A thematic marketplace:
The SSH Open Marketplace

Franciska de Jong

executive director CLARIN ERIC

f.m.g.dejong@uu.nl

3rd ESFRI RIs -EOSC Workshop, 26 January 2022

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

- SSH = Social Sciences and Humanities, *aka*
  SCI = Social and Cultural Innovation (ESFRI-specific jargon)

- How can the thematic services of SSHopencloud.eu Social Sciences & Humanities Resources add value to the Research Infrastructure community?

- SSH Open Marketplace Social Sciences and Humanities Open Marketplace - a thematic service discovery platform rooted in disciplinary communities, compatible with existing scholarly commons and aligned with international standards

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC social sciences & humanities open cloud

The role of thematic marketplaces in the EOSC ecosystem

- Stimulating user engagement *(Zoe Cournia)*
- Community-driven guidance for generating metadata *(Anne Fouilloux)*
- Contextualisation through rich curation of data, workflows, scenarios
- Recommendation functionality as part of discovery workflows
- Fit for integration in RI reward framework, leveraged visibility of resources
- Comfort zone for individual researchers: link with peers, disciplinary traditions
- Starting point for cross-domain or cross-cluster research projects
- EOSC Platform not much appreciated (yet) (attractive for 20%)
- Provides room for building the ground work first without losing the link with contexts of use.

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

SSHOC: a cluster project contributing to the maturing of EOSC and the visibility of thematic services

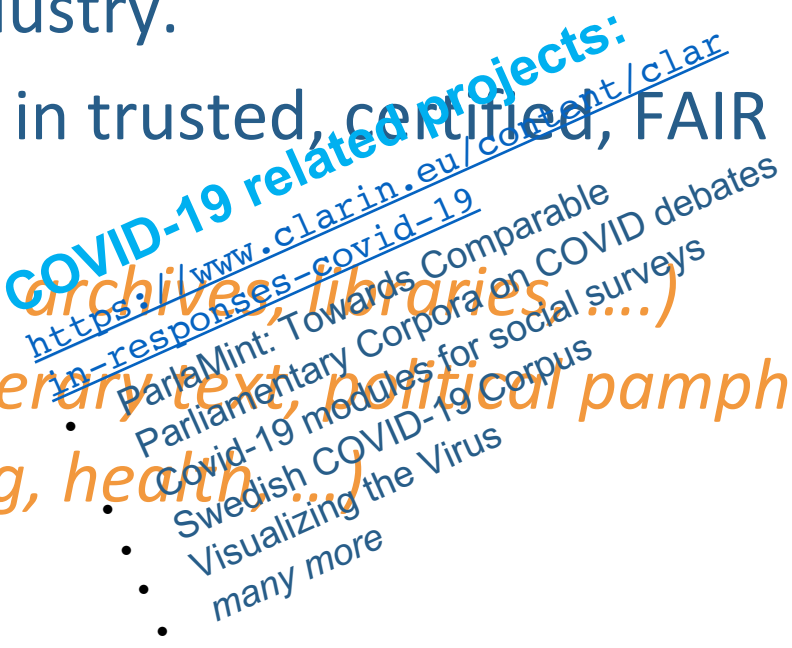
- Working towards common data and thematic service protocols that will serve the interoperability between the disciplines involved, and thereby strengthen the support for 500.000 million SSH researchers (academic and non-academic) and the data science industry.
- Focussing on cluster-specific data types in trusted, certified, FAIR repositories (>100). Examples:
  - Cultural heritage content *(museums, archives, libraries, ….)*
  - Language resources *(newspapers, literary text, political pamphlets, …)*
  - Social survey data *(income, wellbeing, health, …)*
  - …
- Providing room for generalising thematic services more effectively than via direct focus on EOSC as a discovery platform.

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

SSHOC: a cluster project contributing to the maturing of EOSC and the visibility of thematic services

- Working towards common data and thematic service protocols that will serve the interoperability between the disciplines involved, and thereby strengthen the support for 500.000 million SSH researchers (academic and non-academic) and the data science industry.
- Focussing on cluster-specific data types in trusted, certified, FAIR repositories (>100). Examples:
  - Cultural heritage content *(museums, archives, libraries, …)*
  - Language resources *(newspapers, literary text, political pamphlets, …)*
  - Social survey data *(income, wellbeing, health, …)*
  - …
- Providing room for generalising thematic services more effectively than via direct focus on EOSC as a discovery platform.

