An Introduction to Hybrid Human-Machine Information Systems

Gianluca Demartini
University of Queensland
g.demartini@uq.edu.au

Djellel Eddine Difallah
Center for Data Science
New York University
djellel@nyu.edu

Ujwal Gadiraju
L3S Research Center
Leibniz Universität Hannover
gadiraju@L3S.de

Michele Catasta
Stanford University
pirroh@cs.stanford.edu

now
the essence of knowledge
Boston — Delft

Foundations and Trends® in Web Science

Published, sold and distributed by:
now Publishers Inc.
PO Box 1024
Hanover, MA 02339
United States
Tel. +1-781-985-4510
www.nowpublishers.com
sales@nowpublishers.com

Outside North America:
now Publishers Inc.
PO Box 179
2600 AD Delft
The Netherlands
Tel. +31-6-51115274

The preferred citation for this publication is

G. Demartini, D. E. Difallah, U. Gadiraju and M. Catasta. *An Introduction to Hybrid Human-Machine Information Systems*. Foundations and Trends® in Web Science, vol. 7, no. 1, pp. 1–87, 2016.

This Foundations and Trends® issue was typeset in LaTeX using a class file designed by Neal Parikh. Printed on acid-free paper.

ISBN: 978-1-68083-374-4

Foundations and Trends® in Web Science

Volume 7, Issue 1, 2016

Editorial Scope

Topics

Foundations and Trends® in Web Science publishes survey and tutorial articles in the following topics:

- Agents and the semantic web
- Collective intelligence
- Content management
- Databases on the web
- Data mining
- Democracy and the web
- Dependability
- Economics of information and the web
- E-crime
- E-government
- Emergent behaviour
- Ethics
- Hypertext/Hypermedia
- Identity
- Languages on the web
- Memories for life
- Mobile/Pervasive
- Network infrastructures
- Performance
- Privacy
- Scalability
- Security
- Semantic web
- Social networking
- Standards
- The law and the web
- The web as an educational tool
- The web in the developing world
- Trust and provenance
- Universal usability
- User interfaces
- Virtual reality
- Web art
- Web governance
- Search
- Web services

Information for Librarians

Foundations and Trends® in Web Science, 2016, Volume 7, 4 issues. ISSN paper version 1555-077X. ISSN online version 1555-0788. Also available as a combined paper and online subscription.

Foundations and Trends® in Web Science
Vol. 7, No. 1 (2016) 1–87

DOI: 10.1561/1800000025

An Introduction to Hybrid Human-Machine Information Systems

Gianluca Demartini
University of Queensland
g.demartini@uq.edu.au

Djellel Eddine Difallah
Center for Data Science
New York University
djellel@nyu.edu

Ujwal Gadiraju
L3S Research Center
Leibniz Universität Hannover
gadiraju@L3S.de

Michele Catasta
Stanford University
pirroh@cs.stanford.edu

Contents

Abstract

Hybrid Human-Machine Information Systems leverage novel architectures that make systematic use of Human Computation by means of crowdsourcing. These architectures are capable of scaling over large amounts of data and simultaneously maintain high-quality data processing levels by introducing humans into the loop. Such hybrid systems have been developed to tackle a variety of problems and come with inter-disciplinary challenges. They need to deal with the full spectrum of challenges from the social science standpoint, such as understanding crowd workers behavior and motivations when performing tasks. These systems also need to overcome highly technical challenges like constraint optimization and resource allocation based on limited budgets and deadlines to be met.

In this paper, we introduce the area of Human Computation and present an overview of different applications for which Hybrid Human-Machine Information Systems have already been used in the realms of data management, information retrieval, natural language processing, semantic web, machine learning, and multimedia to better solve existing problems. Finally, we discuss current research directions, opportunities for the future development of such systems and their application in practice.

G. Demartini, D. E. Difallah, U. Gadiraju and M. Catasta. *An Introduction to Hybrid Human-Machine Information Systems*. Foundations and Trends® in Web Science, vol. 7, no. 1, pp. 1–87, 2016.
DOI: 10.1561/1800000025.

1

Crowdsourcing and Human Computation

1.1 Motivation

Crowdsourcing is a broad term that encompasses different approaches to collect ideas, opinions, or data from a group of anonymous individuals, typically on-line [Howe, 2006]. Crowdsourcing for Human Computation is commonly interpreted as the manual execution at scale of micro-tasks for data processing or analysis: The idea is that some computational tasks are still easier for humans than for machines and algorithms to perform. Examples of such tasks include image understanding, detecting sarcasm in text, transcribing audio files to text, and others. A common property of crowdsourcing is that the contributors are considered anonymous other than having a username and some profiling information related to them (e.g., Wikipedia editors). This aspect triggers possible issues of trust and content quality (e.g., vandalism in Wikipedia [Potthast et al., 2008]). On the other hand, it opens the doors to large numbers of individuals who can contribute to the crowdsourcing campaign. The popularity of Human Computation approaches is evident with the growth of commercial paid micro-task crowdsourcing platforms[1] like Amazon Mechanical Turk[2] and CrowdFlower[3] that enable the creation of hybrid human-machine information systems harnessing the wisdom of the crowd at scale. Such systems are designed to leverage both the scalability of computational machines over large amounts of data as well as the power of human intelligence by building human-in-the-loop systems that can get input from manual processing of data. There are numerous applications of Human Computation to real-world problems, both in the industry and in academia. Popular examples of commercial applications are the annotation of online videos and the transcription of audio content. Examples of academic research include systems for classifying the sentiment of social media content and large-scale execution of online surveys.

In this work, we introduce the cross-domain area of hybrid human-machine systems. We provide an overview of the systems that have proposed and evaluated by different research communities relevant to Web Science. We also highlight current open research challenges (from the social to the technical ones) that need to be tackled to make such systems reliable and dependable.

1.2 Crowdsourcing Platforms

Howe [2006] defined crowdsourcing as the approach of tapping into human intelligence at scale by accessing crowds of people online. Crowdsourcing is a very broad term that includes leveraging human intelligence to solve complex problems like innovation challenges[4] [Travis, 2008], to support scientific discoveries[5] [Lintott et al., 2008], up to simple micro-tasks platforms. For example, InnoCentive brings together tens of thousand scientists that collaboratively aim to solve difficult research problems. Problems are typically provided by large organizations like, for example, national space agencies or pharmaceutical com-

[1] In this work we focus on systems built using micro-task crowdsourcing platforms, and we will thus use the term 'crowdsourcing platform' to refer to micro-task platforms only.

[2] `http://mturk.com`

[3] `http://crowdflower.com`

[4] `http://innocentive.com/`

[5] `http://galaxyzoo.org/`

panies. This platform allows for discussion and collaborative problem solving of these grand challenges. Another example of crowdsourcing is GalaxyZoo which is a large citizen science project where volunteers contribute by annotating images or other scientific data artifacts with the purpose of supporting scientific discovery without the need to be an expert in the field. For example, given a space image was taken by a telescope, members of the crowd are asked to categorize it in one of few possible galaxy types depicted using stylized icons.

The individual tasks available on a micro-task crowdsourcing platform are typically called *Human Intelligence Tasks* (HITs) and are completed by individuals in the crowd also known as *workers*. On the other hand, HITs are published on these platforms by so-called *requesters*, who, in a paid crowdsourcing setting, would attach a monetary reward to be assigned to workers who complete the HIT. Such platforms have been gaining popularity over time and are used for both commercial products as well as for academic research [Difallah et al., 2015]. On the crowd worker side, because tasks are paid the same amount independently on where the worker is physically located, such platforms are used for a number of different reasons including both as a leisure activity (in these cases workers would look for tasks that are interesting and fun) as well as a means to have a full-time job (in these cases workers would look for highly rewarded tasks) [Kuek et al., 2015].

