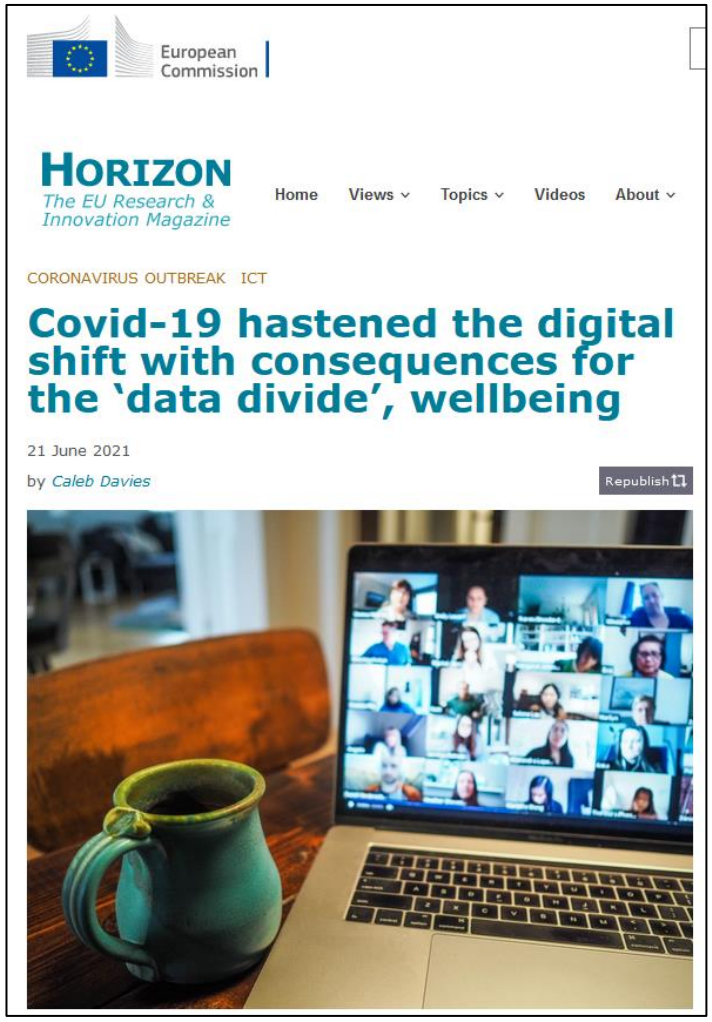
Digital humanities research & infrastructures in a time of pandemic, diversity and media consolidation

Paul Spence
Department of Digital Humanities,
King's College London

2021eko azaroaren 26an
Euskararentzako CLARIAH-EUS hizkuntza-teknologia humanitateetan eta zientzia sozialetan garatzeko azpiegitura diseinatzen jardunaldia



"The Covid-19 pandemic has put rocket boosters under the transition to a digital society."

https://ec.europa.eu/research-and-innovation/en/horizon-magazine/covid-19-hastened-digital-shift-consequences-data-divide-wellbeing

- The move to digital methods has resulted in greater flexibility and creativity
- "The digital divide has become more complex"
- Important issues around the ethics and care of research online

https://blogs.lse.ac.uk/impactofsocialsciences/2020/10/26/how-the-pandemic-has-transformed-research-methods-and-ethics-3-lessons-from-33-rapid-responses/