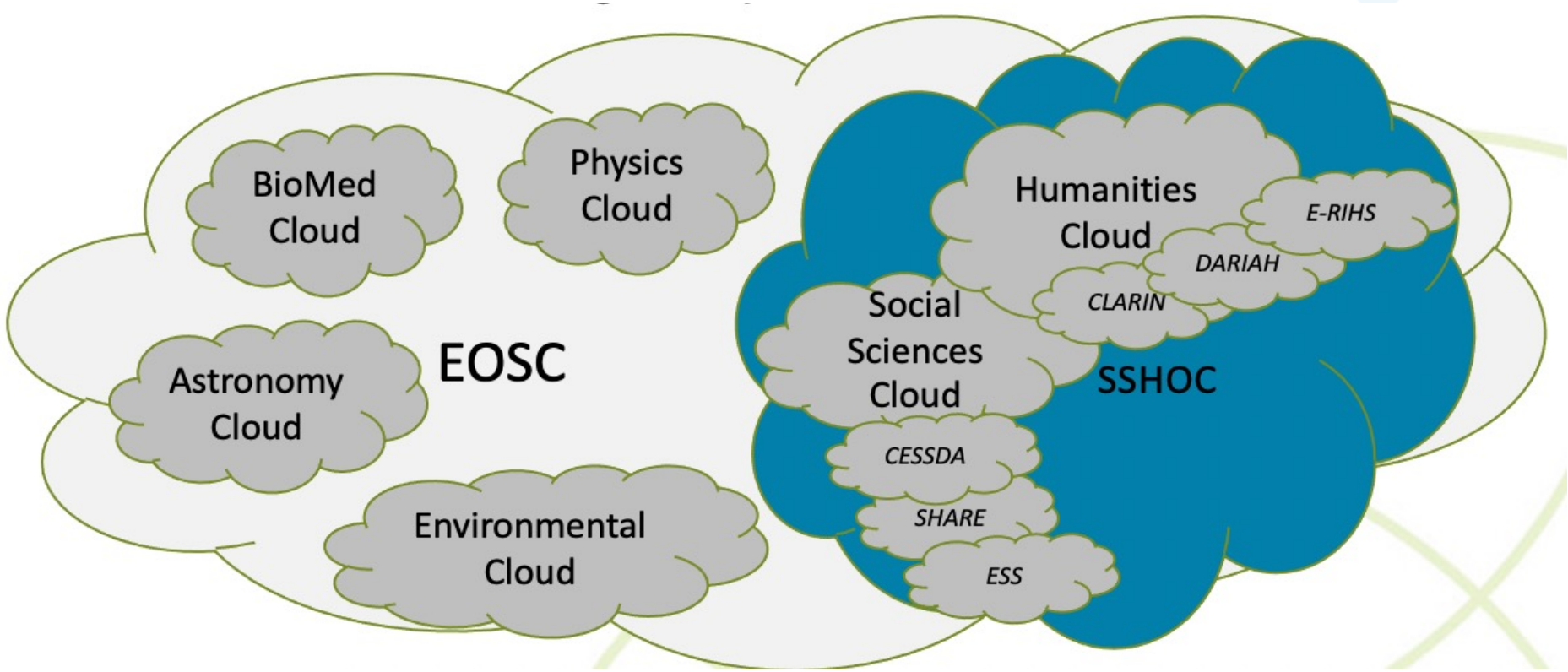


This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

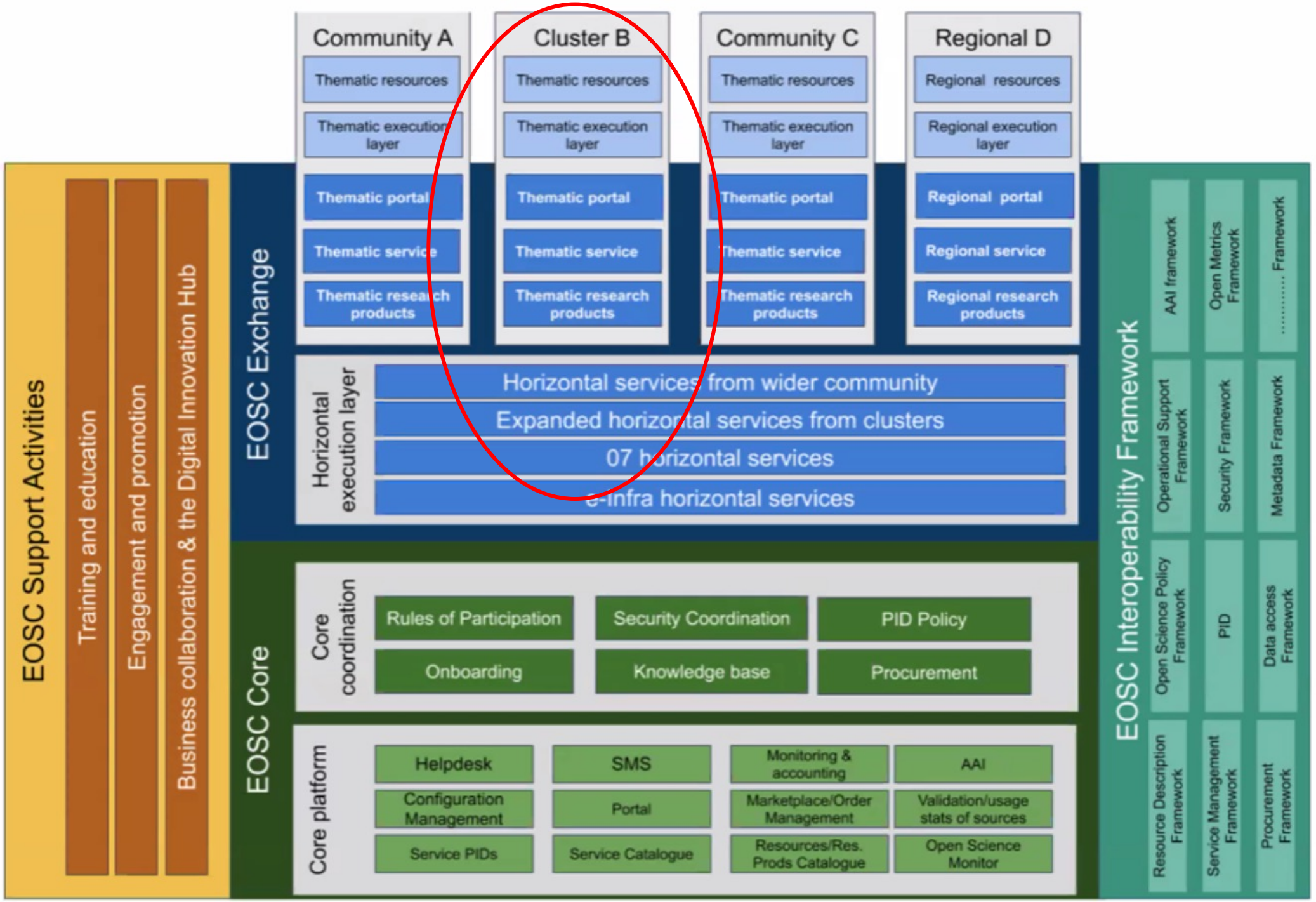
Support for EOSC as Cloud of Clouds, System of Systems, Marketplace of Marketplaces

a distributed infrastructure, with shareable resources, optimized access for researchers and cost-efficient service offer based on federation and collaboration.



This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud


Community A
Thematic resources
Thematic execution layer
Thematic portal
Thematic service
Thematic research products
Cluster B
Thematic resources
Thematic execution layer
Thematic portal
Thematic service
Thematic research products
Community C
Thematic resources
Thematic execution layer
Thematic portal
Thematic service
Thematic research products
Regional D
Regional resources
Regional execution layer
Regional portal
Regional service
Regional research products
EOSC Exchange
Horizontal execution layer
Horizontal services from wider community
Expanded horizontal services from clusters
07 horizontal services
e-infra horizontal services
EOSC Core
Core coordination
Rules of Participation
Security Coordination
PID Policy
Onboarding
Knowledge base
Procurement
Core platform
Helpdesk
SMS
Monitoring & accounting
AAI
Configuration Management
Portal
Marketplace/Order Management
Validation/usage stats of sources
Service PIDs
Service Catalogue
Resources/Res. Prods Catalogue
Open Science Monitor
EOSC Support Activities
Training and education
Engagement and promotion
Business collaboration & the Digital Innovation Hub
EOSC Interoperability Framework
AAI framework
Operational Support Framework
Open Metrics Framework
Security Framework
Metadata Framework
Open Science Policy Framework
PID
Data access Framework
Resource Description Framework
Service Management Framework
Procurement Framework

Critical first steps towards success for EOSC 1/2

The main focus and value of the MVE is to connect disciplinary infrastructures and research data to enhance disciplinary, cross-disciplinary and transnational research, leading to new scientific discoveries and new insights for society. Hence, **the federation of research data infrastructures *via* the thematic clustering of Research Infrastructures, e-Infrastructures and regional projects is seen as a critical first step to promote early successes for EOSC,** and as an enabler to help address society's current and future global challenges.