1.3 The Amazon Mechanical Turk Marketplace

Amazon Mechanical Turk (MTurk) is arguably one of the oldest and currently the most popular micro-task crowdsourcing platform. It has a continuous flow of workers and requesters. It provides a programmatic Application Programming Interfaces (APIs) as well as a Web interface for requesters to design and deploy online tasks. Its activity logs are available to the public[6] [Ipeirotis, 2010a] and were used to perform an analysis tracing its evolution [Difallah et al., 2015]. This analysis has shown that the amount of tasks and reward available on the MTurk platform as well as the number of active requesters have been increas-

[6] `http://www.mturk-tracker.com`

ing over time. This is a sign of increased interest in the use micro-task crowdsourcing. Looking at the most active MTurk requesters, it is possible to observe a mix of academic users as well as industrial organizations. Most common tasks types include the execution of surveys (e.g., for social science studies) and commercial applications like, for example, audio transcription and image annotation. Gadiraju et al. [2014] presented a goal-oriented taxonomy of microtasks which identified the following six main categories of task type depending on the goal of the requesters – content access, content creation, information finding, interpretation and analysis, surveys, and verification and validation.

Demographic studies by [Ipeirotis, 2010b, Ross et al., 2010] have shown that the vast majority of crowd workers on MTurk is split between the United States and India. On average, MTurk workers have high education levels and tend to be younger in India than in the US. Requesters on MTurk have been initially required to be based in the US (for financial reasons) but are now allowed from a small but growing number of mainly English-speaking countries.

MTurk adopts a *pull* crowdsourcing methodology, where all the published tasks are publicly presented to workers on a search-based dashboard. The workers can then pick their preferred tasks on a first-come-first-served basis.

From a requester perspective, the pull crowdsourcing approach has several advantages including simplicity and minimization of task completion times, since any available worker from the crowd can pick and perform any HIT, provided that they meet some pre-requisites set by the requester. From a worker perspective, it creates competition among requesters, and potentially leads to high HIT standards, both in terms of interface design, quality, and pricing.

On the other hand, pull crowdsourcing limits the possibilities of the platform to offer any form of service guarantees to its customers (i.e., the requesters). For example, this mechanism cannot guarantee priority to a requester who has a deadline, and often the only effective lever consists in increasing the unit reward of the HITs to attract more workers [Alonso and Baeza-Yates, 2011]. It also cannot guarantee that the worker who performs the task is the best fit, as more knowledgeable workers might be available within the crowd, but are unable to pick the task on time.

1.4 A Definition of Hybrid Human-Machine Information Systems

Modern crowdsourcing platforms offer programmatic APIs to post HITs, monitor their progress, collect the results and distribute the rewards to participating workers. Hence, the idea of combining Human Computation and computers to produce a new breed of hybrid human-machine algorithms found an opportunity to concretize. Not only can the crowd be invoked programmatically, using a declarative language, but this very process can also be parametrized, monitored and embedded in long-running jobs [Law and Ahn, 2011].

A direct application of this idea goes naturally with the class of machine-learning algorithms that produce their results along with a confidence score. A generic hybrid scheme consists in falling back to the crowd to increase the precision of the results whenever the confidence of the generated solution falls below a predefined threshold. Another application is in active learning, where a classification algorithm would repeatedly collect training labels from the crowd – as opposed to a limited number of human operators [Mozafari et al., 2014]. Likewise, we refer to the class of information systems that would involve the crowd at some point in their execution as *Hybrid Human-Machine Information Systems.*

1.5 Challenges of Hybrid Human-Machine Systems

While crowdsourcing platforms enable the design and development of novel information systems that benefit both from the scalability of machine processing as well as from the power of human intelligence, there are some challenges that need to be tackled to make sure that such systems are efficient and effective [Demartini, 2015].

The first challenge is how to best combine human intelligence and machine processing power in the most efficient way knowing that humans are naturally slower but more capable than machine-based algorithms. Two main approaches to combine human intelligence with machine processing have been used so far. The first approach is using crowdsourcing as a data *pre-processing* step. An example of this is the creation of large-scale manually labeled datasets to train machine learning algorithms [Mozafari et al., 2014]. The second approach is crowdsourcing to *post-process* machine-based algorithmic results. An example of this is the manual quality check and filtering of a ranked list of results for a search query [Teevan et al., 2014].

Another common property of such hybrid systems is the use of monetary incentives. Rewarding workers for their contributions help to scale easily the size of the crowd. This, however, brings in challenges of trust and data quality. The use of financial incentives creates another challenge which is the presence of malicious worker who will perform with low-quality to quickly collect the reward attached to the task [Gadiraju et al., 2015]. This has a direct implication on the effectiveness of the hybrid human-machine system that relies on quality data from the crowdsourcing platform. While the use of financial incentives introduces certain challenges, lessons learned from other types of crowdsourcing (e.g., volunteer crowdsourcing like Wikipedia or gamification approaches used in games with a purpose) can be adapted and applied to paid crowdsourcing platforms too, for example, retain workers and foster quality work.

The third challenge of the use of crowdsourcing in hybrid human-machine systems is the latency introduced by crowd-based data processing. A popular approach in such hybrid systems is to perform *batch data processing* rather than executing real-time jobs. When such hybrid systems make humans and machines work in combination, the obvious efficiency bottleneck lies on the crowd side. This is due to the intrinsic latency introduced by the use of humans completing tasks in crowdsourcing applications which limits the potential for real-time responses. Moreover, it is very difficult to predict the completion time of a batch of crowdsourcing tasks as different batches are competing for workers attention on the platform.

On a temporal perspective, after a focus on developing hybrid systems for specific problems across disciplines, more recently, research attention has moved on solving core crowdsourcing problems (e.g., incentives, retention, quality assurance, etc.) rather than building new systems and applications [Demartini, 2015].

1.6 Opportunities of Hybrid Human-Machine Systems

In the remainder of this book, we present an overview of the different approaches to building hybrid information systems adopted by different communities in Web Science, thus highlighting the opportunities of such systems and how different communities have dealt with the key challenges outlined in the previous section.

We start with discussing work in the database area where Human Computation has been applied to problems like query interpretation and data integration (Chapter 2). Hybrid human-machine systems like CrowdDB [Franklin et al., 2011] have been proposed to address missing data problems and provide subjective ordering capabilities to databases.

We discuss work carried out by the information retrieval community that has used crowdsourcing as a methodology to create evaluation collection as well as part of a search system either to interpret search queries or to answer to complex information needs (Chapter 3). In this domain, the opportunity that hybrid systems can leverage is the power of human intelligence to understand search queries and to identify pieces of relevant content to be presented back to users having an information need.

We then present work in the natural language processing area that looked into the use of crowdsourcing to develop hybrid systems for information extraction tasks like, for example, named entity recognition and entity linking (Chapter 4). Again, the power of human intelligence here is harnessed for its ability to understand natural language and its idiosyncrasies like, for example, the use of sarcasm which is an open challenge for purely machine-based sentiment classification methods.

We then discuss hybrid systems developed by the semantic web community for schema matching and knowledge acquisition (Chapter 5). In this domain, hybrid systems leverage the natural ability of humans for semantic understanding. Thus, they can help machine-based algorithms in defining concept relations and in merging schemas by manually interpreting data semantics.

We present work in the area of machine learning that is using crowdsourcing to collect training data for supervised models either in batch or within an active learning setting (Chapter 6). In this case, the crowd-based component can be used to provide examples to machine-based models that can then scale the classification to large datasets.

