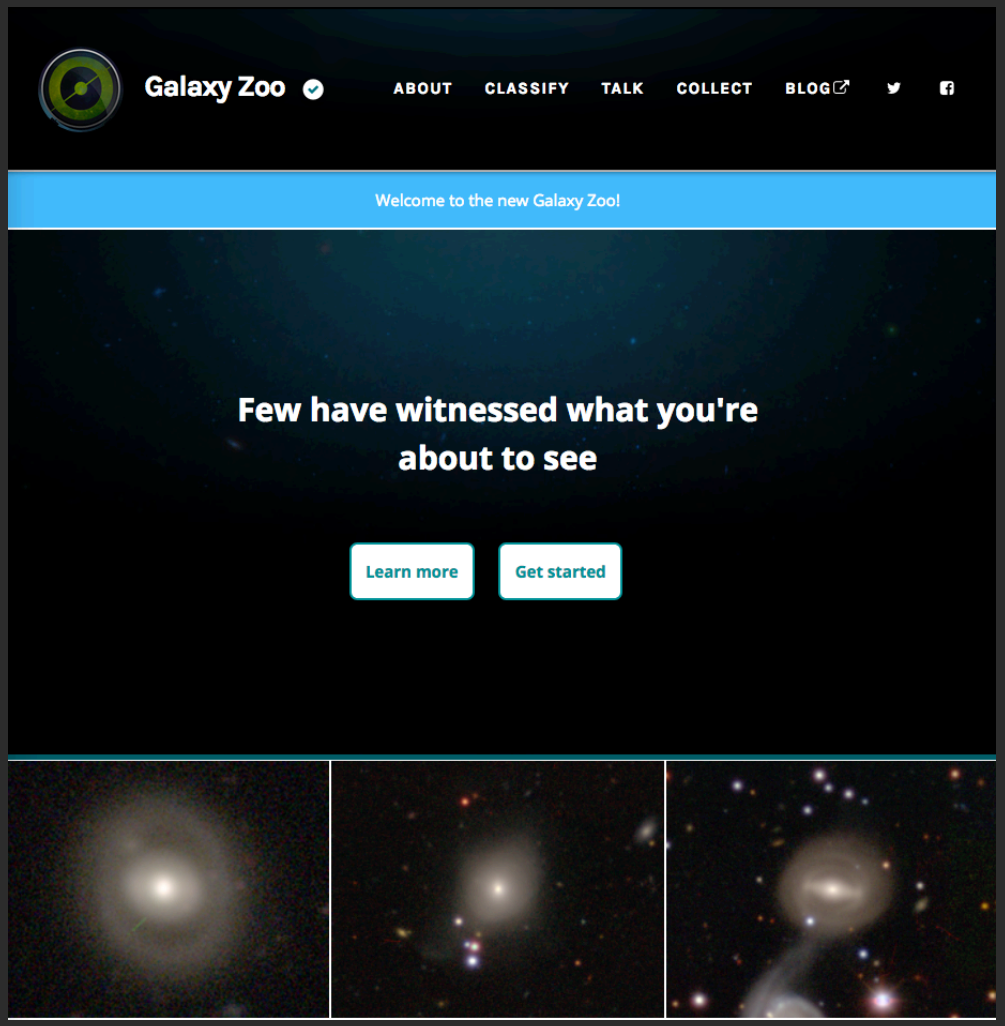
Sources:
https://www.nature.com/news/the-time-machine-reconstructing-ancient-venice-s-social-networks-1.22147
https://vtm.epfl.ch/page-109337.html

Citizen Science

Bürgerwissenschaft

Sciences Participatives

Source: https://de.wikipedia.org/wiki/Wikipedia:Wien/Archiv/2015#/media/File:20140415_TFM_Edit-a-thon_Wien_0301.jpg CC BY-SA 3.0 at