Solutions for a Sustainable EOSC; A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group (2020)

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

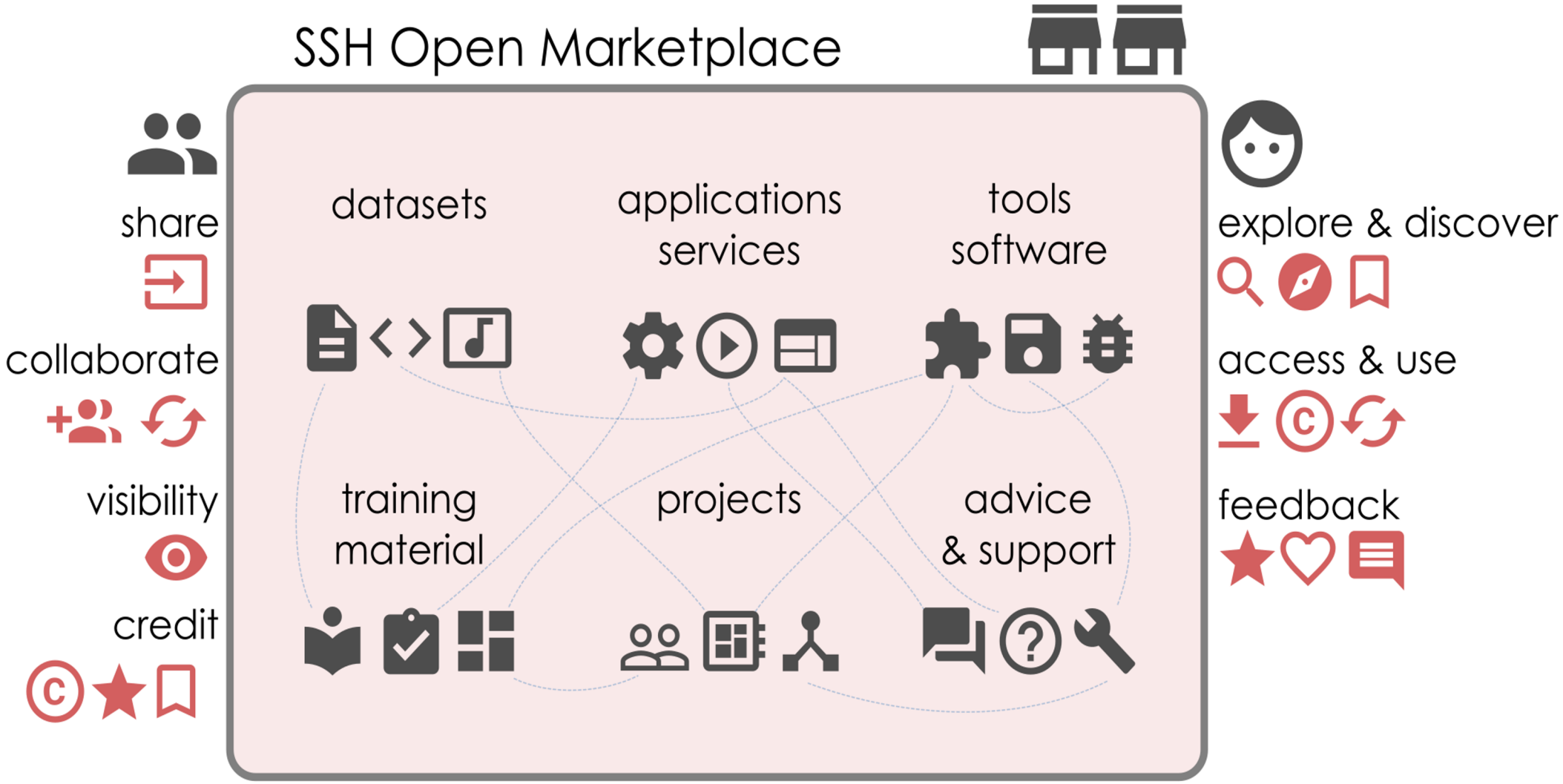
Critical first steps towards success for EOSC 2/2

… resources and **thematic services** identified as ready for integration by the Research Infrastructure clusters should be prioritised for on-boarding because they **represent high-quality resources valued by research communities** and the **Research Infrastructures** have demonstrated their eagerness to contribute to EOSC and **have already established communication channels**, and **at cluster-level best practices are being developed to ensure the interoperability of federated services and data …**

Solutions for a Sustainable EOSC; A FAIR Lady (olim Iron Lady) report from the EOSC Sustainability Working Group (2020)

