In the past, we often discussed how helpful ORCIDs are for our work. An ORCID (Open Researcher and Contributor ID) is a unique personal identifier that scientists can attach to their work. The ORCID ensures that this work is linked to the correct scientist an not to someone else with the same or similar name. We at dblp use ORCIDs to create clean bibliographies. A bibliography should list the work of a single researcher and of course a unique identifier is very helpful here. In this post I will give a short overview on how we handle ORCID and how prevalent it is in DBLP just now. If you do not have an ORCID, consider getting one (for free) at orcid.org. Please make sure that it is attached to your publications whenever possible.

We started experimenting with ORCID in 2016. A more complex integration began in 2017 when we also started to show ORCIDs in bibliographies and individual publications. At the same time we made ORCIDs available with our data releases. We obtain most ORCIDs directly from the publishers together with other publication meta data such as title and author names. ORCID was established in 2012 and many publishers started to attach ORCIDs to their publications only recently (or do not do that at all). But authors can claim such works on their own. This information is provided by ORCID via their annual data dump which we also map to our data set. This means that ORCID has become a common type of data in our collection. Below you see the fraction of signatures in dblp for which an ORCID is known. A signature is a pair of author name and paper. So a paper with five authors has five signatures.
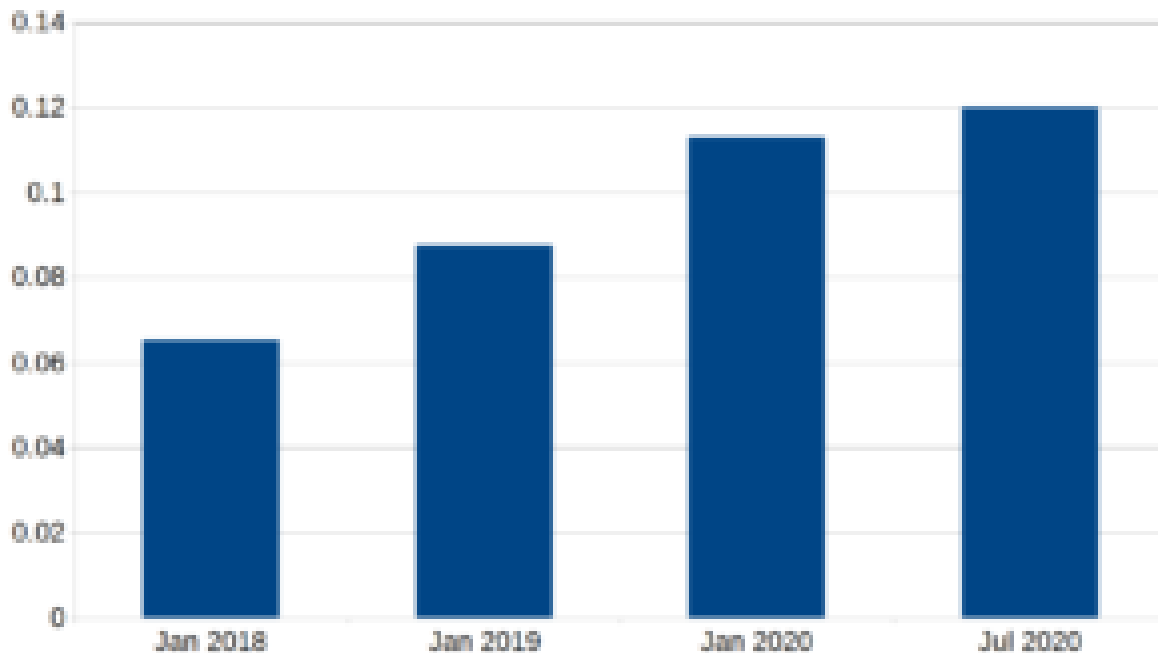
Fraction of signatures with ORCID

An ORCID is now available for 12% of all our signatures and that number is going up. At the moment, we add ORCIDs to dblp in batches. This means that a publication can appear in dblp without any ORCIDs. A few days later they are added. We are working to streamline this process for a faster integration.

(more…)