

# **MICROTASK CROWDSOURCING TO SOLVE SEMANTIC WEB PROBLEMS**

**FULL-DAY TUTORIAL @ISWC2013**

**SYDNEY, AUSTRALIA**

10/21/13

Tutorial@ISWC2013



# TUTORIAL OVERVIEW

- **Microtask crowdsourcing is increasingly used to augment the results of algorithms solving Semantic Web problems**
- **This tutorial covers**
  - Fundamentals of microtask crowdsourcing
  - How-to Mechanical Turk (and CrowdFlower)
  - Latest research in microtask crowdsourcing



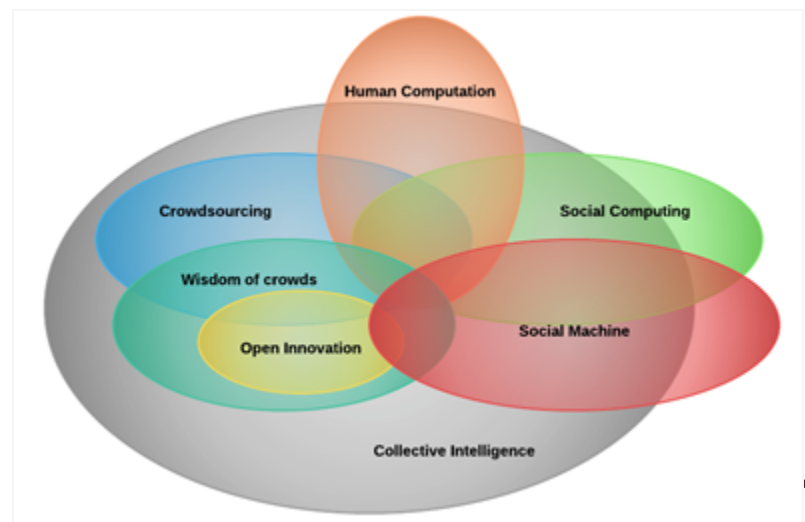
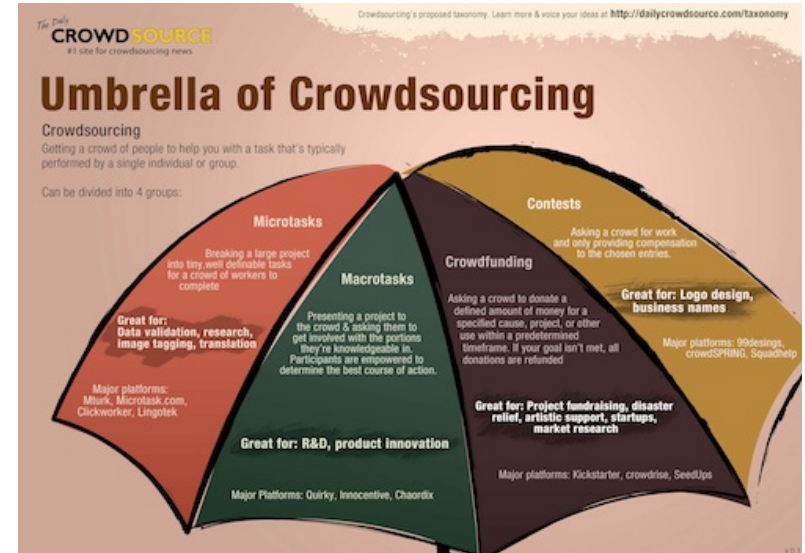
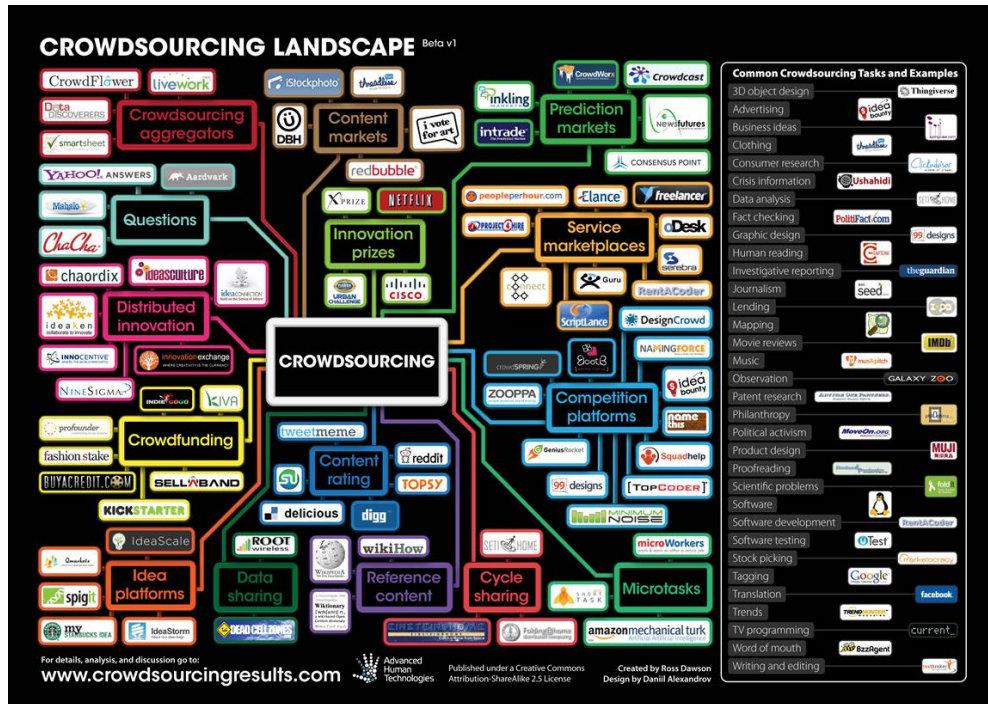
# CROWDSOURCING: PROBLEM SOLVING VIA OPEN CALLS