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud



Slide by DARIAH-EU

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

SSH Open Marketplace

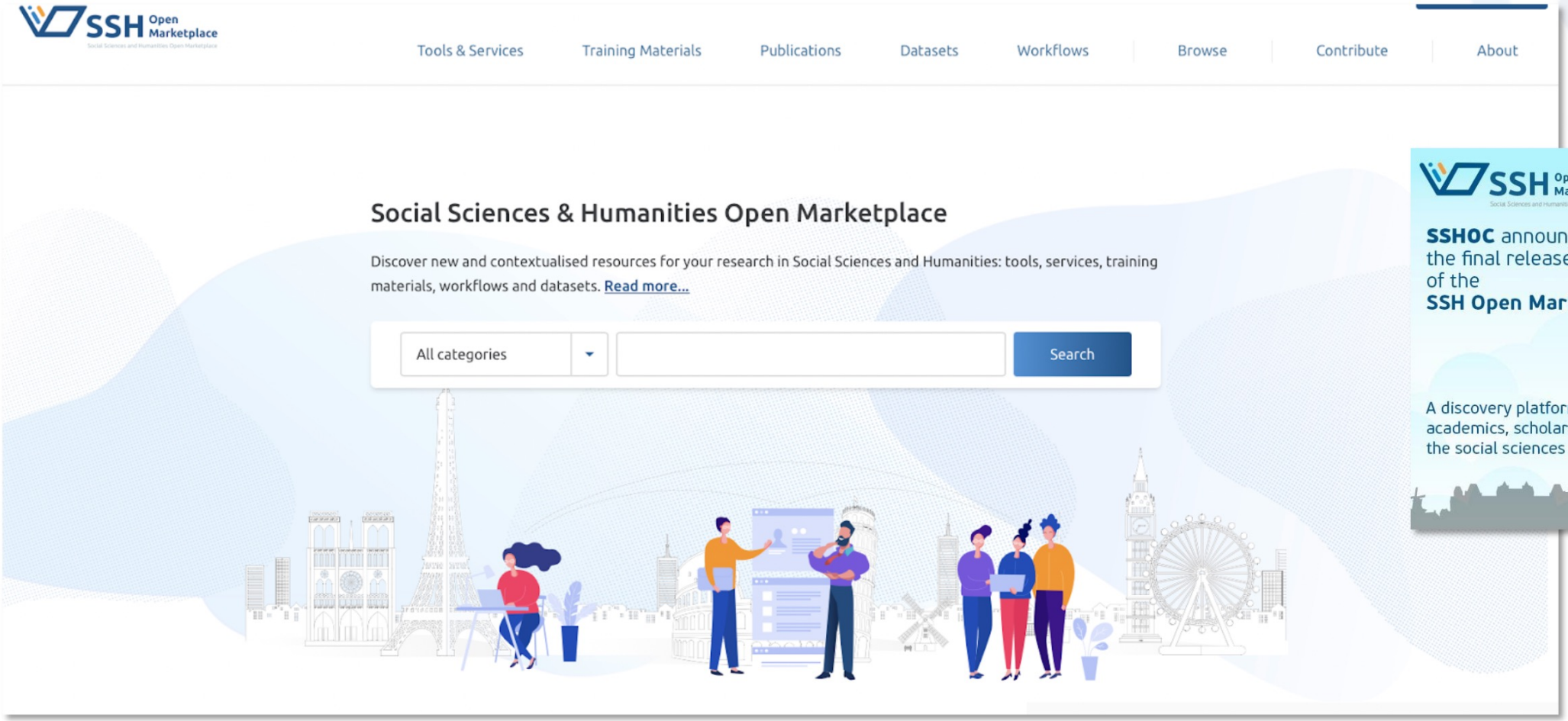
Social Sciences and Humanities Open Marketplace

SSHOC announces the final release of the **SSH Open Marketplace:**

A discovery platform of resources for academics, scholars, and students of the social sciences and humanities.