The changing relationship of humanities to data/digital media

Learning and pedagogy

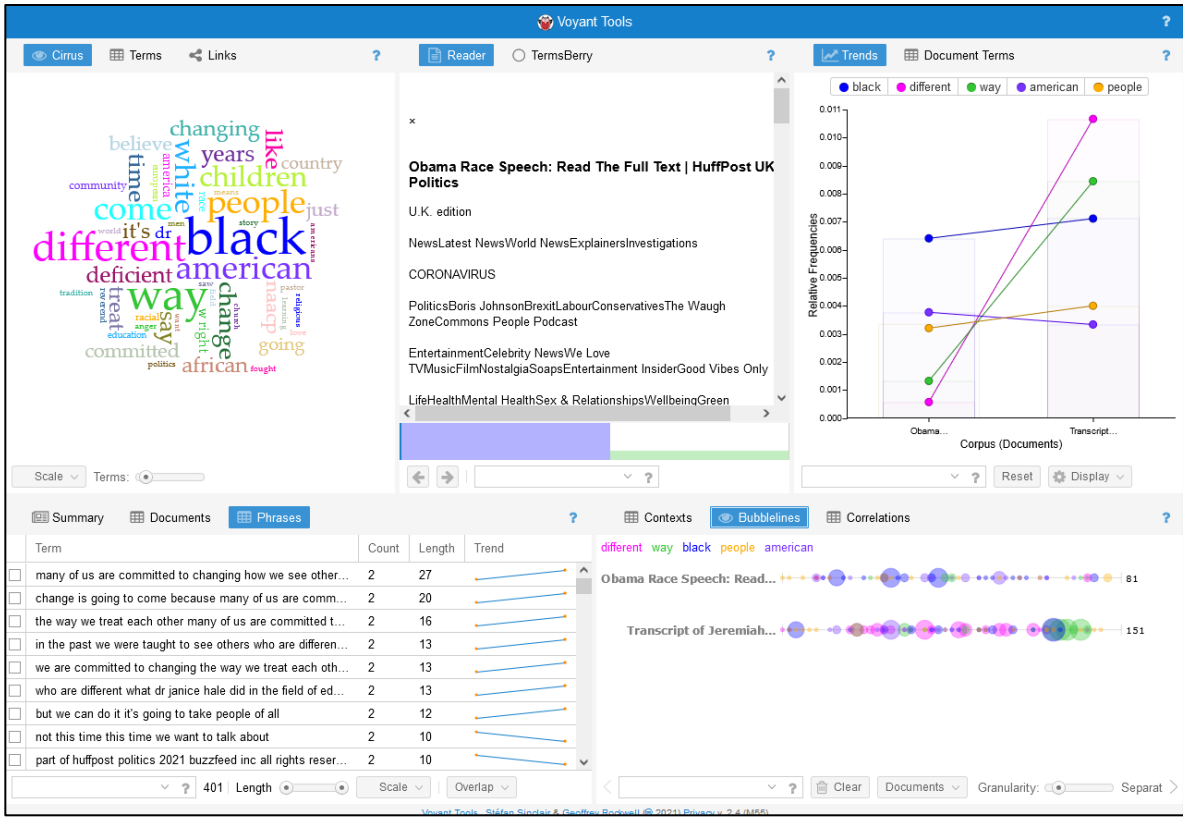
Global knowledge flows

Identity construction

The changing relationship to digital scholarship

Digital Humanities as bridge

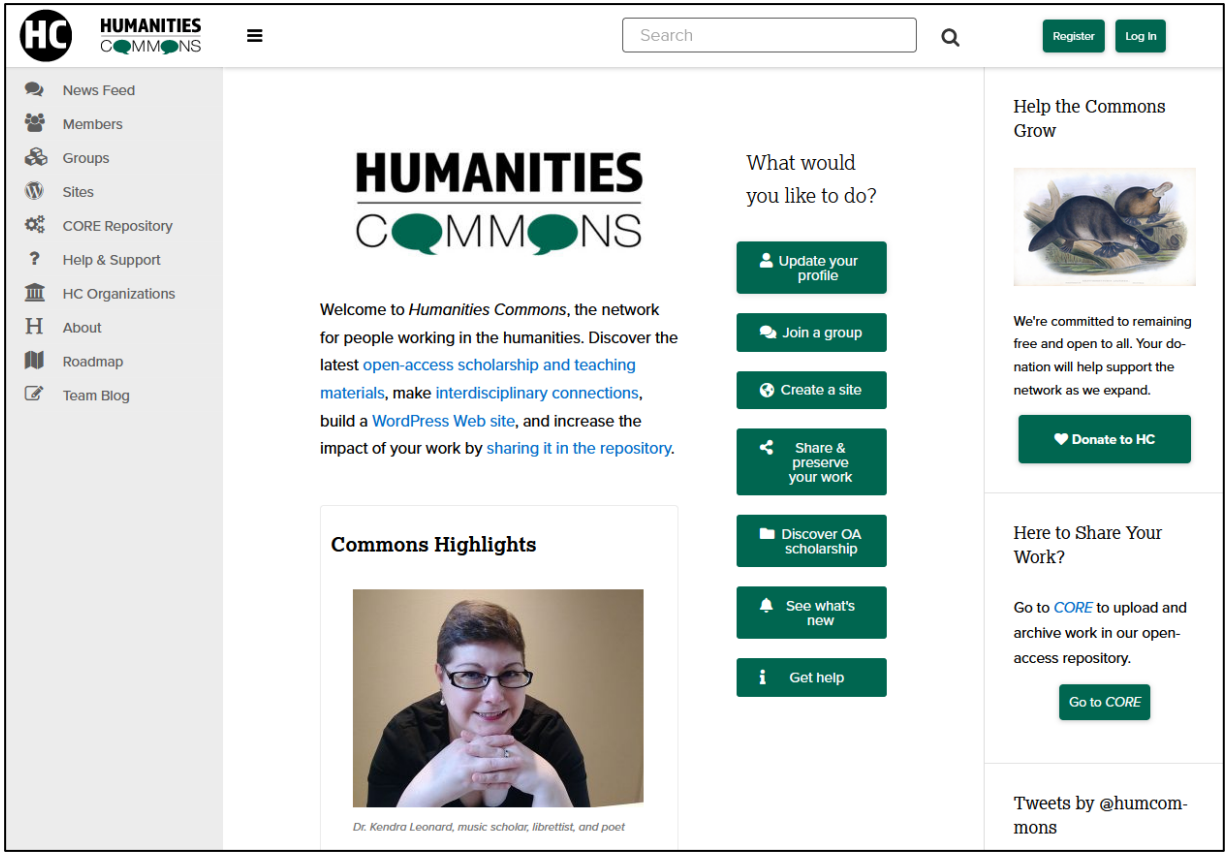
- Advanced application of digital tools to address humanities and social sciences research questions



Hermeneutica: Computer-Assisted Interpretation in the Humanities
Geoffrey Rockwell and Stéfan Sinclair 2016
https://voyant-tools.org

Digital Humanities as bridge

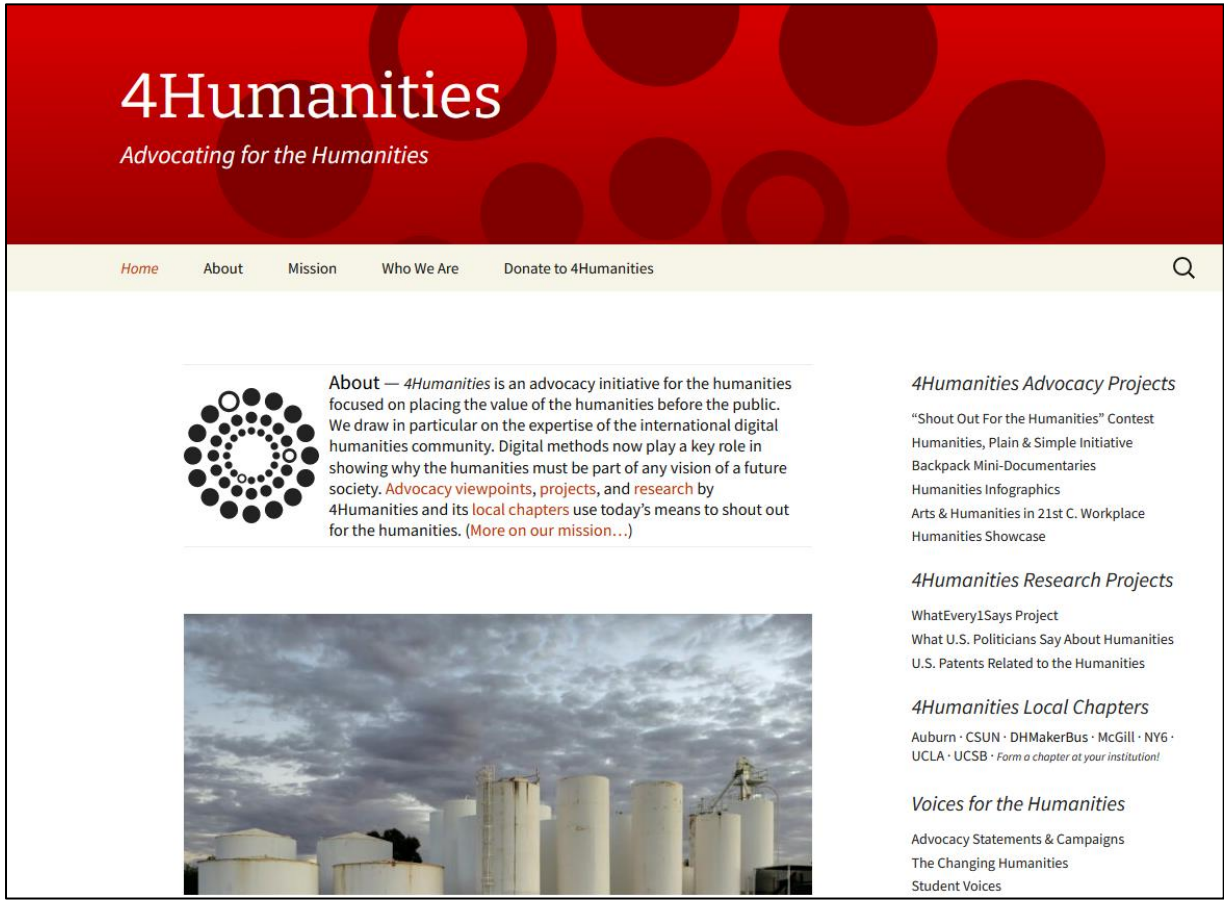
- Community building and discussion around new models for digital scholarship