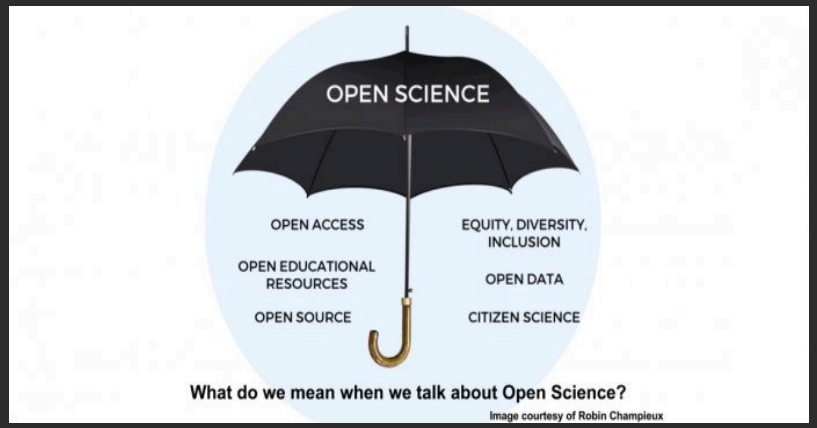
Galaxy Zoo: The most successful Citizen Science project ever (so far…)

Screenshot of: https://www.zooniverse.org/projects/zookeeper/galaxy-zoo/

What is Citizen Science?

"Citizen Science refers to the **general public engagement in scientific research activities** when **citizens actively contribute** to science either with their intellectual effort or surrounding knowledge or with their tools and resources."

Source: Serrano Sanz et al. 2014, p. 8.



Source: https://twitter.com/OHBM/status/971793038543671296

Open Innovation Strategy for Austria 2016

Citizen Science and the Digital Transformation

- “Digital Transformation” (Web 2.0, new technologies, new philosophy of sharing) provides huge potential for Citizen Science
- many successful Citizen Science projects take place in the digital world (VREs)
- digital solutions mobilize individual and collective engagement for co-creation
- new forms of scientific communication play a huge role

