Visions, needs and requirements for (future) research environments: An exploration with ERC Grantee Stephan Schiffels

Katharina Flicker (TU Wien), Bernd Saurugger (TU Wien), Stephan Schiffels (Max Planck Institute)

Researchers are at the very heart of the EOSC: So what do researchers really need to do cutting-edge research? How do they think the EOSC could support them in their endeavours? Let's see what Stephan Schiffels - group leader for population genetics at the Department for Archeogenetics of the Max Planck Institute for the Science of Human History in Jena - has to say.

"Of middle tier researchers, publishing systems, and social outreach in research"

TU Wien: What does your research currently focus on?

SS: I am a population geneticist working in the field of archaeogenetics. We are trying to investigate the human past by looking at genomes from long dead people found in excavations and in archeological contexts. My specific role is to develop methods and to analyse these genomes in ways that can elucidate past movements of people, mixtures or social processes, such as sex biased admixture, or kinship in prehistoric cemeteries. We want to know, for example, how people were related to each other and what these relationships can tell us about their population structure.

TU Wien: What data are you working with to answer your research questions?

SS: I am working mainly with genetic data but also with archaeological and sometimes anthropological data. The genetic data is the biggest part. Thus, we work with next generation sequencing data from living people today and from ancient genomes. Typically, we analyse this data in comparative ways, meaning that we are looking at differences between people's genomes and focus on positions that are different between people and groups. We then try to understand how they stratify across geography, time, ethnic groups, and archaeological cultures. This kind of data is also rather big data: an individual genome is represented by millions of positions that are informative about ancestry and that can be used for genetic comparative analysis.

TU Wien: In your field of research, what kind of opportunities come with big data? What challenges do you face?

SS: There are huge opportunities currently in ancient genomics. In archaeo- or paleo-genetics we have moved towards population scale sequencing. In a single project, we can now analyse hundreds, sometimes thousands of samples that we then use to understand population structure through time. Before that, we analysed only a few samples to full depth. Ramping up the number of individuals and samples used when we do population scale sequencing comes with an entirely different set of questions. However, many genetic metadata and archeological data are often not available in

EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-2018-4, grant Agreement number 831644

machine-readable formats. I am talking FAIR principles here. Findable. Accessible. Interoperable. Reusable. This is crucial, but certainly not currently fulfilled for many such data.

The change is not only in quantity, but also in quality, because we are able to analyse hundreds of genomes instead of only two, we can look into questions of (social) group belonging and procreation or of how ancestry stratifies with wealth and status, as assessed for example using grave goods. In addition, we are now raising questions relating to rather recent history. This also means a shift to engaging more with different scientific communities, such as historians, and also different types of data.

TU Wien: Can you think of any services that might help you face these kind of challenges?

SS: In Germany, there is an initiative called Nationale Forschungsdaten Infrastruktur (NFDI). It involves both humanities and natural sciences. What are we aiming at? Imagine a specific research object you want to analyse and you want to know what kind of analysis has been done on it already. This kind of knowledge management system does not exist yet, but we want it. This is exactly what we are working on with the NFDI4Objects Initiative. The NFDI, however, is a national initiative, but co-coordinating something like that in Europe – or across the world even – would be fantastic.

> "We should aim at standardizing datasets that enable us to link the worlds of natural sciences to that of humanities and support multidisciplinary approaches"

In terms of other services, I would say, we should aim at normalizing and standardizing datasets that enable us to link the worlds of natural sciences to that of humanities and support multidisciplinary approaches, when it comes to the studies of the human past. One would have to think about how to enable researchers from one discipline to make sense of data from another scientific community, or how to translate scientific concepts from one discipline to another. Not a trivial task!

> "We would need access to scripts, tools, primary data etc. to make research reproducible and to put results under scrutiny"

TU Wien: What would you want the EOSC to be or to offer in terms of services?

SS: I think we need services that allow researchers to put in arbitrary data in a computable format and with good documentation. We would need access to scripts, tools, primary data etc. to make research reproducible and to put results under scrutiny. And I think I would like to see something like *eduroam*: everyone uses it and it just works. All I need is to log in.

TU Wien: Finally, considering research environments more generally, what do you need, or want to support your research endeavours?

In Germany, at least, I would say, the situation is quite nice. We typically have access to public funding, for example through the DFG. There are the Max Planck Institutes, and there are great universities. In comparison to other countries there seem to be many jobs in

EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-2018-4, grant Agreement number 831644

academia. However, there are three issues that I do feel passionate about and that should be addressed: the system of having fixed-term contract middle tier researchers, the peer-review and publishing systems and crediting social and public outreach in hiring committees.

> "I would love to see more credits and value given to other forms of publishing, also to public outreach, when it comes to evaluating scientists for grants and jobs"

Middle tier researchers are researchers below the professor or director level, but above PhDs, such as postdocs and group leaders. They are given fixed term contracts with no long-term job security. This kind of policy does not allow for e.g. family planning, which increases gender biases even further. After all, it is still women who are more affected by getting children. This social issue won't be fixed without permanent jobs being available at this level.

The publishing system really has been an issue for a while as well. The traditional system of academic publishing has not changed much during the last 50 years, despite the media landscape having undergone massive changes, with Social Media, blog, online magazines, etc. But for hiring committees, we still are mainly judged by our peer-review publications, and nothing else. On top of that, the peer review process itself is somewhat bizarre. Scientist all over the world peer-review papers for free, all while the authors of papers have to pay huge fees to get their paper published in a journal. Often, that paper is then not even published open access, and readers have to pay a second time to read the paper.

So, I would love to see more credits and value given to other forms of publishing, also to public outreach, when it comes to evaluating scientists for grants and jobs. After all, more and more researchers try to increase their public outreach and explain what they do to the public. They are on Twitter, Facebook, or Instagram. They write blogs, run webpages that explain stuff, or give public talks. Credit or reward systems should take that kind of work into account, so that not only publication records matter but also social media and public outreach.

Stephan Schiffels received his PhD in 2012 from the University of Cologne. He worked as a Postdoc at the Wellcome Sanger Institute in Hinxton (UK). From 2015, until 2020 he was group leader at the Max Planck Institute for the Science of Human History in Jena, and since September 2020 is group leader at the Max Planck Institute for Evolutionary Anthropology in Leipzig. Stephan's research focuses on genomic analyses to learn about the human past.

EOSCsecretariat.eu has received funding from the European Union's Horizon Programme call H2020-INFRAEOSC-2018-4, grant Agreement number 831644