



## **A case for user-generated sensor metadata**

Daniel Nüst

Institute for Geoinformatics, University of Münster, Münster, Germany

Cheap and easy to use sensing technology and new developments in ICT towards a global network of sensors and actuators promise previously unthought of changes for our understanding of the environment. Large professional as well as amateur sensor networks exist, and they are used for specific yet diverse applications across domains such as hydrology, meteorology or early warning systems. However the impact this “abundance of sensors” had so far is somewhat disappointing. There is a gap between (community-driven) sensor networks that could provide very useful data and the users of the data. In our presentation, we argue this is due to a lack of metadata which allows determining the fitness of use of a dataset. Syntactic or semantic interoperability for sensor webs have made great progress and continue to be an active field of research, yet they often are quite complex, which is of course due to the complexity of the problem at hand. But still, we see the most generic information to determine fitness for use is a dataset’s provenance, because it allows users to make up their own minds independently from existing classification schemes for data quality.

In this work we will make the case how curated user-contributed metadata has the potential to improve this situation. This especially applies for scenarios in which an observed property is applicable in different domains, and for set-ups where the understanding about metadata concepts and (meta-)data quality differs between data provider and user. On the one hand a citizen does not understand the ISO provenance metadata. On the other hand a researcher might find issues in publicly accessible time series published by citizens, which the latter might not be aware of or care about. Because users will have to determine fitness for use for each application on their own anyway, we suggest an online collaboration platform for user-generated metadata based on an extremely simplified data model. In the most basic fashion, metadata generated by users can be boiled down to a basic property of the world wide web: many information items, such as news or blog posts, allow users to create comments and rate the content. Therefore we argue to focus a core data model on one text field for a textual comment, one optional numerical field for a rating, and a resolvable identifier for the dataset that is commented on.

We present a conceptual framework that integrates user comments in existing standards and relevant applications of online sensor networks and discuss possible approaches, such as linked data, brokering, or standalone metadata portals. We relate this framework to existing work in user generated content, such as proprietary rating systems on commercial websites, microformats, the GeoViQua User Quality Model, the CHARMe annotations, or W3C Open Annotation. These systems are also explored for commonalities and based on their very useful concepts and ideas; we present an outline for future extensions of the minimal model. Building on this framework we present a concept how a simplistic comment-rating-system can be extended to capture provenance information for spatio-temporal observations in the sensor web, and how this framework can be evaluated.