Models of Citizen Engagement in Science

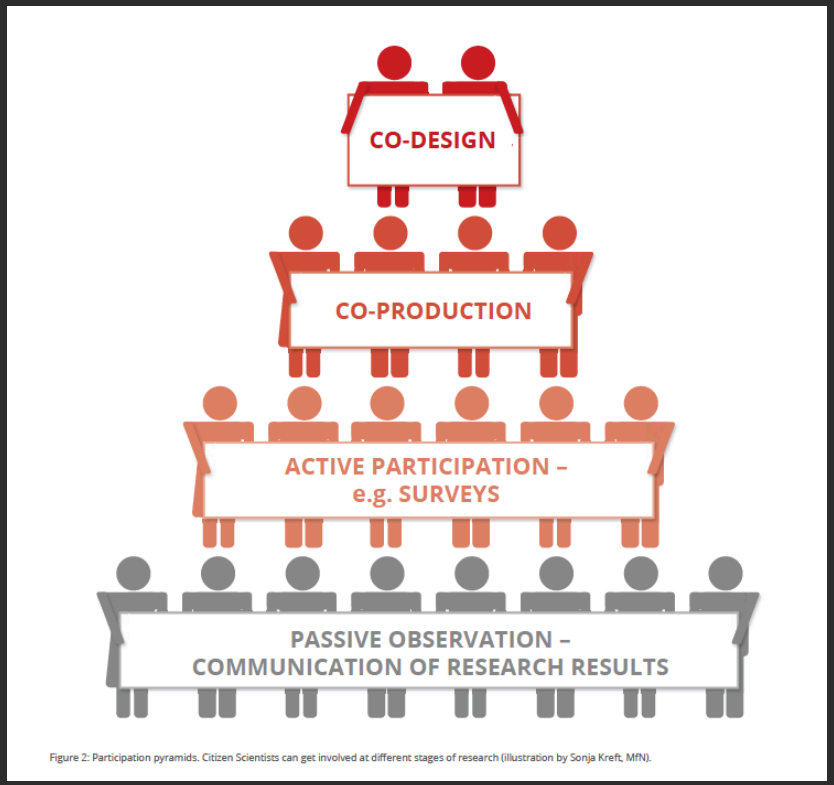
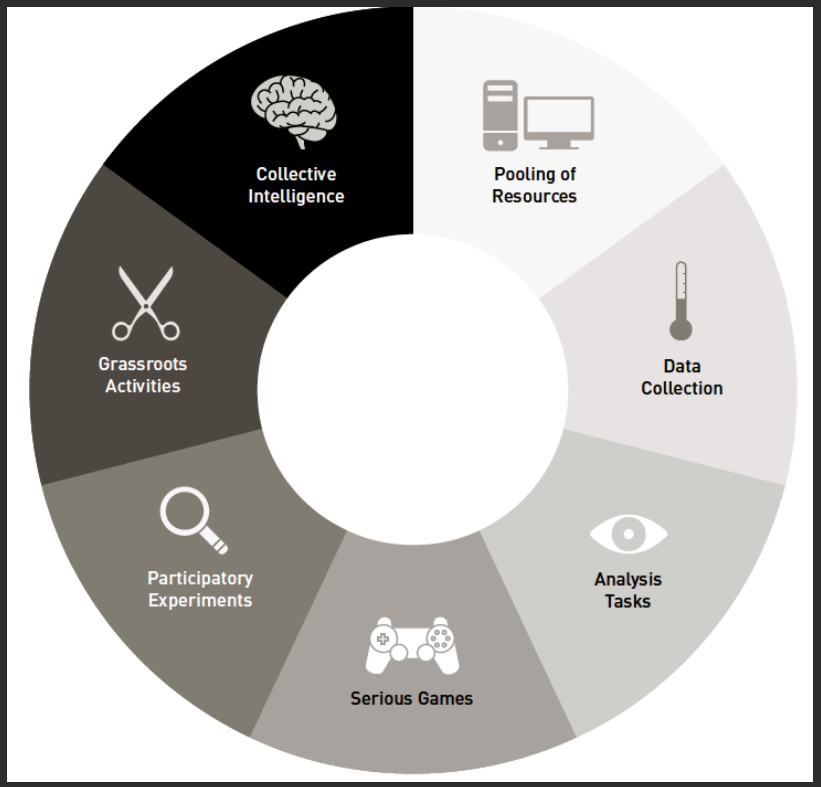


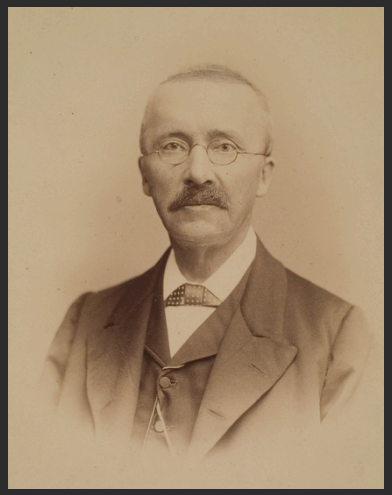
Figure 2: Participation pyramids. Citizen Scientists can get involved at different stages of research (illustration by Sonja Kreft, MfN).

Source: Bonn et al. 2016, p. 17.



Source: Serrano Sanz et al. 2014, p. 9.