https://hcommons.org/

Digital Humanities as bridge

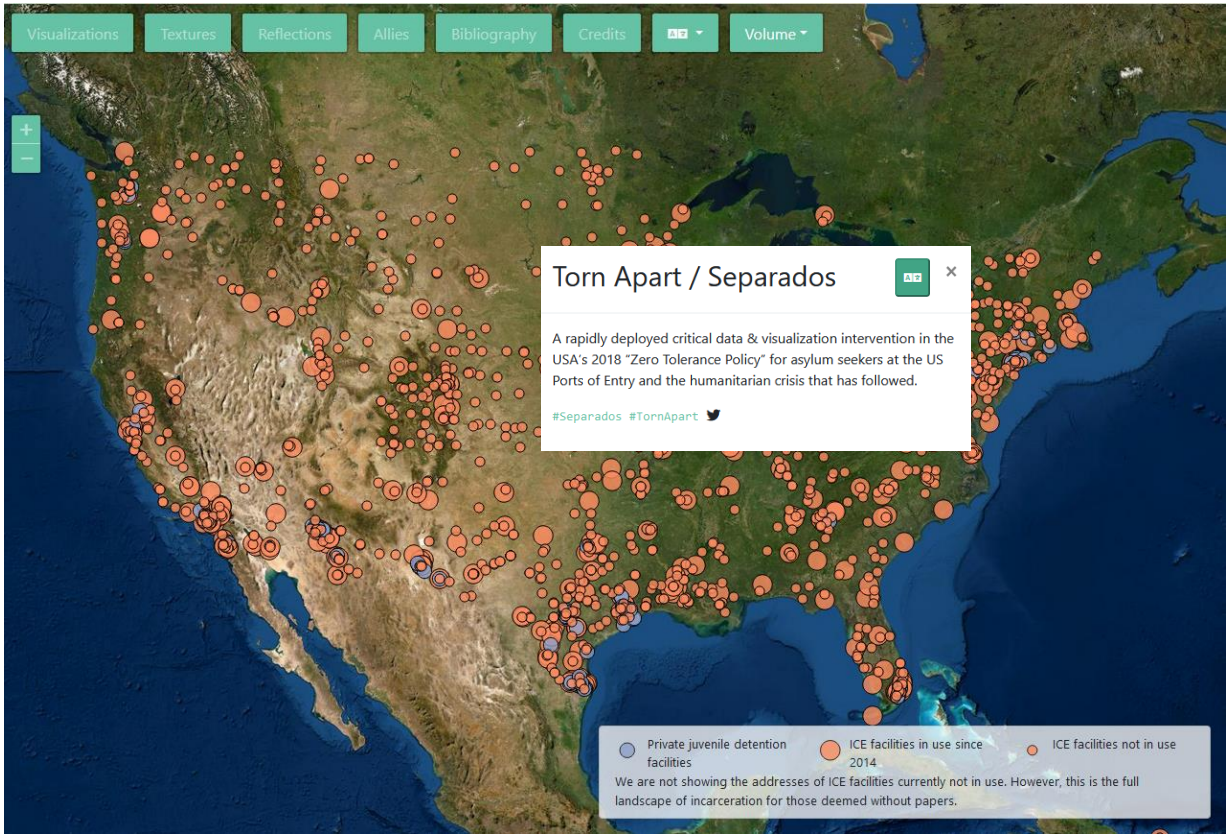
- Digital advocacy



https://4humanities.org/

Digital Humanities as bridge

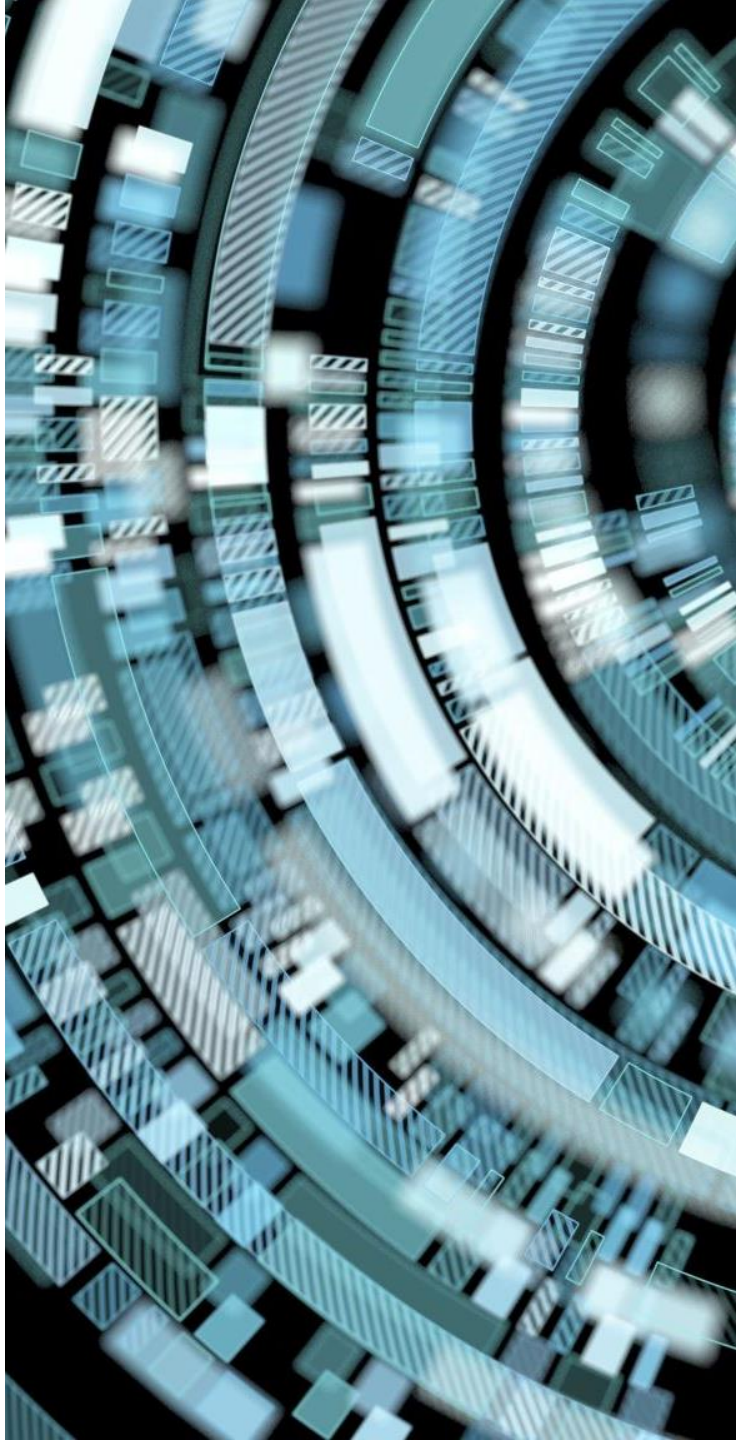
- Digital activism

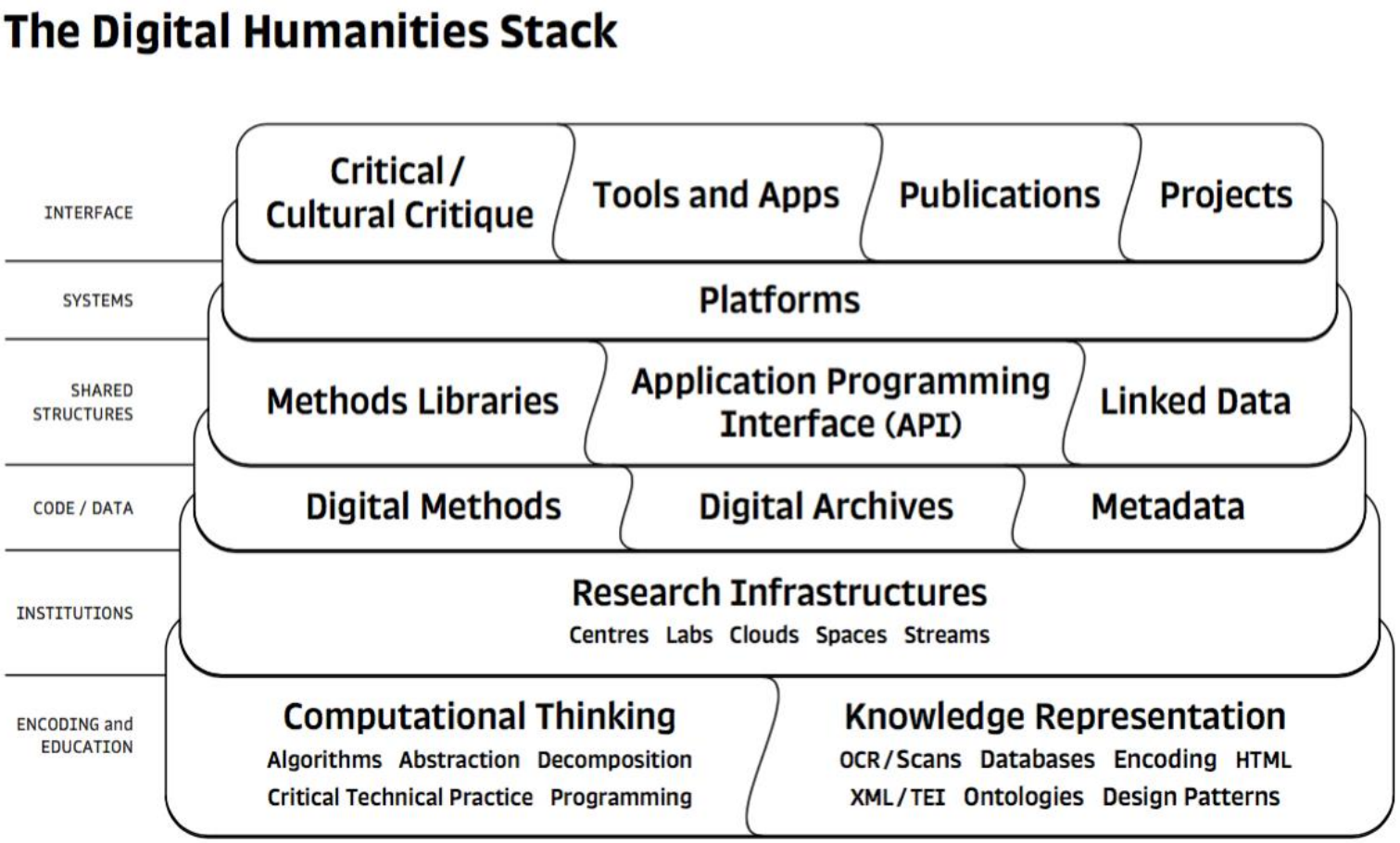


https://xpmethod.columbia.edu/torn-apart/volume/1/

Digital methods and infrastructure for the humanities

- How can humanities researchers can acquire the digital literacies they need?
- How can they learn to identify, learn and apply critically digital methods to their research?
- How can the humanities contribute to the design of the digital systems they need?





Berry, David M.; Fagerjord (2017). Digital Humanities: Knowledge and Critique in a Digital Age.

Research Infrastructures: Communities, not just technologies

"digital ecosystems are an emerging new concept of infrastructures that recognises the need for a flexible combination of humans, machines, content and things to work together on a common task"

"In this definition, digital ecosystems are derived from communities rather than technologies. As open systems, digital research ecosystems will rely on communities and community involvement in a scenario where anyone can participate."

Sheila *Anderson* & Tobias *Blanke 2012*. Taking the Long View: From e-Science Humanities to Humanities. Digital Ecosystems.

