FAIR ENOUGH?
BUILDING DH RESOURCES IN AN UNEQUAL WORLD

Antonio Rojas Castro
Berlin-Brandenburgische Akademie der Wissenschaften (BBAW)
Goals

• A critique of the Digital Humanities as a universalist, not situated, and scientific field whose epistemological frameworks, methods and tools can be applied anywhere, anytime and under all conditions.

• An examination of the FAIR Principles drawing on examples derived from two North-South DH projects: The Programming Historian en español (PHes) and the Proyecto Humboldt Digital (ProHD).
Contents

• The World after COVID-19
• Cultural criticism in the Digital Humanities
• The FAIR Principles
• Caveats and dangers
• Conclusions
THE WORLD AFTER COVID-19

• Social distance and confinement
• Closing of borders and control on movement
• Persistent inequalities stemming from capitalism, colonialism, and patriarchy
• Perennial issues of digital access and inclusivity
“Epidemics (the new coronavirus is the most recent manifestation of them) only become a serious global problem when the populations of the richer countries of the Global North are affected. This is what happened with AIDS. In 2016, malaria killed 405,000 people, the vast majority in Africa, and that was not news.”
The Digital Humanities Divide

• Digital Divide
• Governance of digital infrastructures
• Development of standards
• Code Hegemony
• Multicultural and linguistic issues, gender, and the representation of minorities

Based on Fiormonte, Numerico & Tomasi (2015)
In the borderlands

- “Rather than initiating a one size fits all global model, we need to imagine a global digital humanities that lives in the borderlands, a place of connection and contradiction and, mostly importantly, a place that does not try to centralize itself” (Earhardt, 2018).
Racialized and colonial violence

- “The landscape of information and communication technology including the tools used in DH projects, are fully implicated in racialized violence and environmental destruction: from extraction to production, and from consumption to disposal of digital technologies” (Noble, 2019).
- “[…] the need for digital archives that resist colonial violence in content and method, mediating in the gaps and silences in the digital cultural record that can be filled with extant sources” (Risam, 2019).
The FAIR Principles

- **FAIR** = Findable, Accessible, Interoperable, Reusable
- Set of guiding principles and practices created by FORCE 11 in 2014.
- Endorsed by the European Commission - Horizon 2020 Program and by organizations like LIBER and RDA.
- Addressed to both machines and humans.
- To be applied to data, metadata, tools, vocabularies, and infrastructures (Wilkinson et al., 2016).
The FAIR Principles

• “The 15 facets of the FAIR principles are all short sentences. Their brevity gives the impression that they are all items that can be checked off. However, our analysis shows that the FAIR principles are much trickier than this” (Dunning, Smaele & Böhmer, 2017).
To be Findable:

- F1. (meta)data are assigned a globally unique and persistent identifier
- F2. data are described with rich metadata (defined by R1 below)
- F3. metadata clearly and explicitly include the identifier of the data it describes
- F4. (meta)data are registered or indexed in a searchable resource

Caveats:

- Persistent identifiers are not free
- What does “rich” metadata mean?
- Compliance with third parties resources
F1. (meta)data are assigned a globally unique and persistent identifier
F1. (meta)data are assigned a globally unique and persistent identifier

Handle.Net® Registry

Payment

CNRI requires a one-time $50 Registration Fee (for each new prefix allotted), and an Annual Service Fee of $50 per prefix, payable to CNRI. If you wish, you may choose to prepay the Annual Service Fee for up to five years. Payment can be made by check, credit card, wire transfer, or (in some cases) PayPal.

Before making a payment, both new users who are requesting a prefix and users who wish to pay the annual service fee for a prefix they have already been allotted must complete the Prefix Registration Form. Registration generates an email confirmation message that includes an HNR Transaction Number which is required for processing payment. After you register you will be contacted by the HNR Administrator who will provide you with payment instructions and an invoice. For your information, the types of payment CNRI accepts are listed below.

Credit Cards

CNRI accepts Visa, MasterCard, and American Express. Please call the Handle.Net Registry Administrator, Jane Euler, at (703) 620-8990 with your credit card information. You may also fax that information to (703) 620-0913. If sending by fax, please include your HNR Transaction Number and/or prefix number, organization name, and phone number on the cover sheet. Credit card payments may be made using PayPal as described in the section below.