We present work from the multimedia community which developed hybrid systems for processing different content types (e.g., images, audio, video), and the ubiquitous computing community at large (Chapter 7). Tasks like audio transcription and image labeling are among the most popular in crowdsourcing platforms. The opportunity for hybrid systems applied to this domain is to achieve human-level understanding in automatic multimedia content processing.

Finally, in Chapter 8, we present the common techniques used across communities in Web Science and highlight the lessons learned and the open research questions that need to be addressed to produce better hybrid human-machine information systems.

References

Maribel Acosta, Amrapali Zaveri, Elena Simperl, Dimitris Kontokostas, Sören Auer, and Jens Lehmann. Crowdsourcing linked data quality assessment. In *The Semantic Web - ISWC 2013 - 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II*, pages 260–276, 2013. URL http://dx.doi.org/10.1007/978-3-642-41338-4_17.

Eugene Agichtein, Eric Brill, and Susan Dumais. Improving web search ranking by incorporating user behavior information. In *Proceedings of the 29th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 19–26. ACM, 2006.

Salman Ahmad, Alexis Battle, Zahan Malkani, and Sepander Kamvar. The jabberwocky programming environment for structured social computing. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 53–64. ACM, 2011.

Omar Alonso. Implementing crowdsourcing-based relevance experimentation: An industrial perspective. *Information Retrieval*, 16(2):101–120, 2013.

Omar Alonso and Ricardo A. Baeza-Yates. Design and Implementation of Relevance Assessments Using Crowdsourcing. In *ECIR*, pages 153–164, 2011.

Omar Alonso and Stefano Mizzaro. Relevance criteria for e-commerce: A crowdsourcing-based experimental analysis. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 760–761, New York, NY, USA, 2009. ACM. URL http://doi.acm.org/10.1145/1571941.1572115.

Omar Alonso and Stefano Mizzaro. Using crowdsourcing for TREC relevance assessment. *Inf. Process. Manage.*, 48(6):1053–1066, 2012. URL http://dx.doi.org/10.1016/j.ipm.2012.01.004.

Vamshi Ambati, Stephan Vogel, and Jaime Carbonell. Active learning-based elicitation for semi-supervised word alignment. In *Proceedings of the ACL 2010 Conference Short Papers*, ACLShort '10, pages 365–370, Stroudsburg, PA, USA, 2010a. Association for Computational Linguistics. URL http://dl.acm.org/citation.cfm?id=1858842.1858909.

Vamshi Ambati, Stephan Vogel, and Jaime G. Carbonell. Active learning and crowd-sourcing for machine translation. In *Proceedings of the International Conference on Language Resources and Evaluation, LREC 2010, 17-23 May 2010, Valletta, Malta*, 2010b. URL http://www.lrec-conf.org/proceedings/lrec2010/summaries/244.html.

Nate Anderson. The youtube effect: Http traffic now eclipses p2p. *Ars Technica*, 19, 2007.

Jaime Arguello, Fernando Diaz, Jamie Callan, and Jean-Francois Crespo. Sources of evidence for vertical selection. In *Proceedings of the 32Nd International ACM SIGIR Conference on Research and Development in Information Retrieval*, SIGIR '09, pages 315–322, New York, NY, USA, 2009. ACM. URL http://doi.acm.org/10.1145/1571941.1571997.

Lora Aroyo and Chris Welty. The three sides of crowdtruth. *Journal of Human Computation*, 1:31–34, 2014.

Yoram Bachrach, Thore Graepel, Tom Minka, and John Guiver. How to grade a test without knowing the answers—a bayesian graphical model for adaptive crowdsourcing and aptitude testing. *Proceedings of the 29th International Conference on Machine Learning (ICML 2012)*, 2012.

Albert Bandura. *Self-efficacy*. Wiley Online Library, 1994.

Tim Berners-Lee, James Hendler, and Ora Lassila. The semantic web. *Scientific american*, 284(5):28–37.

Michael S. Bernstein, Jaime Teevan, Susan Dumais, Daniel Liebling, and Eric Horvitz. Direct answers for search queries in the long tail. In *CHI '12*, pages 237–246. ACM, 2012. URL http://doi.acm.org/10.1145/2207676.2207710.

Michael S Bernstein, Greg Little, Robert C Miller, Björn Hartmann, Mark S Ackerman, David R Karger, David Crowell, and Katrina Panovich. Soylent: a word processor with a crowd inside. *Communications of the ACM*, 58(8):85–94, 2015.

Mikhail Bilenko and Raymond J. Mooney. Adaptive duplicate detection using learnable string similarity measures. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, KDD '03, pages 39–48, New York, NY, USA, 2003. ACM. URL `http://doi.acm.org/10.1145/956750.956759`.

Roi Blanco, Harry Halpin, Daniel M Herzig, Peter Mika, Jeffrey Pound, Henry S Thompson, and Thanh Tran Duc. Repeatable and reliable search system evaluation using crowdsourcing. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information - SIGIR '11*, page 923, 2011. URL `http://portal.acm.org/citation.cfm?doid=2009916.2010039`.

Kalina Bontcheva, Ian Roberts, Leon Derczynski, and Dominic Rout. The GATE Crowdsourcing Plugin: Crowdsourcing Annotated Corpora Made Easy. In *Proceedings of Demonstrations at the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 97–100. Association for Computational Linguistics, 2014.

Ria Mae Borromeo and Motomichi Toyama. Automatic vs. crowdsourced sentiment analysis. In *Proceedings of the 19th International Database Engineering & Applications Symposium*, IDEAS '15, pages 90–95, New York, NY, USA, 2014. ACM. URL `http://doi.acm.org/10.1145/2790755.2790761`.

Alessandro Bozzon, Ilio Catallo, Eleonora Ciceri, Piero Fraternali, Davide Martinenghi, Marco Tagliasacchi, and M Tagliasacchi. A framework for crowdsourced multimedia processing and querying. *CrowdSearch*, 842:42–47, 2012.

Anthony Brew, Derek Greene, and Pádraig Cunningham. Using crowdsourcing and active learning to track sentiment in online media. In *Proceedings of the 2010 Conference on ECAI 2010: 19th European Conference on Artificial Intelligence*, pages 145–150, Amsterdam, The Netherlands, The Netherlands, 2010. IOS Press. URL `http://dl.acm.org/citation.cfm?id=1860967.1860997`.

Robin Brewer, Meredith Ringel Morris, and Anne Marie Piper. "why would anybody do this?": Understanding older adults' motivations and challenges in crowd work. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems, San Jose, CA, USA, May 7-12, 2016*, pages 2246–2257, 2016. URL `http://doi.acm.org/10.1145/2858036.2858198`.

R. Bunescu and M. Pasca. Using encyclopedic knowledge for named entity disambiguation. In *Proceedings of EACL*, volume 6, 2006.

Jeffrey A Burke, Deborah Estrin, Mark Hansen, Andrew Parker, Nithya Ramanathan, Sasank Reddy, and Mani B Srivastava. Participatory sensing. *Center for Embedded Network Sensing*, 2006.

Michele Catasta, Alberto Tonon, Djellel Eddine Difallah, Gianluca Demartini, Karl Aberer, and Philippe Cudré-Mauroux. TransactiveDB: Tapping into Collective Human Memories. *Proceedings of the VLDB Endowment*, 7(14), 2014.

Joseph Chee Chang, Aniket Kittur, and Nathan Hahn. Alloy: Clustering with crowds and computation. In *Proceedings of the 2016 CHI Conference on Human Factors in Computing Systems*, pages 3180–3191. ACM, 2016.

Georgios Chatzimilioudis, Andreas Konstantinidis, Christos Laoudias, and Demetrios Zeinalipour-Yazti. Crowdsourcing with smartphones. *IEEE Internet Computing*, 16(5):36–44, 2012.