Top photos by Bruno Kelzer and Fré Sonneveld on Unsplash
Bottom photo from UK Ireland DH Network at https://dhnetwork.org/

DH RIs: design criteria
[DH = Digital Humanities, RI = Research Infrastructure]

Conceptual parameters

Design Criteria for HUMLab in Sweden:

- Translucence – "see what is going on"
- Flexibility
- Memory and context
- Multiplexity
  - "simultaneous engagement with many types of materials, ideas and perspectives through the screens, interaction technology, sensor technology and audio"
  - "includes a variety of technologies, systems and software"
- Non-linearity and messiness
- Engagement
- Adaptability and co-evolution

Svensson 2011 *From Optical Fiber To Conceptual Cyberinfrastructure* Digital Humanities Quarterly

DH RIs: design criteria

Practical parameters

- Long-term preservation and persistent identifiers
- Sharing of data, tools and services
- Active collection and development of new data and tools
- Applicable Business and sustainability models

European Science Foundation report 2011

http://www.esf.org/fileadmin/user_upload/esf/RI_DigitalHumanities_B42_2011.pdf

How can we protect humanities research and publications from digital commodification?

Image: Pixabay

How can humanities researchers (of any level of digital competence) access and effectively use digital research tools?

CLARIN

"a digital infrastructure offering data, tools and services to support research based on language resources"

CLARIN

https://www.clarin.eu/

DARIAH

"The pan-European infrastructure for arts & humanities scholars"

DARIAH-EU

https://www.dariah.eu/

How can humanities researchers (of any level of digital competence) access and effectively use digital research tools?

- Develop a digital agenda designed for and by humanities researchers
- Connect communities and to support digital literacies
- Foster reuse and integration of digital methods, tools and data within a licensed framework with training in their use
- Strengthen meaningful humanities engagement with the European Open Science Cloud
- Ensure mechanisms of reliability, transparency and quality review
- Give the humanities protagonism and visibility in the ethical and policy issues which our engagement with digital brings

DARIAH

"The pan-European infrastructure for arts & humanities scholars"

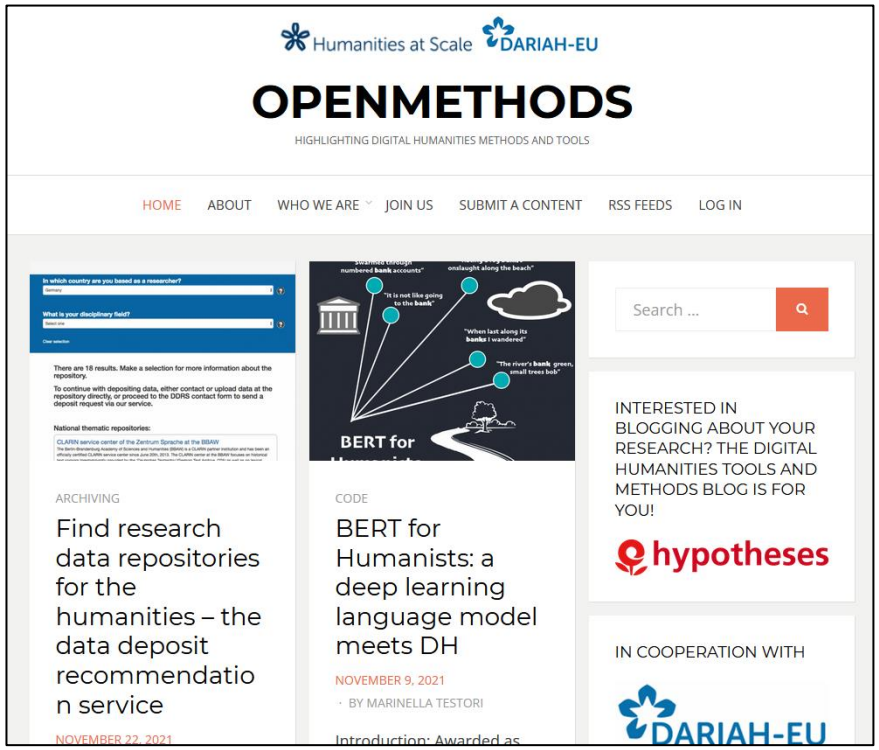
DARIAH-EU

https://www.dariah.eu/

How can humanities researchers (of any level of digital competence) access and effectively use digital research tools?

- Information about digital humanities methods and tools
- Curated open access blog posts, papers, publications, videos, podcasts …
- A variety of methods from text annotation to data visualization, and from sentiment analysis to deep learning language models
- Transdisciplinary collaborations through a light-touch community workflow which draws on multilingual content
- Fosters critical and theoretical engagement with digital tools and methods

Open Methods



https://openmethods.dariah.eu/

Taxonomy of Digital Research Activities in the Humanities, TADIRAH

- 1 Capture
  - conversion
  - data recognition
  - discovering
  - gathering
  - imaging
  - recording
  - transcription
- 2 Creation
  - designing
  - programming
  - web development
  - writing
- 3 Enrichment
  - annotating
  - cleanup
  - editing
- 4 Analysis
  - content analysis
  - network analysis
  - relational analysis
  - spatial analysis
  - structural analysis
  - stylistic analysis
  - visualization
- 5 Interpretation
  - contextualizing
  - modeling
  - theorizing
- 6 Storage
  - archiving
  - identifying
  - organizing
  - preservation
- 7 Dissemination
  - collaboration
  - commenting
  - communicating
  - crowdsourcing
  - publishing
  - sharing

http://tadirah.dariah.eu

What role does critical thinking have in relation to Digital Research Infrastructures and methods?

- "infrastructure" … "has much of the same scale, complexity, and general cultural impact as the idea of "culture" itself"
- the digital humanities are "uniquely placed to interpret and critique culture at the level of infrastructure"

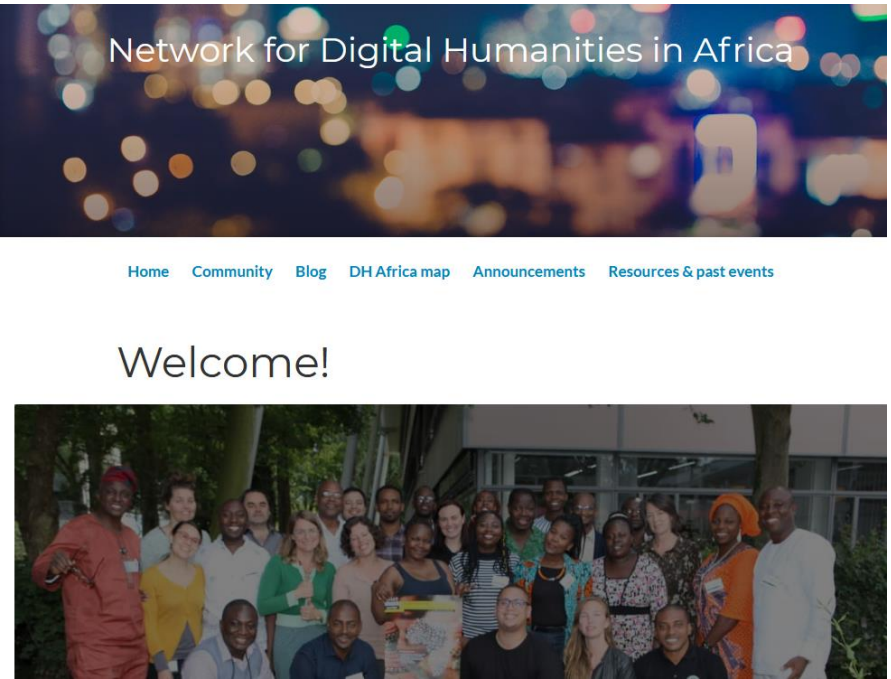
https://liu.english.ucsb.edu/what-infrastructure-means-to-me-kings-college/