"Simply defined, crowdsourcing represents the act of a company or institution taking a function once performed by employees and outsourcing it to an undefined (and generally large) network of people in the form of an open call. This can take the form of peer-production (when the job is performed collaboratively), but is also often undertaken by sole individuals. The crucial prerequisite is the use of the open call format and the large network of potential laborers."

[Howe, 2006]



# CROWDSOURCING COMES IN DIFFERENT FORMS AND FLAVORS



10/21/13

Tutorial@ISWC2013



# IN THIS TUTORIAL: CROWDSOURCING AS HUMAN COMPUTATION

Outsourcing tasks that machines find difficult to solve to humans



# IN THIS TUTORIAL: MICROTASK CROWDSOURCING

Work is broken down into smaller pieces that can be solved independently

The collage features several prominent elements:

- servio**: A blue banner with the text "Need to get work done?" and "Want to earn money now?".
- CloudCrowd**: A blue banner with the text "500,000+ Hyper-Specialized Workers On Demand".
- microTASK**: A yellow banner with the text "you send samasource a project", "samasource breaks it down into microwork", and "work is allocated to our service partners".
- One Billion MobileWorks**: A blue banner with the text "Crowdsourcing On Demand. Faster. More accurate. Painless." and "thousands of tasks possible!".
- The World's Largest Workforce**: A large orange banner with the text "Instantly hire millions of people to collect, filter, and enhance your data." and "Real-Time Crowd Labor".
- CloudCrowd**: A blue banner with the text "We're working on it. Lots of us."
- Freebase**: A banner featuring three glass jars containing blue, brown, and green powders.
- CloudFlare**: A banner with the text "RTFM Real Time Frito Moderator" and "Sentinel Sentiment Analysis".
- MobileWorks**: A banner with the text "Try out applications that use the MobileWorks crowd" and "Build your Own".

At the bottom right, there is a large number **6**.

