

Machines are People Too

Paul Groth

Elsevier Labs, Elsevier Inc., USA

Abstract. The theory and practice of digital libraries provides a long history of thought around how to manage knowledge ranging from collection development, to cataloging and resource description. These tools were all designed to make knowledge findable and accessible to people. Even technical progress in information retrieval and question answering are all targeted to helping answer a human's information need.

However, increasingly demand is for data. Data that is needed not for people's consumption but to drive machines. As an example of this demand, there has been explosive growth in job openings for Data Engineers – professionals who prepare data for machine consumption. In this talk, I overview the information needs of machine intelligence and ask the question: Are our knowledge management techniques applicable for serving this new consumer?

Paul Groth is Disruptive Technology Director at Elsevier Labs. He holds a Ph.D. in Computer Science from the University of Southampton (2007) and has done research at the University of Southern California and the Vrije Universiteit Amsterdam. His research focuses on dealing with large amounts of diverse contextualized knowledge with a particular focus on the web and science applications. This includes research in data provenance, data science, data integration and knowledge sharing. He leads architecture development for the Open PHACTS drug discovery data integration platform. Paul was co-chair of the W3C Provenance Working Group that created a standard for provenance interchange. He is co-author of "Provenance: An Introduction to PROV" and "The Semantic Web Primer: 3rd Edition" as well as numerous academic articles. He blogs at <http://thinklinks.wordpress.com>. You can find him on twitter: @pgroth.