Citizen Science examples from eHumanities and eHeritage



Heinrich Schliemann,
the German pioneer of archeology who discovered Troy,
was a businessman, he never studied archeology

Pictures:
https://en.wikipedia.org/wiki/Heinrich_Schliemann#/media/File:Heinrich_Schliemann_(HeidICON_28763)_(cropped).jpg CC BY-SA 4.0,
https://en.wikipedia.org/wiki/Heinrich_Schliemann#/media/File:MaskeAgamemnon.JPG CC BY-SA 3.0

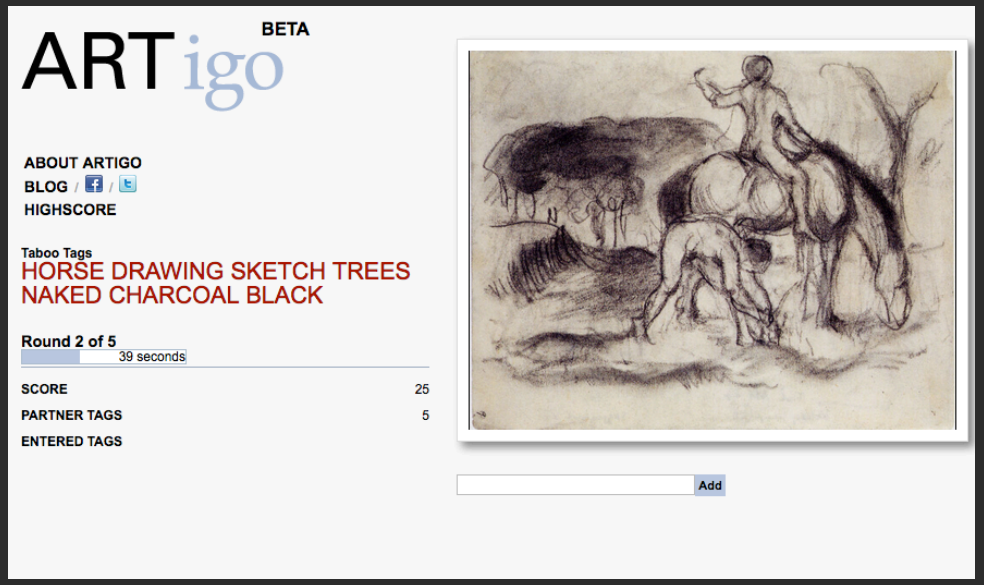
Letters of 1916

- aim: create a crowd-sourced digital collection of letters written around the Easter Rising (Ireland)
- participants transcribe (edit and tag) letters and also add letters to the collection
- huge public engagement factor, educational aspects
- **data collection, analysis**

Screenshot of: http://letters1916.maynoothuniversity.ie/

ARTigo: Informatics and (digital) art history

- aim: develop a semantic search engine for digitised works of art
- several games to promote art recognition: citizen scientists describe the content of artwork with keywords > collect a wide range of keywords in order to enhance search functionality and answer complex research questions
- **gamification** (serious games)



Screenshot of: http://www.artigo.org/tabooTaggingGame.html;jsessionid=AC62C9A64E97191E89CA1F07AAF56B28.www2

Citizen Science @ Libraries

- as part of regional studies e.g. contribution to regional bibliographies, adding to Wikipedia articles, transcription of digitised materials on Wikisource
- get general access to tools and general research support
- georeferencing of maps (Virtuelles Kartenforum 2.0), SLUB Dresden
- contributions to collective research of history (German Example: History of Cycling)

Source examples: https://www.slub-dresden.de/en/open-science/citizen-science/

Potentials and Challenges of Citizen Science

Potentials

- opening of the scientific process and greater research diversity: create scientific knowledge outside institutionalized science with “unusual suspects”
- develop new research questions and topics: societal relevance (real problems, concrete answers) and broader public engagement/understanding
- wider dissemination of research results to the public (science communication)
- Create large data sets & solve work intensive tasks

Challenges

- lack of resources (time, money, spaces, infrastructure)
- lack of fair sharing and compensation models for crowdwork
- data protection / Intellectual Property Right (IPR)

Always keep in mind the two perspectives:

- **Researchers** ask themselves: What do I gain by working with volunteers?
- **Volunteers** ask themselves: What do I gain by working with scientists?

Citizen Science and Research Infrastructures

https://twitter.com/OPERASEU/status/976769535922630662

Poll: What can Research Infrastructures do for Citizen Science?

