# Crowdsourcing Linguistic Resources

Jožef Stefan Institute, Ljubljana

Mathias Müller

## What are we talking about?

Nowadays, crowdsourcing is used to create linguistic resources

ommunity    What's On    Jobs    Motoring    Real Estate    Obituaries    Classifieds    ALL

## couple may crowdsource son's name
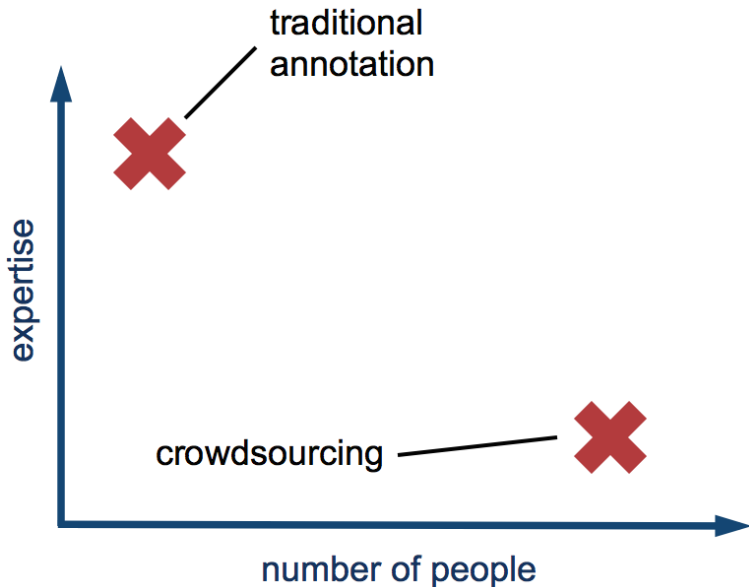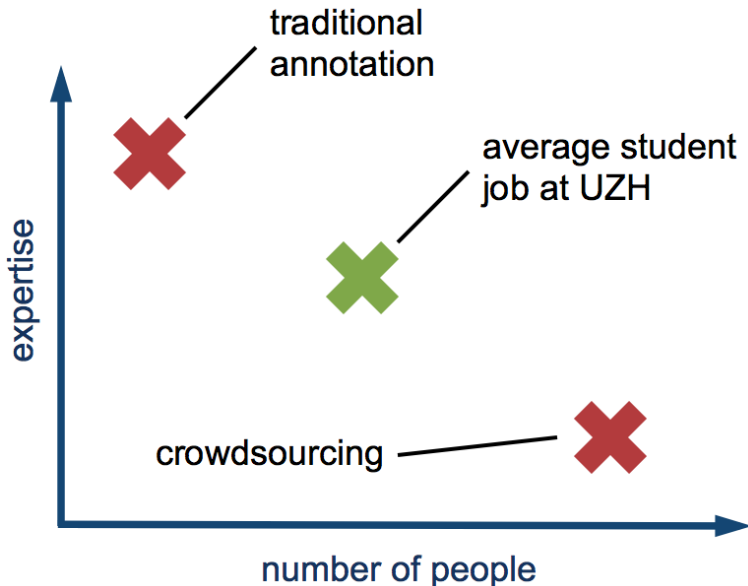
⬚ USE THIS CONTENT

## RRS BOATY MCBOATFACE

**JAMES HAND**

# Versus traditional annotation

# Versus traditional annotation



traditional annotation

average student job at UZH

expertise

crowdsourcing

number of people

## Versus traditional annotation

Taking into account more relevant factors:

|  | **traditional** |
|---|---|
| expertise | high |
| number of people | few |
| cost | high |
| risk | low |
| time span | slow |
| fun | no |

How is crowdsourcing different?

**Versus traditional annotation**

How is crowdsourcing different?

|                  | **traditional** | **crowdsourcing** |
|------------------|-----------------|-------------------|
| expertise        | high            | low               |
| number of people | few             | a lot             |
| cost             | high            | low               |
| risk             | low             | low               |
| time span        | slow            | fast              |
| fun              | no              | no                |

# Meta analyses

# Sampling the literature

Felt et al. (2015a)
Ambati et al. (2010)
Boyd-Graber and Satinoff (2012)
Feizabadi and Padó (2014)
Gao et al. (2015)
Graham et al. (2013)
Hovy et al. (2014)
Kawahara et al. (2014)
Li and Lee-Urban (2012)
Ling et al. (2013)
Madnani and Chodorow (2011)
Meena and Gustafson (2014)
Nguyen et al. (2014)
Post et al. (2012)
Asheghi et al. (2014)
Sayeed et al. (2011)
Takabatake (2015)
Tschirsich and Hintz (2013)
Wang et al. (2012)
Wang et al. (2015)
Zaidan and Callison-Burch (2011)
Borg and Gatt (2014)
Hsueh et al. (2009)
Demartini et al. (2013)
Kunchukuttan et al. (2013)
Amancio and Specia (2014)
Sabou et al. (2014)

Felt et al. (2015b)
Aroyo and Welty (2013)
Cocos et al. (2015)
Filatova (2012)
Götze and Boye (2015)
Bonial et al. (2014)
Jha et al. (2010)
Khapra et al. (2014)
Clematide et al. (2016)
Littell et al. (2014)
Madnani et al. (2015)
Mitchell and Bohus (2014)
Novotney and Callison-Burch (2010)
Prabhakaran et al. (2012)
Rumshisky (2011)
Schmidt et al. (2015)
Søgaard et al. (2013)
Vertanen and Kristensson (2011)
Wang et al. (2014a)
Wray et al. (2015)
Zbib et al. (2013)
Bessho et al. (2012)
Bontcheva et al. (2014b)
Goldman et al. (2014)
Trisedya and Manurung (2012)
Munro et al. (2010)
Wang et al. (2013)