Alessandro Checco, Kevin Roitero, Eddy Maddalena, Stefano Mizzaro, and Gianluca Demartini. Let's agree to disagree: Fixing agreement measures for crowdsourcing. In *The 5th AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2017)*, HCOMP '17, 2017.

Kuan-Ta Chen, Chen-Chi Wu, Yu-Chun Chang, and Chin-Laung Lei. A crowdsourceable qoe evaluation framework for multimedia content. In *Proceedings of the 17th ACM international conference on Multimedia*, pages 491–500. ACM, 2009.

Peter Christen. A survey of indexing techniques for scalable record linkage and deduplication. *IEEE Trans. on Knowl. and Data Eng.*, 24(9):1537–1555, September 2012. URL `http://dx.doi.org/10.1109/TKDE.2011.127`.

Joana Costa, Catarina Silva, Mário Antunes, and Bernardete Ribeiro. On using crowdsourcing and active learning to improve classification performance. In *Intelligent Systems Design and Applications (ISDA), 2011 11th International Conference on*, pages 469–474. IEEE, 2011.

S. Cucerzan. Large-scale named entity disambiguation based on Wikipedia data. In *Proceedings of EMNLP-CoNLL*, volume 2007, pages 708–716, 2007.

Mohamed Yehia Dahab, Hesham A Hassan, and Ahmed Rafea. Textontoex: Automatic ontology construction from natural english text. *Expert Systems with Applications*, 34(2):1474–1480, 2008.

Susan B. Davidson, Sanjeev Khanna, Tova Milo, and Sudeepa Roy. Using the Crowd for Top-k and Group-by Queries. In *Proceedings of the 16th International Conference on Database Theory*, ICDT '13, pages 225–236, New York, NY, USA, 2013. ACM. URL `http://doi.acm.org/10.1145/2448496.2448524`.

Alexander Philip Dawid and Allan M Skene. Maximum likelihood estimation of observer error-rates using the em algorithm. *Applied statistics*, pages 20–28, 1979.

Adriel Dean-Hall, Charles L Clarke, Jaap Kamps, Paul Thomas, Nicole Simone, and Ellen Voorhees. Overview of the trec 2013 contextual suggestion track. In *NIST Special Publication 500-302: The Twenty-Second Text REtrieval Conference Proceedings (TREC 2013)*, 2013.

Gianluca Demartini. Hybrid human–machine information systems: Challenges and opportunities. *Computer Networks*, 90:5–13, 2015.

Gianluca Demartini, Djellel Eddine Difallah, and Philippe Cudré-Mauroux. Zencrowd: leveraging probabilistic reasoning and crowdsourcing techniques for large-scale entity linking. In *Proceedings of the 21st international conference on World Wide Web*, pages 469–478. ACM, 2012.

Gianluca Demartini, Djellel Eddine Difallah, and Philippe Cudré-Mauroux. Large-scale linked data integration using probabilistic reasoning and crowdsourcing. *The VLDB Journal*, 22(5):665–687, 2013a.

Gianluca Demartini, Beth Trushkowsky, Tim Kraska, and Michael J. Franklin. CrowdQ: Crowdsourced Query Understanding. In *CIDR 2013, Sixth Biennial Conference on Innovative Data Systems Research, Asilomar, CA, USA, January 6-9, 2013, Online Proceedings*, 2013b. URL `http://cidrdb.org/cidr2013/Papers/CIDR13_Paper137.pdf`.

Leon Derczynski, Kalina Bontcheva, and Ian Roberts. Broad twitter corpus: A diverse named entity recognition resource. In *COLING 2016, 26th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, December 11-16, 2016, Osaka, Japan*, pages 1169–1179, 2016. URL `http://aclweb.org/anthology/C/C16/C16-1111.pdf`.

Ernesto Diaz-Aviles and Ricardo Kawase. Exploiting twitter as a social channel for human computation. In *CrowdSearch*, pages 15–19, 2012.

Djellel Eddine Difallah, Gianluca Demartini, and Philippe Cudré-Mauroux. Pick-a-crowd: tell me what you like, and i'll tell you what to do. In *Proceedings of the 22nd international conference on World Wide Web*, pages 367–374. International World Wide Web Conferences Steering Committee, 2013.

Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, and Philippe Cudré-Mauroux. Scaling-up the crowd: Micro-task pricing schemes for worker retention and latency improvement. In *Second AAAI Conference on Human Computation and Crowdsourcing difallah-scaleup. pdf Google Scholar BibTex*, 2014.

Djellel Eddine Difallah, Michele Catasta, Gianluca Demartini, Panagiotis G. Ipeirotis, and Philippe Cudré-Mauroux. The dynamics of micro-task crowdsourcing: The case of amazon mturk. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, pages 238–247, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee. URL `https://doi.org/10.1145/2736277.2741685`.

Djellel Eddine Difallah, Gianluca Demartini, and Philippe Cudré-Mauroux. Scheduling human intelligence tasks in multi-tenant crowd-powered systems. In *Proceedings of the 25th International Conference on World Wide Web*, WWW '16, pages 855–865, Republic and Canton of Geneva, Switzerland, 2016. International World Wide Web Conferences Steering Committee. URL `https://doi.org/10.1145/2872427.2883030`.

Guoru Ding, Jinlong Wang, Qihui Wu, Linyuan Zhang, Yulong Zou, Yu-Dong Yao, and Yingying Chen. Robust spectrum sensing with crowd sensors. *IEEE Transactions on Communications*, 62(9):3129–3143, 2014.

X. Dong, A. Halevy, and J. Madhavan. Reference reconciliation in complex information spaces. In *Proceedings of the 2005 ACM SIGMOD international conference on Management of data*, pages 85–96. ACM, 2005.

Siamak Faradani, Björn Hartmann, and Panagiotis G Ipeirotis. What's the right price? pricing tasks for finishing on time. In *Human Computation*, 2011.

Oluwaseyi Feyisetan, Markus Luczak-Rösch, Elena Simperl, Ramine Tinati, and Nigel Shadbolt. Towards hybrid NER: A study of content and crowdsourcing-related performance factors. In *The Semantic Web. Latest Advances and New Domains - 12th European Semantic Web Conference, ESWC 2015, Portoroz, Slovenia, May 31 - June 4, 2015. Proceedings*, pages 525–540, 2015. URL `http://dx.doi.org/10.1007/978-3-319-18818-8_32`.

Michael J. Franklin, Donald Kossmann, Tim Kraska, Sukriti Ramesh, and Reynold Xin. CrowdDB: answering queries with crowdsourcing. In *Proceedings of the 2011 ACM SIGMOD International Conference on Management of data*, SIGMOD '11, pages 61–72, New York, NY, USA, 2011. ACM. URL `http://doi.acm.org/10.1145/1989323.1989331`.

Ujwal Gadiraju and Stefan Dietze. Improving learning through achievement priming in crowdsourced information finding microtasks. In *Proceedings of the Seventh International Learning Analytics & Knowledge Conference*, pages 105–114. ACM, 2017.

Ujwal Gadiraju, Ricardo Kawase, and Stefan Dietze. A taxonomy of microtasks on the web. In *Proceedings of the 25th ACM conference on Hypertext and social media*, pages 218–223. ACM, 2014.

Ujwal Gadiraju, Ricardo Kawase, Stefan Dietze, and Gianluca Demartini. Understanding malicious behavior in crowdsourcing platforms: The case of online surveys. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems, CHI 2015, Seoul, Republic of Korea, April 18-23, 2015*, pages 1631–1640, 2015.

Ujwal Gadiraju, Alessandro Checco, Neha Gupta, and Gianluca Demartini. Modus operandi of crowd workers: The invisible role of microtask work environments. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies*, 1(3):49, 2017a.

