Open Workflows for Open Regional Science

[**Dani Arribas-Bel**]

This text is written as a thought piece to accompany a shortened version of Serge Rey's WRSA 2014 presidential address, entitled *Open Regional Science*. Both of them are to be published in the June edition of the NARSC newsletter. The manuscript, as well as this site, are stored in an open repository on this url. A pdf version is available here. Comments and feedback most welcome.

In his WRSA presidential address, Rey (2014) makes a "*call to arms for regional scientists to engage with open science and open source*". His premise is that Regional Science, as a discipline, has kept itself very much distant from recent developments in other fields in the way the scientific process is carried out. In particular, he considers the movement of open science, as contrasted with the notion of *captured* science. This involves a set of practices around the way academics write software (open source), manage data (open data), carry out modeling experiments (open modeling), collaborate (open collaboration) and distribute their results (open publishing). I could not agree more with both Rey's principles and his view of the current state of affairs in Regional Science. In this piece[1], I would like to add to his plea to "go open" with a component of open science that, although not explicitly mentioned in the address, underlies most of its philosophy: open workflows.

The concept of a workflow is something us regional scientists can all relate to, even though we might not realize straight away. After all, we all have a particular set of habits, techniques and tools we use to carry out the diverse activities that involve being a modern scholar. However, not much thought usually goes into the design and adoption of a *good* workflow, let alone of an *open* one. Rather, it is something seldom taught in school that the student (i.e. future researcher/scientist) is usually expected to "just know", as if its acquisition and learning curve were non-existent or just surpased by simply enrolling in a PhD program.

Thinking and designing a good open workflow is not (only) about tool choice. It is much more about shaping every aspect of how we carry out research, from idea generation to its implementation to its dissemination, in a way that can easily be captured, made transparent and, if need be, reproduced again in the future. Tools, of course, play a big role in enhancing (or dampening) this process. Some have been designed from the ground up with these principles in

---

[1]This text draws considerable inspiration from a workshop I co-organized with Thomas De Graaff from the VU University in Amsterdam in January 2014. We developed it in the spirit of its contents, which can be openly accessed at the following url: http://darribas.org/WooW

mind and thus embody them much better than others created under different premises. In any case, the umbrella of the workflow covers beyond the simple choice of one piece of software over another.

Besides their better alignment with the scientific ethos, there are plenty of practical reasons for adopting open workflows. Purely self-interested minds will find they are more efficient in the long run, make collaboration easier and appear as an invaluable sanity keeper when it is required to revisit a research project after some time. However, the true advantages become clear when one considers the colective benefits of an open workflow culture. A structured and documented process facilitates the transition from private to public, lowering the barriers to publish and share with the wider community; transparency and reproducibility make it easier for scientists to fully grasp the methods and results of peers, even when they do not know each other; modularity and portability allow to reuse components from project to project, avoiding "reinventing the wheel".

A good open workflow is not something one "just knows". As most things in life worth pursuing, it takes interest and some effort to master; however, its benefits clearly exceed its costs. Today, we know how Galileo Galilei made his discoveries because he kept detailed research diaries of his steps. We need to ensure the researchers of tomorrow will be able to look back and be able to reconstruct how we reached the insights that have made regional science a meaninful discipline. The answer definitely lies in a more open regional science; but that can only happen with more open workflows.

References

Rey, S. J. (2014). Open Regional Science. *WRSA Presidential Address.*