

> **19/08/2015**

Aggregating Labels in Crowdsourcing Data

CERN Openlab Summer Students
Lightning Talks Sessions

Maria Priisalu

Crowdsourcing

- › **Many workers**
- › **Easy tasks**
- › **Multiple answers**
- › **Little or no prior knowledge**
- › **Aggregate answers**



Advantages of Crowdsourcing

- › **Costs less**
- › **Faster**
- › **Educate the worker**



Research in Crowdsourcing

- › Hybrid systems
- › Task scheduling
- › Learning algorithms
- › Aggregation

@WalmartLabs



Our implementation


FEATURED
10


SOCIAL
151


ART
39


HUMANITIES
21


BIOLOGY
7


ECONOMICS
13

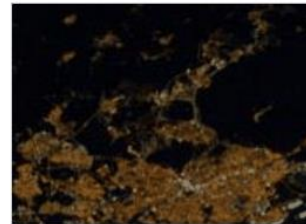

SCIENCE
180



Dark Skies ISS
Image pattern recognition



Lost at Night
Locate a City in images f...



Night Cities ISS
Georeference City images...



Air Quality with Biomarkers: Lichens
Air Quality using Lichens...



> Dawid-Skene model (EM)

Future (and current) work

- › **Aggregation of other types of labels**
- › **Sociological aggregation**
- › **Educating the crowdworker**



Maria Priisalu

Note: Image from Windows Clipart



#cernwebfest

CERN Summer Student Webfest 2015



July 31 – Aug. 2
Restaurant 1



the LHC
+ tech
+ data
+ creativity
+ challenges
+ community
+ you

A collage of various purple icons representing different fields: a lightbulb, a coffee cup, a key, a house, a tree, a guitar, a smartphone, a DNA helix, a microscope, a trophy, an airplane, a globe, a rocket, a test tube, a leaf, a gear, a speech bubble, a server rack, a signal wave, a battery, a microscope, a globe, a rocket, a test tube, a leaf, a gear, a speech bubble, a server rack, a signal wave, a battery.

Come develop world-changing apps, collaborate with incredible people, and innovate the future of the web, the LHC, and the way we understand the world.



Science Lab



Pori



Citizen CyberLab



19/08/2014

design: rellie, igor, jenny, daniel, tim, trebble, bill & greg, alexandra, cecilia

The Webfest