Ujwal Gadiraju, Besnik Fetahu, Ricardo Kawase, Patrick Siehndel, and Stefan Dietze. Using worker self-assessments for competence-based pre-selection in crowdsourcing microtasks. *ACM Transactions on Computer-Human Interaction (TOCHI)*, 24(4):30, 2017b.

Ujwal Gadiraju, Jie Yang, and Alessandro Bozzon. Clarity is a worthwhile quality: On the role of task clarity in microtask crowdsourcing. In *Proceedings of the 28th ACM Conference on Hypertext and Social Media, HT 2017, Prague, Czech Republic, July 4-7, 2017*, pages 5–14. ACM, 2017c.

Mihai Georgescu, Dang Duc Pham, Claudiu S Firan, Ujwal Gadiraju, and Wolfgang Nejdl. When in doubt ask the crowd: Employing crowdsourcing for active learning. In *Proceedings of the 4th International Conference on Web Intelligence, Mining and Semantics (WIMS14)*, page 12. ACM, 2014.

Stephen Guo, Aditya Parameswaran, and Hector Garcia-Molina. So who won?: dynamic max discovery with the crowd. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pages 385–396. ACM, 2012.

Neha Gupta, David Martin, Benjamin V Hanrahan, and Jacki O'Neill. Turklife in india. In *Proceedings of the 18th International Conference on Supporting Group Work*, pages 1–11. ACM, 2014.

Martin Halvey and Robert Villa. Evaluating the effort involved in relevance assessments for images. In *The 37th International ACM SIGIR Conference on Research and Development in Information Retrieval, SIGIR '14, Gold Coast , QLD, Australia - July 06 - 11, 2014*, pages 887–890, 2014. URL http://doi.acm.org/10.1145/2600428.2609466.

Xianpei Han and Jun Zhao. Named entity disambiguation by leveraging wikipedia semantic knowledge. In *Proceeding of the 18th ACM conference on Information and knowledge management*, CIKM '09, pages 215–224, New York, NY, USA, 2009. ACM. URL http://doi.acm.org/10.1145/1645953.1645983.

Tenshi Hara, Thomas Springer, Gerd Bombach, and Alexander Schill. Decentralised approach for a reusable crowdsourcing platform utilising standard web servers. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 1063–1074. ACM, 2013.

Bernhard Haslhofer, Elaheh Momeni, Manuel Gay, and Rainer Simon. Augmenting europeana content with linked data resources. In *Proceedings of the 6th International Conference on Semantic Systems*, I-SEMANTICS '10, pages 40:1–40:3, New York, NY, USA, 2010. ACM. URL http://doi.acm.org/10.1145/1839707.1839757.

Joseph M Hellerstein and David L Tennenhouse. Searching for jim gray: a technical overview. *Communications of the ACM*, 54(7):77–87, 2011.

Mehdi Hosseini, Ingemar J. Cox, Natasa Milic-Frayling, Gabriella Kazai, and Vishwa Vinay. On aggregating labels from multiple crowd workers to infer relevance of documents. In *Advances in Information Retrieval - 34th European Conference on IR Research, ECIR 2012, Barcelona, Spain, April 1-5, 2012. Proceedings*, pages 182–194, 2012. URL http://dx.doi.org/10.1007/978-3-642-28997-2_16.

Tobias Hoßfeld, Michael Seufert, Matthias Hirth, Thomas Zinner, Phuoc Tran-Gia, and Raimund Schatz. Quantification of youtube qoe via crowdsourcing. In *Multimedia (ISM), 2011 IEEE International Symposium on*, pages 494–499. IEEE, 2011.

Jeff Howe. The rise of crowdsourcing. *Wired magazine*, 14(6):1–4, 2006.

Nguyen Quoc Viet Hung, Nguyen Thanh Tam, Ngoc Tran Lam, and Karl Aberer. An evaluation of aggregation techniques in crowdsourcing. In *WISE (2)*, pages 1–15, 2013.

Panagiotis G Ipeirotis. Analyzing the amazon mechanical turk marketplace. *XRDS: Crossroads, The ACM Magazine for Students*, 17(2):16–21, 2010a.

Panagiotis G Ipeirotis. Demographics of mechanical turk. 2010b.

Lilly C. Irani and M. Six Silberman. Turkopticon: Interrupting Worker Invisibility in Amazon Mechanical Turk. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 611–620, New York, NY, USA, 2013. ACM. URL http://doi.acm.org/10.1145/2470654.2470742.

Alejandro Jaimes, Nicu Sebe, and Daniel Gatica-Perez. Human-centered computing: a multimedia perspective. In *Proceedings of the 14th ACM international conference on Multimedia*, pages 855–864. ACM, 2006.

Puneet Jain, Justin Manweiler, Arup Acharya, and Kirk Beaty. Focus: clustering crowdsourced videos by line-of-sight. In *Proceedings of the 11th ACM Conference on Embedded Networked Sensor Systems*, page 8. ACM, 2013.

M.A. Jaro. Advances in record-linkage methodology as applied to matching the 1985 census of Tampa, Florida. *Journal of the American Statistical Association*, 84(406):414–420, 1989.

Shawn R Jeffery, Liwen Sun, Matt DeLand, Nick Pendar, Rick Barber, and Andrew Galdi. Arnold: Declarative crowd-machine data integration. In *CIDR*, 2013.

Lili Jiang, Yafang Wang, Johannes Hoffart, and Gerhard Weikum. Crowdsourced entity markup. In *Proceedings of the 1st International Workshop on Crowdsourcing the Semantic Web, Sydney, Australia, October 19, 2013*, pages 59–68, 2013. URL `http://ceur-ws.org/Vol-1030/paper-04.pdf`.

Ece Kamar, Severin Hacker, and Eric Horvitz. Combining human and machine intelligence in large-scale crowdsourcing. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems - Volume 1*, AAMAS '12, pages 467–474, Richland, SC, 2012. International Foundation for Autonomous Agents and Multiagent Systems. URL `http://dl.acm.org/citation.cfm?id=2343576.2343643`.

Salil S Kanhere. Participatory sensing: Crowdsourcing data from mobile smartphones in urban spaces. In *Mobile Data Management (MDM), 2011 12th IEEE International Conference on*, volume 2, pages 3–6. IEEE, 2011.

David R Karger, Sewoong Oh, and Devavrat Shah. Budget-optimal task allocation for reliable crowdsourcing systems. *Operations Research*, 62(1): 1–24, 2014.

Gabriella Kazai, Jaap Kamps, Marijn Koolen, and Natasa Milic-Frayling. Crowdsourcing for book search evaluation: impact of hit design on comparative system ranking. In *Proceedings of the 34th international ACM SIGIR conference on Research and development in Information - SIGIR '11*, page 205, 2011a. URL `http://dl.acm.org/citation.cfm?id=2009916.2009947`.

Gabriella Kazai, Jaap Kamps, and Natasa Milic-Frayling. Worker types and personality traits in crowdsourcing relevance labels. In *Proceedings of the 20th ACM international conference on Information and knowledge management*, pages 1941–1944. ACM, 2011b.

Gabriella Kazai, Jaap Kamps, and Natasa Milic-Frayling. An analysis of human factors and label accuracy in crowdsourcing relevance judgments. *Information Retrieval*, 16(2):138–178, 2013.

David Kent, Morteza Behrooz, and Sonia Chernova. Crowdsourcing the construction of a 3d object recognition database for robotic grasping. In *2014 IEEE International Conference on Robotics and Automation (ICRA)*, pages 4526–4531. IEEE, 2014.

