[SORSE] Little Minions in Archeology, Wikidata, and better scientific collaborations

Report of Contributions

Contribution ID: 1

Type: not specified

Wikidata as a research tool for data modelling and integration in the humanities - Examples from the German Wikimedia Fellow Program Free Knowledge

Wednesday, March 3, 2021 10:00 AM (30 minutes)

Research Software Engineers and their work is becoming increasingly important in digital humanities research just as digital humanities gain traction. Not just developing and extending software are important for this process, data creation, data publication and data management are equally critical; as the idiom goes, "data matures like wine, applications like fish". Therefore, RSEs need to store their 'wine' FAIRly to make it accessible in the future. In this talk we present Wikidata as a data integration hub for FAIR Linked Open Data (LOD), that can be used with many research software tools, e.g. SPARQLing Unicorn, QGIS plugin and Scholia. Wikidata is the information storage system in the Wikimedia Universe. In particular, Wikidata is ideal for data modeling and immediate publishing as any user can create their own (semantic) data model and map it to Wikidata. A number of tools and APIs already exist for applying Wikidata to open research. As an example use case we will present two projects from the German Wikimedia Fellow-Programme Free Knowledge. The first, Irish Ogham Stones in the Wikimedia Universe, focuses on linking different data sets on Ogham Stones within the Wikimedia Universe and the second, Smashed Dishes - Archaeological Sources in Wikidata, aims to create possibilities for amateur archaeologists to collect and publish their data. As a conclusion, we will discuss the advantages we found in using Wikidata. These include enhanced data re-usability and availability, FAIRly published data and data and possibilities for data integration between diverse data sources and research software. Data and software integration may lead to Linked Open Data and, the resulting interoperability opens up new possibilities for RSEs.

Presenters: Mr THIERY, Florian (Römisch-Germanisches Zentralmuseum); SCHMIDT, Sophie (Deutsches Archäologisches Institut)

Session Classification: Talks

Contribution ID: 2

Type: not specified

A new distributed data analysis framework for better scientific collaborations

Wednesday, March 3, 2021 10:30 AM (15 minutes)

A common challenge for projects with multiple involved research institutes is a well-defined and productive collaboration. All parties measure and analyze different aspects, depend on each other, share common methods, and exchange the latest results, findings, and data. Today this exchange is often impeded by a lack of ready access to shared computing and storage resources. In our talk, we present a new and innovative remote procedure call (RPC) framework. We focus on a distributed setup, where project partners do not necessarily work at the same institute, and do not have access to each others resources.

We present an application programming interface (API) developed in Python that enables scientists to collaboratively explore and analyze sets of distributed data. It offers the functionality to request remote data through a comfortable interface, and to share and invoke single computational methods or even entire analytical workflows and their results. The prototype enables researchers to make their methods accessible as a backend module running on their own infrastructure. Hence researchers from other institutes may apply the available methods through a lightweight python or Javascript API. In the end, the overhead for both, the backend developer and the remote user, is very low. The effort of implementing the necessary workflow and API usage equalizes the writing of code in a non-distributed setup. Besides that, data do not have to be downloaded locally, the analysis can be executed "close to the data" while using the institutional infrastructure where the eligible data set is stored.

With our prototype, we demonstrate distributed data access and analysis workflows across institutional borders to enable effective scientific collaboration.

This framework has been developed in a joint effort of the DataHub and Digitial Earth initiatives within the Research Centers of the Helmholtz Association of German Research Centres, HGF.

Presenter: Mr SOMMER, Philipp S. (Helmholtz Zentrum Geesthacht (HZG), Institute for Coastal Research)

Session Classification: Talks

Contribution ID: 3

Type: not specified

Little Minions in Archaeology - an open space for RSE software and small scripts in digital archaeology

Wednesday, March 3, 2021 10:45 AM (10 minutes)

In our daily work, small self-made scripts, home-grown small applications and small hardware devices significantly help us to get work done. These little helpers - "little minions" - often reduce our workload or optimise our workflows, although they are not often presented to the outside world and the research community. Instead, we generally focus on presenting the results of our research and silently use our small tools to generate these, without not even pointing to them, especially to the source code or building instructions. In archaeology, research software is underrepresented in conference talks as well as in research publications. We have to promote these little minions in the RSE community and make them visible. That is why we created a Working Group for Little Minions at the CAA (https://caa-international.org) conference for digital archaeological research and created a session that focus on these "little minions" and invite researchers to share their little helpers, so that the scientific community may benefit and - perhaps - create spontaneously special minion interest groups. In such a minion session we are focussing on lightning talks - aka "minion talks" - where a wide range of tools are shared. Each "minion talk" explains the innovative character and mode of operation of the digital tool. Open Science is the aim of the little minions. Therefore, our aim is that all minions are non-proprietary, open and freely available (e.g. GitHub, GitLab, etc.). The last are not obligatory and tools that are not (yet) available can also be accepted.A constantly expanding list of little minions can be found at https://github.com/caa-minions/minions. In this poster we will present our concept to promote research software at scientific conferences and share our experience on the idea.

Presenters: THIERY, Florian (Römisch-Germanisches Zentralmuseum); MENNENGA, Moritz (Lower Saxony Institute of Historical Coastal Research, Germany); VISSER, Ronald (Saxion - University of Applied Sciences, The Netherlands)

Session Classification: Talks