

Charge of the ORCID Works Metadata Working Group

Works data are an important component of ORCID's mission to help individuals distinguish themselves. Works describe research activities: publications, data sets, or other recognized outputs. The ORCID registry provides a mechanism to connect works data to ORCID accounts, and facilities to store, manage and query sets of descriptions of research works associated with a personal identity. The ORCID Registry can thus aggregate and connect works data from a multiplicity of research information sources.

Works are described using a simple metadata schema, sufficient to identify common work types like scholarly published articles, working papers, etc. It may be helpful to distinguish between the 'data model' of works in ORCID (what elements are used to describe a work, in what vocabularies, etc.), and the 'service model' – viz. how works data are used in queries. The service model is to a fairly large extent delimited by the various user interfaces and APIs that ORCID exposes for works. But it also encompasses a larger spectrum of issues around 'best practices', policy, licensing, guidelines for use, etc. that APIs do not define or constrain.

This group will review the current works data and service models, with the aim to answer the following questions:

- Is the current data model sufficiently expressive – meaning can it describe all the types of research activities that could be recognized as 'works'? Are the attributes (metadata) sufficiently precise, flexible?
- How well does the current model conform to the needs of the works metadata suppliers (publishers, information service providers, etc.)? Is it too onerous or complicated? Does it require explicit 'profiles'?
- How well does the current model conform to the needs of works consumers – which might include both internal use of works data in disambiguation algorithms, but also external users and services such as profile systems.
- While works information is generally considered to be public, do ORCID works data need any usage guidelines/practices/descriptions/provenance?
- What service guarantees (or representations) should ORCID make - if any - about works data with respect to accuracy, currency or completeness?
- What works data provider quality restrictions, if any, should the service impose? E.g. must a citation be 'parsable'? Any minimum number of fields, languages, encodings, etc.

These questions are intended to be suggestive of the possible breadth of inquiries, not as implying deficiencies in the current ORCID technical architecture or services.