Vassilis Kostakos, Jakob Rogstadius, Denzil Ferreira, Simo Hosio, and Jorge Goncalves. Human sensors. In *Participatory Sensing, Opinions and Collective Awareness*, pages 69–92. Springer, 2017.

Allan Kuchinsky, Celine Pering, Michael L Creech, Dennis Freeze, Bill Serra, and Jacek Gwizdka. Fotofile: a consumer multimedia organization and retrieval system. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 496–503. ACM, 1999.

Siou Chew Kuek, Cecilia Paradi-Guilford, Toks Fayomi, Saori Imaizumi, Panos Ipeirotis, Patricia Pina, and Manpreet Singh. The global opportunity in online outsourcing. *World Bank, Washington, DC*, 2015. URL `https://openknowledge.worldbank.org/handle/10986/22284`.

Gierad Laput, Walter S Lasecki, Jason Wiese, Robert Xiao, Jeffrey P Bigham, and Chris Harrison. Zensors: Adaptive, rapidly deployable, human-intelligent sensor feeds. In *Proceedings of the 33rd Annual ACM Conference on Human Factors in Computing Systems*, pages 1935–1944. ACM, 2015.

Martha Larson, Mohammad Soleymani, Maria Eskevich, Pavel Serdyukov, Roeland Ordelman, and Gareth Jones. The community and the crowd: Multimedia benchmark dataset development. *IEEE Multimedia*, 19(3):15–23, 2012.

Walter S Lasecki, Christopher Homan, and Jeffrey P Bigham. Architecting real-time crowd-powered systems. *Human Computation*, 1(1), 2014.

Edith Law and Luis von Ahn. Human computation. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 5(3):1–121, 2011.

Edith Law, Ming Yin, Kevin Chen Joslin Goh, Michael Terry, and Krzysztof Z Gajos. Curiosity killed the cat, but makes crowdwork better. *Proceedings of CHI'16*, 2016.

Florian Laws, Christian Scheible, and Hinrich Schütze. Active learning with amazon mechanical turk. In *Proceedings of the Conference on Empirical Methods in Natural Language Processing*, EMNLP '11, pages 1546–1556, Stroudsburg, PA, USA, 2011. Association for Computational Linguistics. URL `http://dl.acm.org/citation.cfm?id=2145432.2145597`.

V.I. Levenshtein. Binary codes capable of correcting deletions, insertions, and reversals. In *Soviet Physics Doklady*, volume 10, pages 707–710, 1966.

Ee-Peng Lim, Jaideep Srivastava, Satya Prabhakar, and James Richardson. Entity identification in database integration. In *Data Engineering, 1993. Proceedings. Ninth International Conference on*, pages 294–301. IEEE, 1993.

Chris J Lintott, Kevin Schawinski, Anže Slosar, Kate Land, Steven Bamford, Daniel Thomas, M Jordan Raddick, Robert C Nichol, Alex Szalay, Dan Andreescu, et al. Galaxy zoo: morphologies derived from visual inspection of galaxies from the sloan digital sky survey. *Monthly Notices of the Royal Astronomical Society*, 389(3):1179–1189, 2008.

Christoph Lofi, Kinda El Maarry, and Wolf-Tilo Balke. Skyline queries in crowd-enabled databases. In *Proceedings of the 16th International Conference on Extending Database Technology*, EDBT '13, pages 465–476, New York, NY, USA, 2013. ACM. URL `http://doi.acm.org/10.1145/2452376.2452431`.

Jerome P Lynch and Kenneth J Loh. A summary review of wireless sensors and sensor networks for structural health monitoring. *Shock and Vibration Digest*, 38(2):91–130, 2006.

Eddy Maddalena, Kevin Roitero, Gianluca Demartini, and Stefano Mizzaro. Considering assessor agreement in ir evaluation. In *Proceedings of the ACM SIGIR International Conference on Theory of Information Retrieval*, ICTIR '17, pages 75–82, New York, NY, USA, 2017. ACM. URL `http://doi.acm.org/10.1145/3121050.3121060`.

Aamer Mahmood, Walid G Aref, Eduard Dragut, and Saleh Basalamah. The palm-tree index: Indexing with the crowd. In *DBCrowd 2013: First VLDB Workshop on Databases and Crowdsourcing*, 2013.

Christina Manzo, Geoff F Kaufman, Sukdith Punjasthitkul, and Mary Flanagan. " by the people, for the people": Assessing the value of crowdsourced, user-generated metadata. *Digital Humanities Quarterly*, 9(1), 2015.

Adam Marcus. *Optimization techniques for human computation-enabled data processing systems*. PhD thesis, Massachusetts Institute of Technology, 2012.

Adam Marcus, Eugene Wu, David Karger, Samuel Madden, and Robert Miller. Human-powered sorts and joins. *Proceedings of the VLDB Endowment*, 5(1):13–24, 2011a.

Adam Marcus, Eugene Wu, David R Karger, Samuel Madden, and Robert C Miller. Crowdsourced databases: Query processing with people. In *Conference on Innovative Data Systems Research*, 2011b.

Adam Marcus, David Karger, Samuel Madden, Robert Miller, and Sewoong Oh. Counting with the crowd. *Proceedings of the VLDB Endowment*, 6(2): 109–120, 2012.

Milan Markovic, Peter Edwards, and David Corsar. A role for provenance in social computation. In *Proceedings of the 1st International Workshop on Crowdsourcing the Semantic Web, Sydney, Australia, October 19, 2013*, pages 93–96, 2013. URL `http://ceur-ws.org/Vol-1030/paper-05.pdf`.

David Martin, Benjamin V Hanrahan, Jacki O'Neill, and Neha Gupta. Being a turker. In *Proceedings of the 17th ACM conference on Computer supported cooperative work & social computing*, pages 224–235. ACM, 2014.

Tyler McDonnell, Matthew Lease, Mucahid Kutlu, and Tamer Elsayed. Why is that relevant? collecting annotator rationales for relevance judgments. In *Proceedings of the 4th AAAI Conference on Human Computation and Crowdsourcing (HCOMP)*, pages 139–148. AAAI, 2016.

Marija Milenkovic and Oliver Amft. An opportunistic activity-sensing approach to save energy in office buildings. In *Proceedings of the fourth international conference on Future energy systems*, pages 247–258. ACM, 2013.

Jonathan Mortensen. Crowdsourcing ontology verification. In *The Semantic Web - ISWC 2013 - 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II*, pages 448–455, 2013. URL `http://dx.doi.org/10.1007/978-3-642-41338-4_30`.

Jonathan Mortensen, Mark A. Musen, and Natasha F. Noy. Crowdsourcing the verification of relationships in biomedical ontologies. In *AMIA 2013, American Medical Informatics Association Annual Symposium, Washington, DC, USA, November 16-20, 2013*, 2013. URL `http://knowledge.amia.org/amia-55142-a2013e-1.580047/t-09-1.582024/f-009-1.582025/a-345-1.582084/a-363-1.582079`.

Jonathan M Mortensen, Evan P Minty, Michael Januszyk, Timothy E Sweeney, Alan L Rector, Natalya F Noy, and Mark A Musen. Using the wisdom of the crowds to find critical errors in biomedical ontologies: a study of snomed ct. *Journal of the American Medical Informatics Association*, 22 (3):640–648, 2015.

Barzan Mozafari, Purna Sarkar, Michael Franklin, Michael Jordan, and Samuel Madden. Scaling up crowd-sourcing to very large datasets: A case for active learning. *Proceedings of the VLDB Endowment*, 8(2), 2014.

Y Yi Mun and Yujong Hwang. Predicting the use of web-based information systems: self-efficacy, enjoyment, learning goal orientation, and the technology acceptance model. *International journal of human-computer studies*, 59(4):431–449, 2003.

