

Online Information Quality

26 - 29 March 2018 @Snellius

Introduction

The Online Information Quality workshop has been held from March 26th to March 29th.

Overall, 23 participants joined the event, coming from the Netherlands, United Kingdom, Italy, Belgium, and France. The audience comprised both senior researchers and Ph.D. students. The goal of the workshop was to investigate information quality online from diverse angles. Understanding, characterizing, describing, and assessing the quality of information online is a compelling issue. To tackle this vast challenge, we need the contribution of diverse practitioners. Computer Scientists are likely candidates to provide tools and methodologies to assess information online, but we need the contribution also of philosophers, sociologists, management experts, digital humanities scholars, and other researchers in order to tackle this challenge. The quality of information online depends, in fact, on the intentions, incentives, and constraints of who produces and who consumes it. To understand these elements, we need domain-specific contributions.

Workshop structure

The programme has been structured as follows: we had three days of themed presentations and group work, followed by a write-a-thon. The themes of the first three days were: digital humanities & cultural heritage, online news, and theoretical & pragmatic aspects of online information quality. The talks on the theme of digital humanities & cultural heritage touched upon two main themes. On the one hand, cultural heritage institutions need to preserve their authoritativeness in a Web context, and thus they put in place specific mechanisms to do so. On the other hand, academia provides methods and tools to assess the quality of information for digital humanities experts, allowing them to consciously assess it. The talks on journalism and social media took the reader angle and focused on the impact of online news on readers, and on how the perception of such quality is influenced also by how readers are exposed to different sources of news. On the third day, two industrial and two academic talks shed light on how the quality of online information is conceived, determined, and handled in specific sectors, namely, social media management, scientific publishing, and management research. Part of each of the three days was dedicated to group work, when the participant addressed the following questions in the context of the theme of the day: (1) What are the minimum requirements for quality? (2) When do you need to engage source/tool criticism? (3) How is reputation/criticism built?

Workshop Outcomes

One important (high level) outcome of the workshop is the recognition that information quality is a multi-dimensional problem. The specific lesson related to this is that we need to recognize this multi-dimensional nature to ensure we do not address only a single dimension or simply the wrong dimension, understand and map its many dimensions, and use this understanding to identify underdeveloped, poorly theorized, or easily ignored aspects of information quality.

As a first approximation, we can add some structure to the problem by distinguishing between:

- the different reasons why IQ is valuable or desirable,
- the different reasons why IQ can be hard to obtain, including:
 - traditional challenges: messy information, the presence of noise, the presence of intentional misinformation and disinformation, accumulation of errors, etc.
 - challenges due to the heterogeneity of sources, actors, standards, expectations, etc.
 - practical challenges: scale, cost, compromises, etc.
 - socio-epistemic and ethical considerations: trust, transparency, accountability,
- the difference between characterizing information quality (definitions, standards, criteria) and recognizing information quality (applying standards, criteria, ...).

These findings, together with a deeper report, will be described in an article to be submitted to the ACM Journal on Information and Data Quality. We envision the possibility to organize a journal special issue on these themes, as well as to organize at least one follow-up workshop, tackling online information quality in additional domains.

Patrick Allo (Vrije Universiteit Brussel, Belgium)

Lora Aroyo (Vrije Universiteit Amsterdam, The Netherlands)

Kaspar Beelen (Universiteit van Amsterdam, The Netherlands)

Davide Ceolin (Centrum Wiskunde en Informatica, The Netherlands)

Vladi Finotto (Ca' Foscari University, Italy)