- Critical Infrastructure studies collective: research connecting thinking about infrastructures and society across numerous fields including science and technology studies, design, ethnography, feminist theory, critical race and ethnicity studies and the creative arts

https://cistudies.org/

How global/multilingual is our digital research, really?

- Urszula Pawlicka Deger 2020:
  - "How does the global distribution and dimensions of infrastructure contribute to the shaping of global representation of the field?"



GLOBAL OUTLOOK::DIGITAL HUMANITIES

GLOBAL OUTLOOK :: DIGITAL HUMANITIES :: GO::DH

Minimal Computing

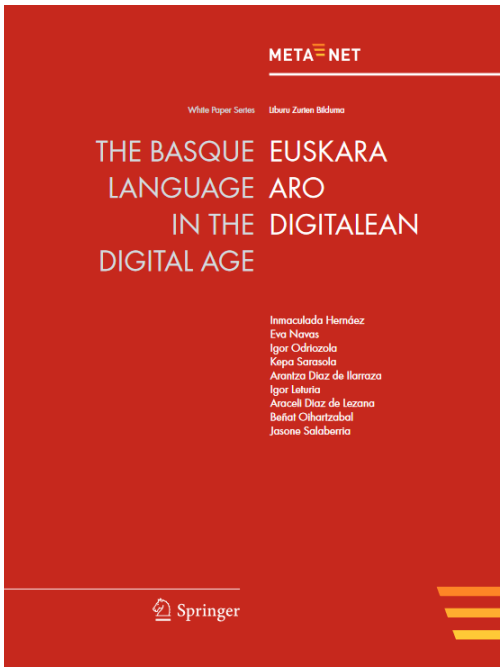
a working group of GO::DH

MULTILINGUALISM IN DIGITAL SPACES

How might a DH volume on multilingualism in digital spaces look?

- case studies by language or language family,
- landscape studies of the underlying digital infrastructure which shapes our communication practices,
- evaluation of the 'multilingual readiness' of DH tools,
- best practices for multilingual DH platform design,
- analysis of DH multilingual data dynamics or competitions to design linguistically inclusive solutions to language challenges.

META.NET
MONITORING DIGITAL LINGUISTIC DIVERSITY IN 2012



Language technologies:

- Machine Translation
- Speech Processing
- Text Analysis
- Speech & Text Resources

Speech and Text Resources

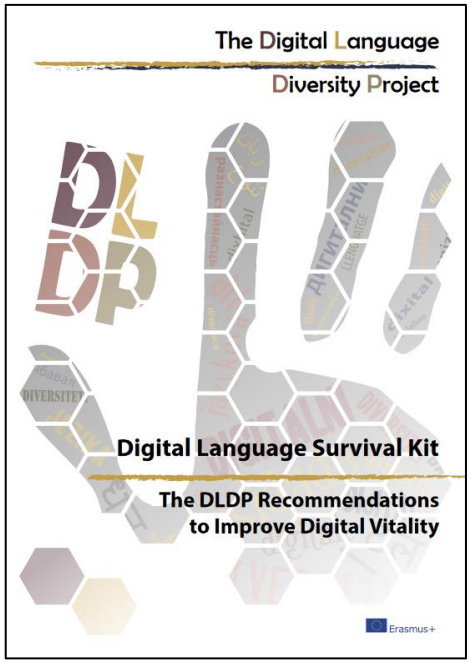
Excellent support	Good support	Moderate support	Fragmentary support	Weak/no support
	▪ English	▪ Czech ▪ Dutch ▪ French ▪ German ▪ Hungarian ▪ Italian ▪ Polish ▪ Spanish ▪ Swedish	▪ Basque ▪ Bulgarian ▪ Catalan ▪ Croatian ▪ Danish ▪ Estonian ▪ Finnish ▪ Galician ▪ Greek ▪ Norwegian (Bokmål, Nynorsk) ▪ Portuguese ▪ Romanian ▪ Serbian ▪ Slovak ▪ Slovene	▪ Icelandic ▪ Irish ▪ Latvian ▪ Lithuanian ▪ Maltese ▪ Welsh

Cross-language comparison http://www.meta-net.eu/whitepapers/key-results-and-cross-language-comparison

http://www.meta-net.eu/

EUROPEAN LANGUAGE GRID

DIGITAL LINGUISTIC DIVERSITY FRAMEWORKS

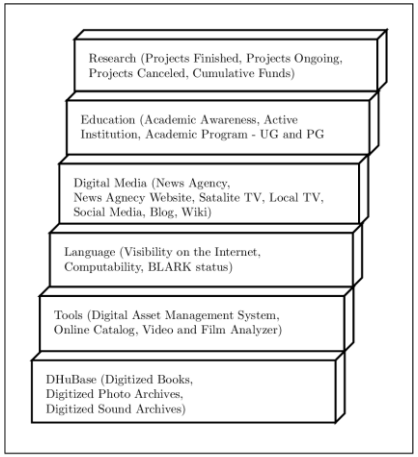


http://www.dldp.eu/en

- *Digital Capacity* (Digital Literacy; Character Encoding, Input and Output Methods; Availability of Language Resources);
- *Digital Presence and Use* (Use for E-Communication; Use on Social Media; Availability of Internet Media; Wikipedia) and;
- *Digital Performance* (Availability of Internet Services; Localised Social Network; Localised Software: Operating Systems and Basic Software; Machine Translation Services; Dedicated Internet Top-Level Domain).