This

The SSH Open Marketplace



Visit us here: marketplace.sshopencloud.eu

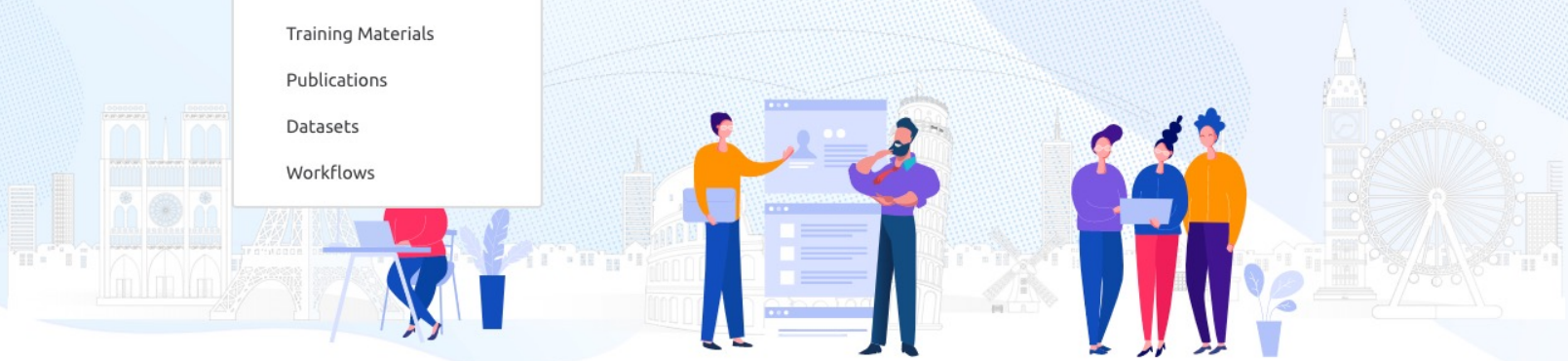
This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

Social Sciences & Humanities Open Marketplace

Discover new and contextualised resources for your research in Social Sciences and Humanities: tools, services, training materials, workflows and datasets. Read more...

All categories | Search

- All categories
- Tools & Services
- Training Materials
- Publications
- Datasets
- Workflows



Browse

Browse by activity

See all

Analyzing (641) Visual Analysis (313) Content Analysis (250) Discovering (180) Capturing (168) Annotating (142) Enriching (138) Disseminating (127) Gathering (123) Creating (116) Publishing (101) Collaborating (97) Organizing (94) Sharing (91) Network Analysis (79) Storing (76) Editing (72) Data Cleansing (61) Web Development (54) Communicating (50)

Browse by keyword

See all

Analysis (558) Other (358) Visualization (335) Content Analysis (240) search (205) american (197) language (177) natural-language-processing (166) 2000s (161) 2010s (154) Capture (148) Discovering (148) historic (135) Enrichment (119) Gathering (117) Annotating (114) Collaboration (107) Dissemination (107) Publishing (107) Organizing (92)

Last added

See what's new

See all

Blender Tutorials

The Organization The Blender Foundation (2002) is an independent public benefit organization. Its spin-off corporation Blender Institute (2007) hosts the foundation's offices and currently employs 24 people who work on the Blender software and creative projects to validate and...

Read more

This p

Recommended

Tools & Services

See all

Gephi

Activities: Capturing, Modeling, Network Analysis, Spatial Analysis, Analyzing
Keywords: Models, Visualization, 2000s, visualization, graphs

Gephi is the leading visualization and exploration software for all kinds of graphs and networks. Gephi is open-source and free. Runs on Windows, Mac OS X and Linux.

Read more

UDPipe

Activities: Analyzing, Natural Language Processing, Annotating, Structural Analysis, Parsing
Keywords: recommended

UDPipe is a trainable pipeline for tokenization, tagging, lemmatization and dependency parsing of CoNLL-U files. UDPipe is language-agnostic and can be trained given annotated data in CoNLL-U format. Trained models are provided for nearly all UD treebanks. UDPipe is available as...

Read more

Training Materials

See all

Corpus Analysis with Antconc

Activities: Analyzing, Content Analysis, Interpreting, Contextualizing
Keywords: distant-reading, recommended

Corpus analysis is a form of text analysis which allows you to make comparisons between textual objects at a large scale (so-called 'distant reading'). You will learn to do "Keyword-in-context" searching, identify patterns surrounding a particular word, look at statistically...

Read more

Exploring and Analyzing Network Data with Python

Activities: Analyzing, Network Analysis
Keywords: recommended

This lesson introduces network metrics and how to draw conclusions from them when working with humanities data. You will learn how to use the NetworkX Python package to produce and work with these network statistics.

Read more

Publications

See all

Visual Network Analysis with Gephi Workshop

Activities: Network Analysis
Keywords: recommended

"On the record" - transcribing and valorizing qualitative interviews with XML-TEI

Activities: Transcribing, Encoding, Annotating
Keywords: ODD, qualitative data, recommended

This proj

application paired with the Qualtrics survey platform to meet the specific needs of cross-national web surveys, for the social sciences and beyond. WPSS enables: a centralised management of survey fielding orchestration...

Read more

TaDiRAH Taxonomy of Digital Research Activities in the Humanities

Activities: Description
Keywords: DARIAH National Resource

The Taxonomy of Digital Research Activities of the Humanities is designed to organize and categorize content, e.g., CLARIAH-DE, DARIAH Zotero bibliography "Doing Digital Humanities", DHCommons, TAPoR (Text Analysis Portal for Research), AGATE (European Science Academies Gateway...