PayPal
F2. data are described with rich metadata
F2. data are described with rich metadata
“Archival description and metadata favour the ‘creator’ […] people who have traditionally not ‘created’ what is considered a record are rarely included in the archives. Even when they are included, marginalized populations, such as the enslaved, do not exist as ‘creators’ of documents—instead they can be found as items in a slave ship’s cargo list, a plantation’s holdings, an owner’s will, or a newspaper runaway ad.” (Levi, 2019)
F3. (meta)data are registered or indexed in a searchable resource
F3. (meta)data are registered or indexed in a searchable resource
The FAIR Principles

To be Accessible:

• A1. (meta)data are retrievable by their identifier using a standardized communications protocol
  • A1.1. the protocol is open, free, and universally implementable
  • A1.2. the protocol allows for an authentication and authorization procedure, where necessary
• A2. metadata are accessible, even when the data are no longer available

Caveats:

• Digital divide is made up of barriers to access and lack of hardware and digital competences
• Web accessibility needs to take into account disability, age and contextual factors
To be Accessible

Individuals using the internet in 2017 (% of population)

Source: World Bank Data
To be Accessible

- Plan de Acción eLAC 2015
- Regional leaders: Chile, Argentina, Brazil and México
- In 66.6% of Latin American countries, less than 50% of population have access to internet (Bisset, Grossi & Borsetti, 2015)
- “the regional lack of knowledge infrastructure” (Wrisley, 2019).
To be Accessible

Leading the web to its full potential

STANDARDS  PARTICIPATE  MEMBERSHIP  ABOUT W3C

ACCESSIBILITY

On this page — why — what — how — learn more — wai — recent press clippings — current status of specific

The power of the Web is in its universality.
Access by everyone regardless of disability is an essential aspect.
Tim Berners-Lee, W3C Director and inventor of the World Wide Web

The Web is fundamentally designed to work for all people, whatever their hardware, software, language, location, or ability. When the Web meets this goal, it is accessible to people with a diverse range of hearing, movement, sight, and cognitive ability.

Thus the impact of disability is radically changed on the Web because the Web removes barriers to communication and interaction that many people face in the physical world. However, when web sites, applications, technologies, or tools are badly designed, they can create barriers that exclude people from using the Web.

Accessibility is essential for developers and organizations that want to create high quality websites and web tools, and not exclude people from using their products and services.

The mission of the Web Accessibility Initiative (WAI) is to lead the Web to its full potential to be accessible, enabling people with disabilities to participate equally on the Web.
The FAIR Principles

To be Interoperable:

- I1. (meta)data use a formal, accessible, shared, and broadly applicable language for knowledge representation
- I2. (metada)data use vocabularies that follow FAIR principles
- I3. (meta)data include qualified references to other (meta)data

Caveats:

- Being interoperable between different information systems is more difficult than within one
- Using external vocabularies depends on human interpretation and expertise
- Using identifiers to disambiguate terms is time consuming
To be interoperable

- Metadata
  - YAML
- Structure
  - Markdown
- Storage
  - GitHub
  - Git
- Validation
  - Travis
- Publishing
  - Jekyll
  - HTML
  - CSS
To be interoperable

- Sharing
  - OAIPMH

- Structure
  - XML
  - XSLT

- Metadata
  - Dublin Core

- Storage
  - SQL

- File formats
  - PDF
  - JPEG
  - PNG

- Publishing
  - HTML
  - CSS

- Preservation
  - OAI
To be interoperable

- Validation
  - DTD
- Storage
  - eXistDB
- Metadata and structure
  - TEI
- Publishing
  - ODD
  - HTML
  - CSS
“Interoperability is] the creation of compatibility standards for communication networks. Scholars who have studied the interconnection of various components of communication networks often use three concepts – infrastructure, platforms, and network effects – to describe the important role of standards that enable compatibility and interoperability. […] They emphasize the potential for a heap of standardized components to combine into a cohesive and flexible network that can, in turn, sustain more complex social and economic activity” (Russell, 2014).
The FAIR Principles

To be Reusable:

- R1. meta(data) are richly described with a plurality of accurate and relevant attributes
- R1.1. (meta)data are released with a clear and accessible data usage license
- R1.2. (meta)data are associated with detailed provenance
- R1.3. (meta)data meet domain-relevant community standards

Caveats:

- Metadata quality and completeness depend on human interpretation and are both time consuming
- Licenses that allow remix are not a guarantee for reuse: resources need to meet the audience’s needs
- Which community and whose standards?
R1.1. (meta)data are released with a clear and accessible data usage license and R1.2. (meta)data are associated with detailed provenance

- Translation can be seen as reusability example; however, instructive texts demand a translation of the sense in order to produce a text that is acceptable, useful and easy to read for the target audience.
- Function rules over equivalence: translators can rephrase, amplify and annotate text, and adapt images, code snippets, and datasets.
“The problem of prioritizing academic integration at a global level, over any other aspect of the use of standards involves four major issues. First, it is a factor in the disqualification of regional variants in favor of a single model of project development; second, as a consequence of this, the normative use of standards and good practices reinforces the region's technological and academic dependence on Anglo-American countries; third, it subordinates regional communities to a central community [...] ; and, four, it prevents the creation of a research community by emphasizing the practical implementation of projects.” (Priani, 2019).
Who is designing and managing the project and who is executing it? Who is giving or receiving training?

