

Realizing the Potential of Research Data: Subjectification as a Precondition for Reuse

Opening up research data has been linked to a host of benefits, not only for the science system but also for broader processes of economic and social innovation. This is why policymakers and funders have taken steps to promote open data and, more recently, data that is findable, accessible, interoperable, and reusable (FAIR). Yet the positive impacts of open/FAIR data can only be realized if such data, once shared, is actually reused. With practices of data citation still in their infancy, reliable measures of the extent and nature of research data reuse remain elusive; too often, it is simply assumed that once a baseline set of conditions are met, reuse will automatically follow.

While scholarly studies of reuse practices are becoming more common, they often limit themselves to specific disciplinary contexts and rarely consider business or societal reusers. Efforts to build more general models (e.g., Kim and Yoon 2017) have tended to ground themselves in psychological theories of planned behavior, which emphasize the interplay of attitudes and norms but say little about their formation. As a result, the dynamic, affectively charged process by which different groups of social actors come to see themselves as data reusers is not well understood.

Our exploratory qualitative study draws on an extensive literature review as well as interviews with 26 research data reusers and facilitators, which were transcribed and systematically coded. Our findings confirm the importance for reuse of enabling factors such as suitability for purpose, trust, data characteristics (including FAIRness), and capabilities, which we take to include skills, support by data stewards, and infrastructure usability. But our study breaks new ground by drawing on theories of scientific subjectification (Sigl 2019), which focus on the development of particular forms of self-understanding. We argue that this subjectification stage is a precondition for other factors to become relevant in enabling data reuse, as well as a passage point for broader open science advocacy.

These results have concrete implications for open science researchers, practitioners, and policymakers alike. Researchers may want to undertake qualitative studies of specific subjectification contexts such as trainings and mentoring relationships, as well as quantitative studies of the relative strength of individual, interpersonal, and organizational dynamics. Infrastructure providers and other facilitators need to see these subjectification contexts as integral to their strategies for promoting reuse. Finally, while research organizations and policymakers can encourage data reuse through formal structures of evaluation, they should also socialize these structures by identifying and supporting “data communities” (Cooper and Springer 2019) within which the work of subjectification increasingly takes place.

References

- Cooper, Danielle, and Rebecca Springer. 2019. “Data Communities: A New Model for Supporting STEM Data Sharing.” New York: Ithaka S+R. <https://sr.ithaka.org/publications/data-communities>.
- Kim, Youngseek, and Ayoung Yoon. 2017. “Scientists’ Data Reuse Behaviors: A Multilevel Analysis.” *Journal of the Association for Information Science and Technology* 68: 2709–19. <https://doi.org/10.1002/asi.23892>.
- Sigl, Lisa. 2019. “Subjectivity, Governance, and Changing Conditions of Knowledge Production in the Life Sciences.” *Subjectivity* 12: 117–36. <https://doi.org/10.1057/s41286-019-00069-6>.