Read more

The R Manuals

On the website, manuals can be downloaded as PDF files, EPUB files, or browsed as HTML An Introduction to R is based on the former "Notes on R", gives an introduction to the language and how to use R for doing statistical analysis and graphics. R Data Import/Export describes the...

Read more

AntConc

Activities: Discovering, Analyzing, Structural Analysis, Concordance
Keywords: american, search, 2000s, collocation, word list

AntConc is free concordance software. It is multi-platform and easy to deploy and use. AntConc is part of a suite of related tools for text processing and analysis, including applications for parallel corpus analysis, word profiling, PDF to text conversion, text structure...

Read more

Publications

See all

Visual Network Analysis with Gephi Workshop

Activities: Network Analysis
Keywords: recommended

Gephi is a free and open source network analysis software used, among other things, in social network analysis. This workshop is intended for beginners as well as confirmed users. First, we offer an introduction to the basics of Gephi, then we explore through practice the...

Read more

"On the record" - transcribing and valorizing qualitative interviews with XML-TEI

Activities: Transcribing, Encoding, Annotating
Keywords: ODD, qualitative data, recommended

Qualitative interviews constitute an important research tool for disciplines such as history, sociology, ethnology or political science. Yet, despite rare initiatives (Cadorel et al., 2018), transcriptions are scarcely shared with other researchers. And their annotation is most...

Read more

Datasets

See all

ParlaMint

Activities: Enriching, Editing, Formatting
Keywords: parliament, covid-19, recommended, DARIAH National Resource

This dataset contains National parliamentary data, a verified communication channel between the elected political representatives and society members in any democracy. The goal is to provide resources and tools for focused observations on trends, opinions, decisions on...

Read more

Moviegalaxies – Social Networks in Movies

Activities: Network Analysis
Keywords: GEXF, movies, recommended

This repository contains network graphs and network metadata from Moviegalaxies, a website providing network graph data from about 773 films (1915–2012). The data includes individual network graph data in Graph Exchange XML Format and descriptive statistics on measures such as...

Read more

Workflows

See all

Linguistic annotation of corpora

Activities: Encoding, Cluster Analysis, Collocation Analysis, Concordance, Named Entity Recognition
Keywords: text, recommended

Create a dictionary in TEI

Activities: Encoding
Keywords: Linguistics, recommended

This scenario sets out the best practices for creating a born-digital

AntConc is free concordance software. It is multi-platform and easy to deploy and use. AntConc is part of a suite of related tools for text processing and analysis, including applications for parallel corpus analysis, word profiling, PDF to text conversion, text structure...

Read more

This pr

Publications

See all

Visual Network Analysis with Gephi Workshop

Activities: Network Analysis
Keywords: recommended

Gephi is a free and open source network analysis software used, among other things, in social network analysis. This workshop is intended for beginners as well as confirmed users. First, we offer an introduction to the basics of Gephi, then we explore through practice the...

Read more

"On the record" - transcribing and valorizing qualitative interviews with XML-TEI

Activities: Transcribing, Encoding, Annotating
Keywords: ODD, qualitative data, recommended

Qualitative interviews constitute an important research tool for disciplines such as history, sociology, ethnology or political science. Yet, despite rare initiatives (Cadorel et al., 2018), transcriptions are scarcely shared with other researchers. And their annotation is most...

Read more

Datasets

See all

ParlaMint

Activities: Enriching, Editing, Formatting
Keywords: parliament, covid-19, recommended, DARIAH National Resource

This dataset contains National parliamentary data, a verified communication channel between the elected political representatives and society members in any democracy. The goal is to provide resources and tools for focused observations on trends, opinions, decisions on...

Read more

Moviegalaxies – Social Networks in Movies

Activities: Network Analysis
Keywords: GEXF, movies, recommended

This repository contains network graphs and network metadata from Moviegalaxies, a website providing network graph data from about 773 films (1915–2012). The data includes individual network graph data in Graph Exchange XML Format and descriptive statistics on measures such as...

