Minutes from the Workshop

'TRAINING FOR OPEN SCIENCE'

Wednesday, June 10th, 2015

Welcome

John Wood – Chair of the RISE Working Group “Science in Transition” – and Volli Kalm – Dean of the University of Tartu – welcomed all participants to the Workshop. John Wood, together with Toivo Maimets – Professor of Cell Biology and Director of the Institute of Molecular and Cell Biology at the University of Tartu – outlined the aims of the workshop, i.e. looking at ways and means of fostering and promoting proper training for the new generation of researchers, who are expected to be able to operate in an Open Science context.

Structure and objectives

To this end, RISE Members and the invited experts presented their contribution, focusing on the ideal settings, methods and communication forms to be used in order to enable young researchers to successfully adapt to, as well as fully exploit the potential of, the on-going transition of an increasingly open science and technology (S&T) sector. The outcome would then be organised so as to feed possible future recommendations on how innovative training methods be streamlined into the ERA and Horizon 2020 objectives.

The format of the workshop therefore entailed a series of presentations followed by in-depth discussion on the most relevant issues, the end of business being dedicated to conclusions and possible points of action.
**Key points of discussion**

- The integration of specific training methods for the new generation of researchers with the current schooling system and academic environment;
- The characteristics of research processes in an Open Science context;
- Opportunities and drawbacks of Open Data;
- The construction of a new research sector.

**Main points of action**

- A new open-data approach should be fed into academic, as well as on-the-field training. Such complex process should start from a major change in the traditional mindset of established institutions. A second "high-school movement" should be created to train data-librarians and other operators to effectively move steady steps into an Open Science environment. A certification system could follow training courses to ensure that new skills are duly recognised;
- Data scientists constitute a profile which should receive more credit, since all other open-data operators "stand on the shoulders of giants". A focus shift from profiles to skills could allow for a more agile re-match between supply and demand for human capital;
- Science has nowadays become data-intensive – data sets deriving from many different sources – and can be depicted as characterised by volume, variety, velocity, and veracity (the so-called "4 Vs"). Given the abundance of data, the key is now to provide for proper data assessment and intelligibility, in order to embed some value added to the final research results. When dealing with Open Data it is vital to always bear in mind the potential commercial value of the pieces of data handled, as the management and transfer thereof will be directly affected. Data curation: proper meta-data creation and management are fundamental to keep the system up and running, although they should be controlled by specialised, ad hoc personnel (not researchers themselves). The maintenance of such a system can be costly but has proved extremely beneficial when put in place in other parts of the world;
- Open access teems with a handful of opportunities but has also some possible drawbacks in store:
  - There is a risk for potential manipulation of data and a consequent need for apt and accurate protection;
  - Data transfer may prove difficult when research data is intertwined with personal data.
  - Intellectual property issues deserve serious consideration by policy- and lawmakers in an open science context, i.e. in terms of individual acknowledgement, regulating copyright in relation to co-authoring and co-writing, etc.;
  - With regard to some geographical as well as disciplinal areas, researchers may be confronted with the payment of variable fees to access – and publish in – open data repositories;
  - Public-funded projects details are sometimes not openly available or published. Research results are mostly reported in a very general fashion, not allowing for further in-depth research on that basis. Moreover, openness should also be applied to research processes, not only final products or services;
  - The system is based on continuous experimentation and processes must be fallible: all operators need to accept and embrace the possibility of an error;
• **Openness does not necessarily have to be everlasting.** As much as protected and copyrighted data is often disclosed after a fixed amount of time, on the other hand openly accessible data could be granted protection at some point to prevent excessive resistance to publication at an earlier stage, and to mitigate the disadvantages of (lack of) acknowledgement in general. In order to create an efficient and more flexible data system, data sharing could be administered also through **opt-in/opt-out preferences**, as long as the latter are carefully established, well-known, and managed;

• Top-down guidance and appropriate funding have yielded to the construction of a solid system in other parts of the world. Good practices could be revised and assembled toward the creation of a new successful model;

• Despite the uncertainties that still surround the topic, policies and regulation in the context of Open Science should be fostered bearing in mind the importance of a **timely intervention** to ensure the maximum positive impact.