# AGENDA FOR TODAY

Microtask crowdsourcing fundamentals	09:15 – 10:30
Coffee break	10:30 – 11:00
MTurk hands-on	11:00 – 12:45
Lunch break	12:45 – 13:45
Microtask management and quality control	13:45 – 15:30
Coffee break	15:30 – 16:00
Applications to the Semantic Web	16:00 – 17:00
Current research directions	17:00 – 17:15
Wrap-up	17:15 – 17:30

# PRESENTERS

- **Gianluca Demartini**
  - MSc University of Udine, Italy
  - PhD University of Hannover, Germany on Entity Retrieval
  - Worked for UC Berkeley (on crowdsourcing); Yahoo! Research, Spain; L3S Research Center, Germany
  - PostDoc eXascale Infolab, University Fribourg, Switzerland
  - Lecturer for Social Computing in Fribourg
  - Tutorial on Entity Search at ECIR 2012
- **Research interests**
  - Information retrieval
  - Social and Semantic Web



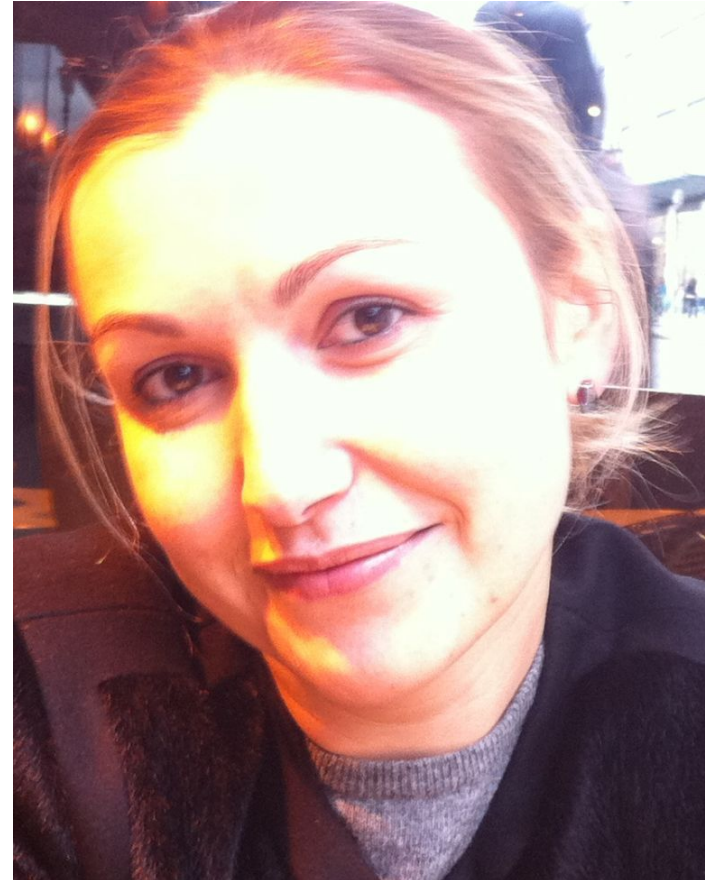
[demartini@exascale.info](mailto:demartini@exascale.info)



# PRESENTERS

- **Elena Simperl**

- PhD in Computer Science  
FU Berlin, Germany
- Worked for FU Berlin,  
Germany; STI Innsbruck,  
Austria; KIT, Germany
- Senior lecturer WAIS,  
University of Southampton,  
UK
- Coordinator of **Insemtives**  
project and tutorials at  
ESWC/ISWC
- **Research interests**
  - Social computing
  - Semantic technologies



[e.simperl@soton.ac.uk](mailto:e.simperl@soton.ac.uk)

# PRESENTERS

- **Maribel Acosta**
  - MSC in Computer Science University Simon Bolivar, Venezuela
  - Worked for University Simon Bolivar, Venezuela
  - PhD student, KIT, Germany
- **Research interests**
  - Data base management
  - Linked Data query processing
  - Social Web



maribel.acosta@kit.edu

**WEB SITE: [HTTPS://SITES.GOOGLE.COM/  
SITE/MICROTASKTUTORIAL/](https://sites.google.com/site/microtasktutorial/)**

**SLIDES AND EXERCISES: [HTTPS://  
GITHUB.COM/MARIBELACOSTA/  
CROWDSOURCING-TUTORIAL](https://github.com/maribelacosta/crowdsourcing-tutorial)**

**FULL-DAY TUTORIAL ISWC2013  
SYDNEY AUSTRALIA**

10/21/13

Tutorial@ISWC2013