Digital Humanities Readiness Assessment Framework: DHuRAF

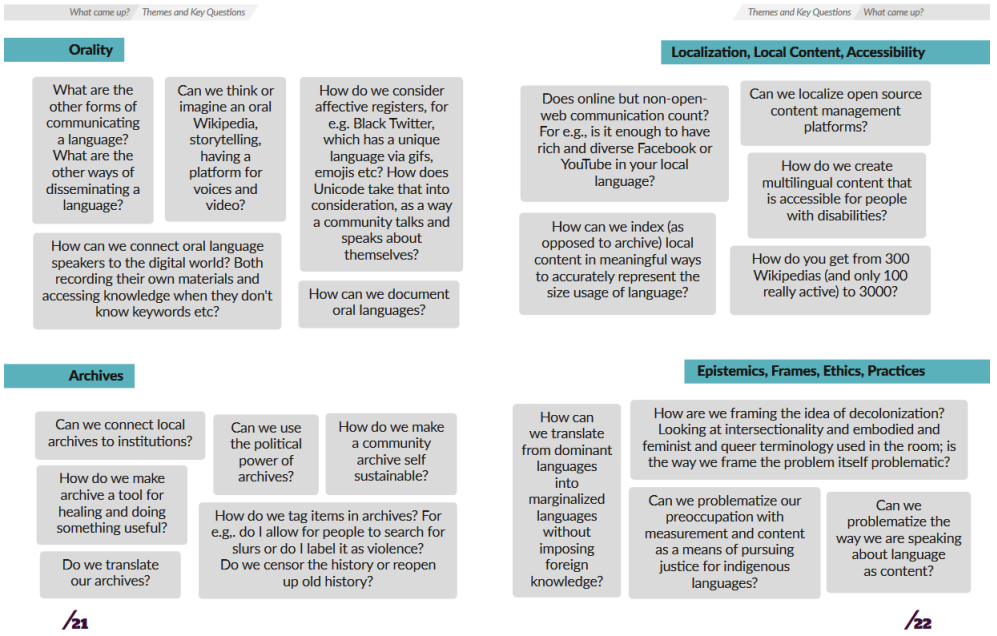
Hossein Hassani, Emir Turajlic, Kemal Taljanovic



https://arxiv.org/abs/1902.06532

Challenge 4

DECOLONISING INTERNET'S LANGUAGES



https://whoseknowledge.org/initiatives/decolonizing-the-internet/ and
Summary report https://whoseknowledge.org/wp-content/uploads/2020/02/DTIL-Report.pdf

Challenge 4

ADHO DH2019 Workshop "Towards Multilingualism In Digital Humanities: Achievements, Failures And Good Practices In DH Projects With Non-latin Scripts"

Date & Time
8th July 2019, 9:00 - 18:00

Workshop organizers:
Martin Lee; martin.lee@fu-berlin.de
Dr. Cosima Wagner; cosima.wagner@fu-berlin.de
Freie Universität Berlin
University Library/Campus Library (Germany)

- ADHO DH2019 Workshop "Towards Multilingualism In Digital Humanities: Achievements, Failures And Good Practices In DH Projects With Non-latin Scripts"
- Exposé
- Workshop Format
- Notes
- Works cited
- Workshop timetable and abstracts
  - 09:00 - 09:15 Introduction to the workshop
  - 09:15 - 12:05 NLS & DH: discussing workflows
    - 09:15 - 09:40: „Ancient Egyptian Hieratic Script – Aspects of Digital Paleography for a NLS"
    - 09:40 - 10:05: "Arabic Script in Digital Humanities Research Software Engineering"
    - 10:05 - 10:30 "SHINE: A Novel API Standard & Data Model to Facilitate the Granular

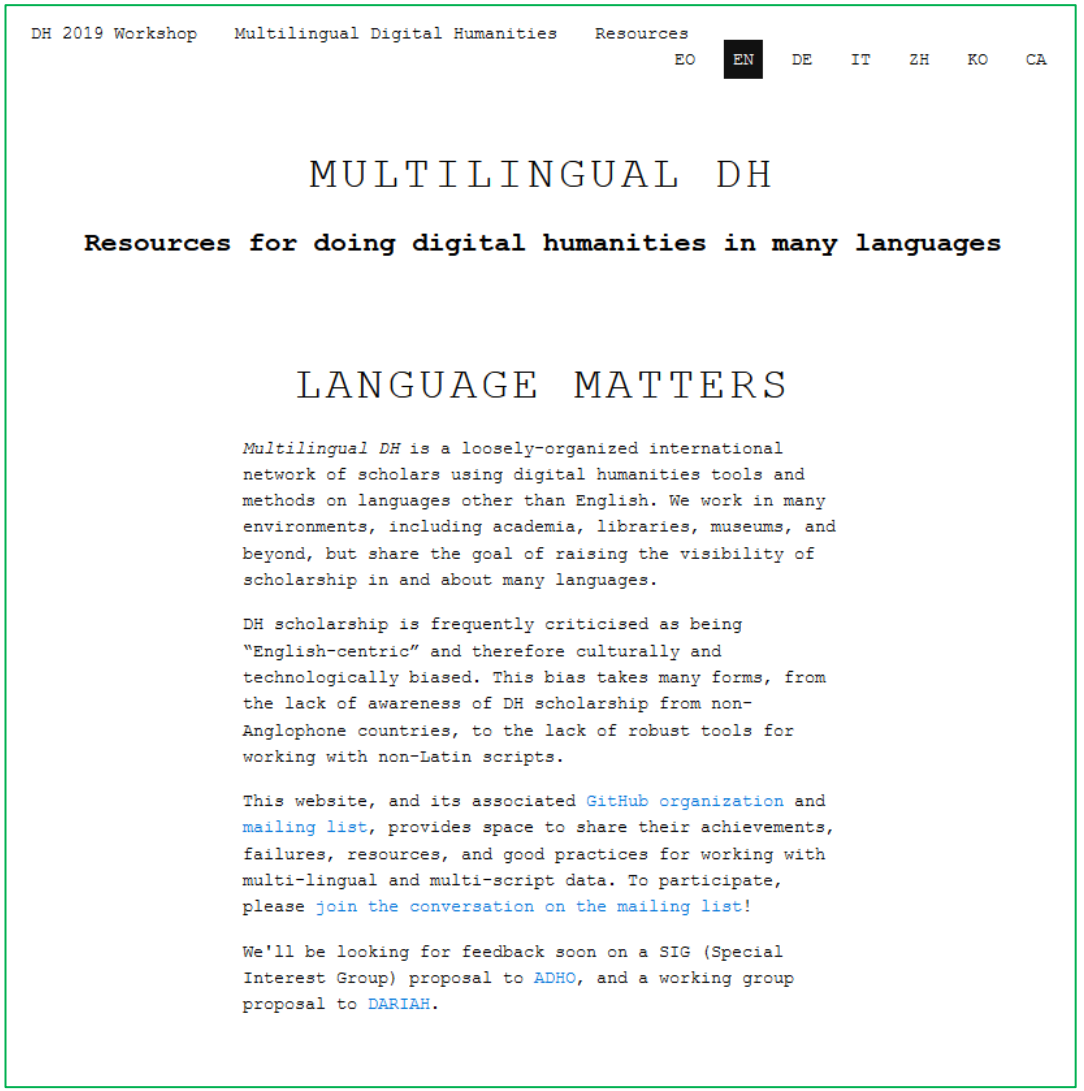
https://hackmd.io/@omhCJrc6T06AFuwMfCrsqw/ry0yFF1oE?type=view

