For We Are Many: The DHTech White Paper on Research Software Engineering for Digital Humanities

Robert Casties, Alexander Czmiel, Julia Damerow, Max Ionov, Albert Meroño Peñuela, Steve Ranford, Catherine Smith, Malte Vogl

NL-RSE19 conference, Amsterdam, 20-11-2019

Who am I?  [https://www.albertmeronyo.org/](https://www.albertmeronyo.org/)

**Academic**

- Postdoctoral researcher, Computer Science, Vrije Universiteit Amsterdam
- AI, Knowledge Graph Construction, Web Query Languages, Web APIs, eScience/DH workflows

**Digital Humanities RSE**

- [CLARIAH](https://www.clariah.eu) Lead Engineer for structured data (TB IG LOD)
- TabLinker, COW, cattle, grlc, midi2rdf
DHTech

Who are we?

- Established at DH 2017 Montreal
- International grass-roots community of Digital Humanities software engineers
- Support the development and reuse of software in the Digital Humanities

Involved in recent RSE community developments

- RSE 2019, RSEConUK 2019, DH-RSE (Germany), NL-RSE
- Address common and DH-specific RSE issues
The White Paper: Background

- DH 2019, Utrecht, NL
- Workshop “I’m the one building the tool!”
  - Software development and academic careers
  - Open Reputation Systems
  - Collaboration vs. data-protection
  - Publication of preliminary results / open source first
  - Visibility of the technological side of DH (digital humanities)
The White Paper: Goals

Awareness of the **importance** of DH RSEs, a clear **career path**, and **academic recognition** (e.g. inadequate publication systems for software and data)

Almost no DH project can be realized without someone who understands the approaches and methods of the research domain and is able to conceptualize and implement the digital or computational part

[https://dh-tech.github.io/dhrse-whitepaper/]
Communication

- Skill of DH RSEs: ability to **mediate** between the technological world and humanities scholars
- Tools need to be produced along the research question
- **Discipline assumptions/tacit knowledge**: never explicitly stated unless asked about them directly
- Continuous RSE - DH scholar dialog needed to **transfer** this knowledge to new RSEs and others
Career Paths

Various roles of DH RSE:

- Software development
- Data analysis
- Representation of results
- Data management

But unclear career path, due to unclear role in research community

- PhD → Postdoc → Professorship? (but software/data != papers)
- Alternatives are lacking
- Call for standardization: keep experienced DH RSE, attract new
Recognition

- No formally agreed way to credit RSE work
- Projects of different scale credit differently

Dependency on career paths

- Classic academic ladder: software & data should count in position applications
- Software & data citations (e.g. FORCE11)
- Include RSE as co-authors
Recognition: Co-author Roles

CRediT – Contributor Roles Taxonomy

CRediT (Contributor Roles Taxonomy) is a high-level taxonomy, including 14 roles, that can be used to represent the roles typically played by contributors to scientific scholarly output. The roles describe each contributor's specific contribution to the scholarly output.

- **Project administration** - Management and coordination responsibility for the research activity planning and execution.
- **Resources** - Provision of study materials, reagents, materials, patients, laboratory samples, animals, instrumentation, computing resources, or other analysis tools.
- **Software** - Programming, software development; designing computer programs; implementation of the computer code and supporting algorithms; testing of existing code components.
- **Supervision** - Oversight and leadership responsibility for the research activity planning and execution, including mentorship external to the core team.
- **Validation** - Verification, whether as a part of the activity or separate, of the overall replication/reproducibility of results/experiments and other research outputs.
- **Visualization** - Preparation, creation and/or presentation of the published work, specifically visualization/data presentation.
- **Writing – original draft** - Preparation, creation and/or presentation of the published work, specifically writing the initial draft (including substantive translation).
- **Writing – review & editing** - Preparation, creation and/or presentation of the published work by those from the original research group.
Funding

Applications: RSEs often detached from planning and writing phases of grant application process → Inadequate resource allocation for software

Reviews by funding bodies: appropriate processes for reviewing technical aspects; RSEs can work with applications to answer these questions
Education

- DH RSE community is very **heterogeneous**, different backgrounds
- Lack of formal CS/SE training leads to buggy, poorly maintainable software
  - Hard to reproduce
  - Incorrect/inaccurate results
- Many DH **education programs** emerging, not necessarily focusing on RSE technical issues
  - Programming **skill**: using a library vs. writing a library
- Resource and course information **scattered** across sources
- DH RSE should **get involved** in DH software educational programs
Community

- DH RSE = DH ∩ RSE
- All RSE face common issues, but DH RSE currently not well represented in either community
- DHTech, DH-RSE
- Proposal: Special Interest Group (SIG) to the Alliance of Digital Humanities Organizations (ADHO)
Channels & more

https://dh-tech.github.io/

https://github.com/dh-tech

https://dh-tech.github.io/join/
Thank you

Questions?

https://dh-tech.github.io/
@albertmeronyo