Read more

Workflows

See all

Linguistic annotation of corpora

Activities: Encoding, Cluster Analysis, Collocation Analysis, Concordance, Named Entity Recognition
Keywords: text, recommended

This scenario explains the steps to take to annotate a corpus in order to conduct linguistic and statistical analysis based on it. This scenario wants to

Create a dictionary in TEI

Activities: Encoding
Keywords: Linguistics, recommended

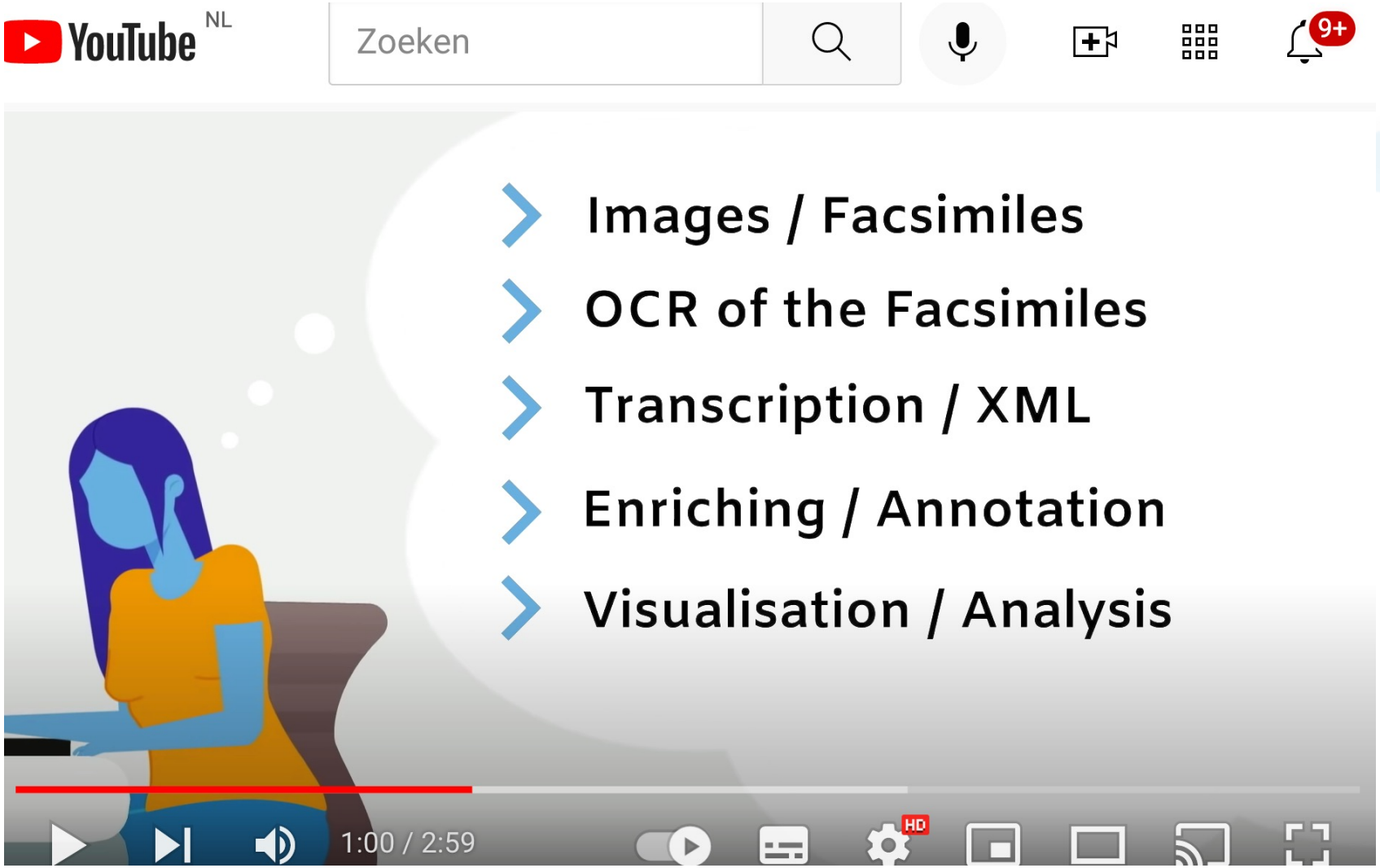
This scenario sets out the best practices for creating a born-digital dictionary, especially with the TEI (Text Encoding Initiative). However,

AntConc is free concordance software. It is multi-platform and easy to deploy and use. AntConc is part of a suite of related tools for text processing and analysis, including applications for parallel corpus analysis, word profiling, PDF to text conversion, text structure...

Read more

This pr

Support for workflow selection



This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782

SSHOC
social sciences & humanities open cloud

The slides were prepared with invaluable input from:

- the SSHOC promotion team
- DARIAH and CLARIN colleagues

***More details*:**

- SSH Open Market video: https://www.youtube.com/watch?v=4tyZ4vhrM9s
- SSH Open Marketplace news: https://www.sshopencloud.eu/news/call-wild-final-release-ssh-open-marketplace

This project is funded from the EU Horizon 2020 Research and Innovation Programme (2014-2020) under Grant Agreement No. 823782