The workshop addressed challenges in a whole range of areas, including

- multilingual data curation,
- Optical Character Recognition (OCR),
- digital research ecosystem design,
- data/text mining,
- and machine translation.

MULTILINGUAL DH

https://multilingualdh.org/



DH 2019 Workshop Multilingual Digital Humanities Resources

EO EN DE IT ZH KO CA

MULTILINGUAL DH

Resources for doing digital humanities in many languages

LANGUAGE MATTERS

Multilingual DH is a loosely-organized international network of scholars using digital humanities tools and methods on languages other than English. We work in many environments, including academia, libraries, museums, and beyond, but share the goal of raising the visibility of scholarship in and about many languages.

DH scholarship is frequently criticised as being "English-centric" and therefore culturally and technologically biased. This bias takes many forms, from the lack of awareness of DH scholarship from non-Anglophone countries, to the lack of robust tools for working with non-Latin scripts.

This website, and its associated GitHub organization and mailing list, provides space to share their achievements, failures, resources, and good practices for working with multi-lingual and multi-script data. To participate, please join the conversation on the mailing list!

We'll be looking for feedback soon on a SIG (Special Interest Group) proposal to ADHO, and a working group proposal to DARIAH.

Disrupting Digital Monolingualism

Workshop shape:

1. **Showcase existing initiatives**
2. **Collectively propose new models and solutions**

- The workshop was hosted by the Language Acts & Worldmaking project
  - with the support of the Cross-Language Dynamics: Reshaping Community project,
  - both projects funded by the AHRC as part of its Open World Research Initiative.
- Led by Paul Spence and Renata Brandao, in collaboration with Naomi Wells

https://languageacts.org/digital-mediations/event/disrupting-digital-monolingualism/

DISRUPTING DIGITAL MONLINGUALISM

1. LINGUISTIC AND GEOCULTURAL DIVERSITY IN DIGITAL KNOWLEDGE INFRASTRUCTURES

How do digital knowledge infrastructures embed multilingual or language-focused practice in social and technical terms? How might digital infrastructures reinforce or redistribute linguistic and cultural influence or authority?

2. WORKING WITH MULTILINGUAL METHODS AND DATA

How well do digital methods and data practices currently model multilingual dynamics, and what new forms of digital, linguistic and cultural criticism are needed to interpret digital methods or data-driven approaches to languages and multilingualism?

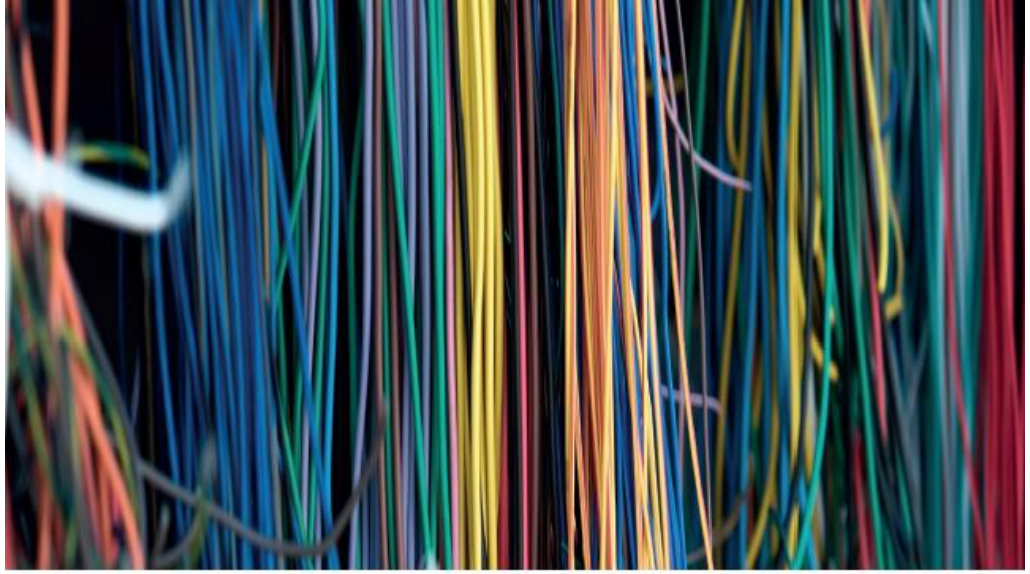
3. TRANSCULTURAL AND TRANSLINGUAL APPROACHES TO DIGITAL STUDY

How can we effectively study intercultural, plurilingual or transborder perspectives using digital tools? What is the role of digital practitioners (in research or industry) and language experts (such as translators, linguists and modern language researchers) in studying these interactions?

4. ARTIFICIAL INTELLIGENCE, MACHINE LEARNING AND NATURAL LANGUAGE PROCESSING IN LANGUAGE WORLDS

What challenges do AI, machine learning and NLP pose for language research fields and professions? What opportunities and limitations does AI enact on the dynamics of linguistic diversity or intercultural communication?

https://languageacts.org/digital-mediations/event/disrupting-digital-monolingualism/



DISRUPTING DIGITAL MONOLINGUALISM

A report on multilingualism in digital theory and practice

Open World Research Initiative OWRI
Arts & Humanities Research Council
Cross-Language Dynamics OWRI
language acts and worldmaking

Thanks:

Arantxa Iriarte

Mikel Iruskieta & workshop organisers

Naomi Wells, James Smithies, Urszula Pawlicka Deger

Renata Brandao and other colleagues on the *Language Acts & Worldmaking* project

https://languageacts.org/

Contact:

Paul Spence

Department of Digital Humanities

King's College London

paul.spence@kcl.ac.uk

Twitter: @politonaiz