Which quality standards, metadata schemes and format files are used and why?

Where is the infrastructure (hardware and software) built and maintained? Where are the servers installed?
Eurocentrism

“There is an extraordinary continuity between the different forms in which Eurocentric knowledge has legitimised the civilising/normalising mission on the basis of the deficiencies - deviations from the normal pattern of the civilised - of other societies. The different historical discourses (evangelization, civilization, the burden of the white man, modernization, development, globalization) all have as their basis the conception that there is a civilizing pattern that is simultaneously superior and normal” (Lander, 2000).
Cultural cloning

• “[…] the preference for the sameness and the construction of homogeneity as an instrument that reinforces privileged positions” (Essed, 2005).
• “[…] social injustice is not only the exclusion of the Other; it is also selective inclusion, and above all the inclusion of more of the same. It is this idea and the practice of preference for the identical […]” (Essed, 2005).
Extractivism

“Extractivism is theft, robbery, pillage. It is a way of being and being in the world, where it appropriates others without consent and without thinking or worrying about the negative impact it generates on the lives of other living beings (human and non-human). The logic of the ontological extractor's attitude is: 'As long as it benefits me, I don't care about the negative consequences on other living beings (human or non-human)'. ” (Grosfoguel, 2016).
Conclusions

• DH scholars need to imagine ways of mitigating the impact of inequalities on project development, communication, training, collaboration and engagement.

• DH is not only about building; the field already has a long tradition of cultural criticism, but DH scholars and practitioners should find time to read and quote each other’s works.

• DH scholars should engage critically with guidelines for building digital resources; compliance to quality standards is a desirable goal, but most of guiding principles are open to interpretation, debate, and negotiation.

• DH projects that involve North-South cooperation must avoid eurocentrism, cultural cloning, and extractivism; they should support an “ecology of knowledge” instead.
Vielen Dank! Thank you! ¡Gracias!

Join the conversation on Twitter: send your questions and comments to @RojasCastroA and use #DHBerlin
Bibliographic References


• Essed, Philomena, “Racisme et préférence pour l’identique : du clonage culturel dans la vie quotidienne”. Translated by Brigitte Marrec, Actuel Marx, volumen 38, issue 3, 2005. DOI : 10.3917/amx.038.0103


Bibliographic References


- Priani Saisó, Ernesto, “Codificación y buenas prácticas. Crítica a la delimitación de las humanidades digitales en América Latina”, *Relaciones Estudios de Historia y Sociedad*, volumen 40, issue 158, 2019. [http://dx.doi.org/10.24901/rehs.v40i158.343](http://dx.doi.org/10.24901/rehs.v40i158.343)


Image credits

- Slide 1: image by Justin Kunimune - Own work, CC BY-SA 4.0, From https://commons.wikimedia.org/w/index.php?curid=66467596
- Slide 2: image by Joaquín Torres García (1878-1949) / Public domain from Wikimedia Commons https://commons.wikimedia.org/wiki/File:Joaqu%C3%ADn_Torres_Garc%C3%ADa_-_Am%C3%A9rica_Invertida.jpg
- Side 3-5: image by PIRO4D from Pixabay
- Slide 6: cover image of The Digital Humanist (Fiormonte, Numerio y Tomasi, 2020) from Punctum Books
- Slide 7: GO::DH logo and screenshot from Around DH 2020
- Slide 8: illustration from Primer nueva Coronica y buen gobierno de Felipe Guaman Poma de Ayala
- Slide 9: image by SangyaPundir - Own work, CC BY-SA 4.0, From https://commons.wikimedia.org/w/index.php?curid=53414062
- Slide 12, 14, 15 & 29: screenshots from the Programming Historian en español
Image credits

- Slide 13: screenshot from Handle.Net Registry
- Slide 16: screenshot from OER Commons
- Slide 17: screenshot from DOAJ
- Slide 22: screenshot from W3C – Accessibility
- Slide 27: cover of Open Standards and the digital age by Andrew L. Russell
- Slide 33: image by Reimund Bertrams from Pixabay
- Slide 34: image by Jan W. from Pixabay