Christian Nieke, Ulrich Güntzer, and Wolf-Tilo Balke. Topcrowd. In *Conceptual Modeling*, pages 122–135. Springer, 2014.

Natalya F Noy, Jonathan Mortensen, Mark A Musen, and Paul R Alexander. Mechanical turk as an ontology engineer?: using microtasks as a component of an ontology-engineering workflow. In *Proceedings of the 5th Annual ACM Web Science Conference*, pages 262–271. ACM, 2013.

David Oleson, Alexander Sorokin, Greg Laughlin, Vaughn Hester, John Le, and Lukas Biewald. Programmatic gold: Targeted and scalable quality assurance in crowdsourcing. In *Proceedings of the AAAI Conference on Human Computation*, AAAIWS'11-11, pages 43–48. AAAI Press, 2011. URL `http://dl.acm.org/citation.cfm?id=2908698.2908706`.

B.W. On, N. Koudas, D. Lee, and D. Srivastava. Group linkage. In *Data Engineering, 2007. ICDE 2007. IEEE 23rd International Conference on*, pages 496–505. IEEE, 2007.

Johan Oomen and Lora Aroyo. Crowdsourcing in the cultural heritage domain: Opportunities and challenges. In *Proceedings of the 5th International Conference on Communities and Technologies*, C&T '11, pages 138–149, New York, NY, USA, 2011. ACM. URL `http://doi.acm.org/10.1145/2103354.2103373`.

George Papadakis, Ekaterini Ioannou, Claudia Niederée, Themis Palpanas, and Wolfgang Nejdl. Beyond 100 million entities: large-scale blocking-based resolution for heterogeneous data. In *Proceedings of the fifth ACM international conference on Web search and data mining*, WSDM '12, pages 53–62, New York, NY, USA, 2012. ACM. URL `http://doi.acm.org/10.1145/2124295.2124305`.

A Parameswaran and N Polyzotis. Answering queries using databases, humans and algorithms. In *Conference on Innovative Data Systems Research*, volume 160, 2011.

Aditya G Parameswaran, Hector Garcia-Molina, Hyunjung Park, Neoklis Polyzotis, Aditya Ramesh, and Jennifer Widom. Crowdscreen: Algorithms for filtering data with humans. In *Proceedings of the 2012 ACM SIGMOD International Conference on Management of Data*, pages 361–372. ACM, 2012.

Vassilis Polychronopoulos, Luca de Alfaro, James Davis, Hector Garcia-Molina, and Neoklis Polyzotis. Human-powered top-k lists. In *WebDB*, pages 25–30, 2013.

Martin Potthast, Benno Stein, and Robert Gerling. Automatic vandalism detection in wikipedia. In *European Conference on Information Retrieval*, pages 663–668. Springer, 2008.

Martin Potthast, Matthias Hagen, Michael Völske, and Benno Stein. Crowdsourcing interaction logs to understand text reuse from the web. In *ACL (1)*, pages 1212–1221, 2013.

Roman Prokofyev, Alberto Tonon, Michael Luggen, Loic Vouilloz, Djellel Eddine Difallah, and Philippe Cudré-Mauroux. Sanaphor: ontology-based coreference resolution. In *International Semantic Web Conference*, pages 458–473. Springer, 2015.

Sasank Reddy, Andrew Parker, Josh Hyman, Jeff Burke, Deborah Estrin, and Mark Hansen. Image browsing, processing, and clustering for participatory sensing: lessons from a dietsense prototype. In *Proceedings of the 4th workshop on Embedded networked sensors*, pages 13–17. ACM, 2007.

Mia Ridge. From tagging to theorizing: Deepening engagement with cultural heritage through crowdsourcing. *Curator: The Museum Journal*, 56(4): 435–450, 2013. URL `http://dx.doi.org/10.1111/cura.12046`.

Mia Ridge. *Crowdsourcing our cultural heritage*. Ashgate Publishing, Ltd., 2014.

Joel Ross, Lilly Irani, M Silberman, Andrew Zaldivar, and Bill Tomlinson. Who are the crowdworkers?: shifting demographics in mechanical turk. In *CHI'10 Extended Abstracts on Human Factors in Computing Systems*, pages 2863–2872. ACM, 2010.

Spencer Rothwell, Steele Carter, Ahmad Elshenawy, Vladislavs Dovgalecs, Safiyyah Saleem, Daniela Braga, and Bob Kennewick. Data collection and annotation for state-of-the-art ner using unmanaged crowds. In *Sixteenth Annual Conference of the International Speech Communication Association*, 2015a.

Spencer Rothwell, Ahmad Elshenawy, Steele Carter, Daniela Braga, Faraz Romani, Michael Kennewick, and Bob Kennewick. Controlling quality and handling fraud in large scale crowdsourcing speech data collections. In *Sixteenth Annual Conference of the International Speech Communication Association*, 2015b.

Jeffrey M Rzeszotarski and Aniket Kittur. Instrumenting the crowd: using implicit behavioral measures to predict task performance. In *Proceedings of the 24th annual ACM symposium on User interface software and technology*, pages 13–22. ACM, 2011.

Cristina Sarasua, Elena Simperl, and Natalya Fridman Noy. CrowdMap: Crowdsourcing Ontology Alignment with Microtasks. In *The Semantic Web - ISWC 2012 - 11th International Semantic Web Conference, Boston, MA, USA, November 11-15, 2012, Proceedings, Part I*, pages 525–541, 2012. URL `http://dx.doi.org/10.1007/978-3-642-35176-1_33`.

Joachim Selke, Christoph Lofi, and Wolf-Tilo Balke. Pushing the boundaries of crowd-enabled databases with query-driven schema expansion. *Proc. VLDB Endow.*, 5(6):538–549, February 2012. URL `http://dl.acm.org/citation.cfm?id=2168651.2168655`.

Vinay Shashidhar, Nishant Pandey, and Varun Aggarwal. Automatic spontaneous speech grading: A novel feature derivation technique using the crowd. In *ACL (1)*, pages 1085–1094, 2015.

Victor S Sheng, Foster Provost, and Panagiotis G Ipeirotis. Get another label? improving data quality and data mining using multiple, noisy labelers. In *Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining*, pages 614–622. ACM, 2008.

Aashish Sheshadri and Matthew Lease. SQUARE: A Benchmark for Research on Computing Crowd Consensus. In *Proceedings of the 1st AAAI Conference on Human Computation (HCOMP)*, pages 156–164, 2013. URL `http://ir.ischool.utexas.edu/square/documents/sheshadri.pdf`.

Edwin Simpson and Stephen Roberts. Bayesian methods for intelligent task assignment in crowdsourcing systems. In *Decision Making: Uncertainty, Imperfection, Deliberation and Scalability*, pages 1–32. Springer, 2015.

Edwin D. Simpson, Matteo Venanzi, Steven Reece, Pushmeet Kohli, John Guiver, Stephen J. Roberts, and Nicholas R. Jennings. Language understanding in the wild: Combining crowdsourcing and machine learning. In *Proceedings of the 24th International Conference on World Wide Web*, WWW '15, pages 992–1002, Republic and Canton of Geneva, Switzerland, 2015. International World Wide Web Conferences Steering Committee. URL `https://doi.org/10.1145/2736277.2741689`.

Philipp Singer, Denis Helic, Andreas Hotho, and Markus Strohmaier. Hyptrails: A bayesian approach for comparing hypotheses about human trails on the web. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1003–1013. ACM, 2015.

Md Smucker, Gabriella Kazai, and Matthew Lease. Overview of the TREC 2013 Crowdsourcing Track. Technical report, DTIC Document, 2014.

