Q&A
15 December 2021
Finding Small Molecules and their Metabolites in Big Data
Associate Professor Emma Schymanski (University of Luxembourg)

Q1: How are you finding the buy into fair templates? We are very pro making our data fair and I know a lot of people are, but equally it's still a problem that not all data is, so how have you found that?

I'm not aware of anyone having directly used the template yet. They really haven't had the articles out there for very long, but basically with the NORMAN Suspect List Exchange we deal with this data and with the mapping - we have been dealing with it for years, PubChem has equally been dealing with it for years. Let's just say we work on a slight modification thereof. We have header mapping files that do this, so we haven't gone back and retrospectively put all the Suspect List Exchange into this template, but rather we're working with our data and we started putting out these articles because we can see there's a very consistent pattern in the information that you want to use, and if people are providing it with very consistent headers, this is extremely easy for people to digest and value add. And it's really not that hard. What we've found with contributors through experiences, if there's no template out there, then people are not sure what to provide. But if there's a template out there and they're insecure, you can just point them to the template and then they're like, "oh, OK, this is not so hard", and so we're hoping over time it would just help raise awareness and keep growing the data. For the NORMAN Suspect List Exchange, we never imagined how big it would be. We've now got 89 lists and we're about a tenth of PubChem contributors (they have 828 contributors). We have 89 contributors, so for a small environmental initiative, it's growing.