Take Home Messages

- check if data and tools that you could use are already available
- start with a pilot to test your approach and then go big
- think and develop your research with Open Science in mind from the start (Open Data, Open Access, documentation)
- Citizen Science projects have a high potential to prove the value of the humanities to the society: Get creative and don't be afraid
- For a Citizen Science project you need to think everything through from the beginning: Citizen Science is not an afterthought
- acknowledge need for upskilling of researchers and citizen scientists (new scientific methods, new topics, e.g. legal aspects, data quality)
- digital citizen science projects need to ensure that applications are suited to the project and that data collection and data use conform to data protection laws (but do not reinvent the wheel)

Collaborate! Get technical expertise/support from experts/RIs!

Bibliography & Reading suggestions #citizenscience:

- Bonn et al. *Green Paper Citizen Science Strategy 2020 for Germany*. Berlin, 2016.
- Ten Principles of Citizen Science from the European Citizen Science Association, 2015
- Open Innovation, Open Science, Open to the World - A vision for Europe, European Commission Directorate-General for Research & Innovation, 2016
- Open Innovation Strategy for Austria, BMWFW, Vienna, 2016
- Pettibone and Vohland, *Citizen Science for all: A Guide for Citizen Science Practitioners*, 2016
- Schreibman 2018, Lecture “From Archives to Activism: Why Public Engagement is Essential for the Future of the Humanities” (Slides)
- Heigl, Florian, Dörler, Daniel, Bartar, Pamela, Brodschneider, Robert, Cieslinski, Marika, Ernst, Marlene, … Ziegler, David. (2018, February 1). Qualitätskriterien für Citizen Science Projekte auf Österreich forscht (Version 1.0). Zenodo. http://doi.org/10.5281/zenodo.1161922
- Bowser and Shanley, *New Visions in Citizen Science*, 2013, chapter Citizen Archivists.
- Serrano Sanz et al. *White Paper on Citizen Science for Europe*. Socientize Consortium, 2014

Some Citizen Science Networks & Associations

- www.buergerschaffenwissen.de (Germany)
- www.citizen-science.at (Austria)
- www.science-et-cite.ch/de/citizen-science (Switzerland)
- www.citizenscienceassociation.org (U.S.)
- https://ecsa.citizen-science.net (Europe)
- https://www.citizenscience.org.au/who-we-are/ (Australia)
- https://www.facebook.com/CitSciAsia (Asia)
- http://citizenscience.org/ (global)
- http://citsci.org (Colorado State University)

Questions & Answers

(Picture CC0 https://pixabay.com/photo-2999583/)

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Announcements

- **New PARTHENOS Training Suite module** around **Developing research Questions** is being developed by PARTHENOS (led by Darja)
- Have a look at the PARTHENOS YouTube Channel: all videos so far produced by PARTHENOS and PARTHENOS Training
- **PARTHENOS Webinar Workshop** at the European Summer University in Digital Humanities 2018 (ESU), Leipzig (apply until 1st of May 2018, scholarships available):
  How Research Infrastructures empower eHumanities and eHeritage Research(ers)

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Questions & Answers

(Picture CC0 https://pixabay.com/photo-2999583/)

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

Thank you for joining us!

(Picture: CCO https://pixabay.com/photo-1889007/)

- ✓ Please help us improving the PARTHENOS Webinar Series and fill out the short Feedback Survey that we'll sent you in a follow-up e-mail.

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.

PARTHENOS
Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies

FH;P Fachhochschule Potsdam University of Applied Sciences

PARTHENOS eHumanities and eHeritage Webinar Series

Join the conversation on Twitter:
#PARTHENOSWebinar
@Parthenos_EU

Unless otherwise stated this work is licensed under a Creative Commons Attribution 4.0 International License.
Graphic: Otto, the Open Access Otter CC-BY 4.0, Katja Diederichs