Cees GM Snoek, Bauke Freiburg, Johan Oomen, and Roeland Ordelman. Crowdsourcing rock n'roll multimedia retrieval. In *Proceedings of the 18th ACM international conference on Multimedia*, pages 1535–1538. ACM, 2010.

Robert Soden and Leysia Palen. From crowdsourced mapping to community mapping: The post-earthquake work of openstreetmap haiti. In *COOP 2014-Proceedings of the 11th International Conference on the Design of Cooperative Systems, 27-30 May 2014, Nice (France)*, pages 311–326. Springer, 2014.

Matthias Stevens and Ellie D'Hondt. Crowdsourcing of pollution data using smartphones. In *Workshop on Ubiquitous Crowdsourcing*, 2010.

Yuyin Sun, Adish Singla, Tori Yan, Andreas Krause, and Dieter Fox. Evaluating task-dependent taxonomies for navigation. In *Proceedings of the 4th AAAI Conference on Human Computation (HCOMP)*, 2016.

Jaime Teevan, Kevyn Collins-Thompson, Ryen W. White, and Susan Dumais. Slow search. *Commun. ACM*, 57(8):36–38, August 2014. URL `http://doi.acm.org/10.1145/2633041`.

Alberto Tonon, Gianluca Demartini, and Philippe Cudré-Mauroux. Pooling-based continuous evaluation of information retrieval systems. *Inf. Retr. Journal*, 18(5):445–472, 2015. URL `http://dx.doi.org/10.1007/s10791-015-9266-y`.

John Travis. Science by the masses, 2008.

Beth Trushkowsky, Tim Kraska, Michael J Franklin, and Purnamrita Sarkar. Crowdsourced enumeration queries. In *Data Engineering (ICDE), 2013 IEEE 29th International Conference on*, pages 673–684. IEEE, 2013.

Vinod Kumar Vavilapalli, Arun C Murthy, Chris Douglas, Sharad Agarwal, Mahadev Konar, Robert Evans, Thomas Graves, Jason Lowe, Hitesh Shah, Siddharth Seth, et al. Apache hadoop yarn: Yet another resource negotiator. In *Proceedings of the 4th annual Symposium on Cloud Computing*, page 5. ACM, 2013.

Matteo Venanzi, John Guiver, Gabriella Kazai, Pushmeet Kohli, and Milad Shokouhi. Community-based bayesian aggregation models for crowdsourcing. In *Proceedings of the 23rd International Conference on World Wide Web*, WWW '14, pages 155–164, New York, NY, USA, 2014. ACM. URL `http://doi.acm.org/10.1145/2566486.2567989`.

Petros Venetis, Hector Garcia-Molina, Kerui Huang, and Neoklis Polyzotis. Max algorithms in crowdsourcing environments. In *Proceedings of the 21st international conference on World Wide Web*, pages 989–998. ACM, 2012.

Sudheendra Vijayanarasimhan and Kristen Grauman. Large-scale live active learning: Training object detectors with crawled data and crowds. *International Journal of Computer Vision*, 108(1):97–114, 2014. URL http://dx.doi.org/10.1007/s11263-014-0721-9.

L. Von Ahn, B. Maurer, C. McMillen, D. Abraham, and M. Blum. recaptcha: Human-based character recognition via web security measures. *Science*, 321(5895):1465–1468, 2008.

Luis Von Ahn. Games with a purpose. *Computer*, 39(6):92–94, 2006.

Luis von Ahn and Laura Dabbish. Labeling images with a computer game. In *CHI '04*, pages 319–326. ACM, 2004. URL http://doi.acm.org/10.1145/985692.985733.

Jiannan Wang, Tim Kraska, Michael J Franklin, and Jianhua Feng. Crowder: Crowdsourcing entity resolution. *Proceedings of the VLDB Endowment*, 5 (11):1483–1494, 2012.

Jiannan Wang, Guoliang Li, Tim Kraska, Michael J Franklin, and Jianhua Feng. Leveraging transitive relations for crowdsourced joins. In *Proceedings of the 2013 international conference on Management of data*, pages 229–240. ACM, 2013.

Yuhui Wang and Mohan S Kankanhalli. Tweeting cameras for event detection. In *Proceedings of the 24th International Conference on World Wide Web*, pages 1231–1241. ACM, 2015.

Steven Euijong Whang, David Menestrina, Georgia Koutrika, Martin Theobald, and Hector Garcia-Molina. Entity resolution with iterative blocking. In *Proceedings of the 2009 ACM SIGMOD International Conference on Management of data*, SIGMOD '09, pages 219–232, New York, NY, USA, 2009. ACM. URL http://doi.acm.org/10.1145/1559845.1559870.

Andrea Wiggins and Yurong He. Community-based data validation practices in citizen science. In *Proceedings of the 19th ACM Conference on Computer-Supported Cooperative Work & Social Computing*, CSCW '16, pages 1548–1559, New York, NY, USA, 2016. ACM. URL http://doi.acm.org/10.1145/2818048.2820063.

Thomas D Wilson. Human information behavior. *Informing science*, 3(2): 49–56, 2000.

W.E. Winkler. The state of record linkage and current research problems. In *Statistical Research Division, US Census Bureau*, 1999.

Samantha Wray and Ahmed Ali. Crowdsource a little to label a lot: Labeling a speech corpus of dialectal arabic. In *Sixteenth Annual Conference of the International Speech Communication Association*, 2015.

Xiao-Feng Xie and Zun-Jing Wang. An empirical study of combining participatory and physical sensing to better understand and improve urban mobility networks. In *Transportation Research Board 94th Annual Meeting*, number 15-3238, 2015.

Rui Yan, Mingkun Gao, Ellie Pavlick, and Chris Callison-Burch. Are two heads better than one? crowdsourced translation via a two-step collaboration of non-professional translators and editors. In *ACL (1)*, pages 1134–1144. Citeseer, 2014.

Tingxin Yan, Vikas Kumar, and Deepak Ganesan. Crowdsearch: Exploiting crowds for accurate real-time image search on mobile phones. In *Proceedings of the 8th International Conference on Mobile Systems, Applications, and Services*, MobiSys '10, pages 77–90, New York, NY, USA, 2010. ACM. URL `http://doi.acm.org/10.1145/1814433.1814443`.

Jie Yang, Judith Redi, Gianluca Demartini, and Alessandro Bozzon. Modeling task complexity in crowdsourcing. In *Proceedings of The Fourth AAAI Conference on Human Computation and Crowdsourcing (HCOMP 2016)*, pages 249–258. AAAI, 2016.

Chih-Hao Yu, Tudor Groza, and Jane Hunter. Reasoning on crowd-sourced semantic annotations to facilitate cataloguing of 3d artefacts in the cultural heritage domain. In *The Semantic Web - ISWC 2013 - 12th International Semantic Web Conference, Sydney, NSW, Australia, October 21-25, 2013, Proceedings, Part II*, pages 228–243, 2013. URL `http://dx.doi.org/10.1007/978-3-642-41338-4_15`.

Andrea Zanella, Nicola Bui, Angelo Castellani, Lorenzo Vangelista, and Michele Zorzi. Internet of things for smart cities. *IEEE Internet of Things journal*, 1(1):22–32, 2014.

L. Zhao, G. Sukthankar, and R. Sukthankar. Incremental relabeling for active learning with noisy crowdsourced annotations. In *2011 IEEE Third International Conference on Privacy, Security, Risk and Trust and 2011 IEEE Third International Conference on Social Computing*, pages 728–733, Oct 2011.

Mengdie Zhuang, Elaine G Toms, and Gianluca Demartini. The relationship between user perception and user behaviour in interactive information retrieval evaluation. In *European Conference on Information Retrieval*, pages 293–305. Springer, 2016.