Al-sabbagh et al. (2014)
Bontcheva et al. (2014a)
Farra et al. (2015)
Finin et al. (2010)
Grady and Lease (2010)
Higgins et al. (2010)
Ng and Kan (2012)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Lopez de Lacalle and Agirre (2015)
Marujo et al. (2013)
Jurgens (2013)
Potthast et al. (2013)
Ramanath et al. (2013)
Sayeed et al. (2010)
Snow et al. (2008)
Talukdar and Cohen (2012)
Voyer et al. (2010)
Wang et al. (2014b)
Yan et al. (2014)
Zeichner et al. (2012)
Aker et al. (2012)
Braslavski (2014)
Hu et al. (2011)
Callison-Burch and Dredze (2010)
Sabou et al. (2012)

**Challenging beliefs**

*"A first dimension of diversification consists of the languages for which resources can be produced.*

*One advantage [of crowdsourcing] is that it allows access to foreign markets with native speakers of many rare languages"* (Sabou et al., 2012)

# By language

Amancio and Specia (2014)
Aroyo and Welty (2013)
Boyd-Graber and Satinoff (2012)
Feizabadi and Padó (2014)
Finin et al. (2010)
Grady and Lease (2010)
Hovy et al. (2014)
Jha et al. (2010)
Kunchukuttan et al. (2013)
Ling et al. (2014)
Madnani et al. (2015)
Mitchell and Bohus (2014)
Negri et al. (2011)
Post et al. (2012)
Ramanath et al. (2013)
Sayeed et al. (2010)
Talukdar and Cohen (2012)
Voyer et al. (2010)
Zaidan and Callison-Burch (2011)
Hsueh et al. (2009)
Higgins et al. (2010)
Khapra et al. (2014)
Post et al. (2012)
Wang et al. (2014b)
Post et al. (2012)
Clematide et al. (2016)
Kawahara et al. (2014)
Trisedya and Manurung (2012)
Clematide et al. (2016)
Hu et al. (2011)
Negri et al. (2011)
Braslavski (2014)
Amancio and Specia (2014)
Sabou et al. (2014)

Aker et al. (2012)
Bessho et al. (2012)
Cocos et al. (2015)
Felt et al. (2015b)
Gao et al. (2015)
Graham et al. (2013)
Hu et al. (2011)
Ng and Kan (2012)
Li and Lee-Urban (2012)
Lopez de Lacalle and Agirre (2015)
Marujo et al. (2013)
Munro et al. (2010)
Nguyen et al. (2014)
Potthast et al. (2013)
Asheghi et al. (2014)
Sayeed et al. (2012)
Tschirsich and Hintz (2013)
Wang et al. (2012)
Zbib et al. (2013)
Al-sabbagh et al. (2014)
Wray et al. (2015)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Wang et al. (2015)
Yan et al. (2014)
Negri et al. (2011)
Takabatake (2015)
Post et al. (2012)
Goldman et al. (2014)
Post et al. (2012)
Post et al. (2012)
Bontcheva et al. (2014b)
Munro et al. (2010)
Wang et al. (2013)

Ambati et al. (2010)
Bontcheva et al. (2014a)
Demartini et al. (2013)
Filatova (2012)
Götze and Boye (2015)
Bonial et al. (2015)
Jamison and Gurevych (2015)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Madnani and Chodorow (2011)
Meena and Gustafson (2014)
Jurgens (2013)
Novotney and Callison-Burch (2010)
Prabhakaran et al. (2012)
Rumshisky (2011)
Snow et al. (2008)
Vertanen and Kristensson (2011)
Yan et al. (2014)
Zeichner et al. (2012)
Farra et al. (2015)
Zbib et al. (2013)
Kunchukuttan et al. (2013)
Wang et al. (2014a)
Gao et al. (2015)
Zaidan and Callison-Burch (2011)
Schmidt et al. (2015)
Littell et al. (2014)
Liao et al. (2011)
Borg and Gatt (2014)
Post et al. (2012)
Ling et al. (2014)
Felt et al. (2015a)
Sabou et al. (2012)

# By language

**English**

Amancio and Specia (2014)
Aroyo and Welty (2013)
Boyd-Graber and Satinoff (2012)
Feizabadi and Padó (2014)
Finin et al. (2010)
Grady and Lease (2010)
Hovy et al. (2014)
Jha et al. (2010)
Kunchukuttan (2013)
Ling et al. (2014)
Madnani et al. (2013)
Mitchell and Bohus (2014)
Negri et al. (2011)
Post et al. (2012)
Ramanath et al. (2013)
Sayeed et al. (2011)
Talukdar and Cohen (2012)
Voyer et al. (2010)
Zaidan and Callison-Burch (2011)
Hsueh et al. (2009)
Higgins et al. (2010)
Khapra et al. (2014)
Post et al. (2012)
Wang et al. (2014b)
Post et al. (2012)
Clematide et al. (2016)
Kawahara et al. (2014)
Trisedya and Manurung (2012)
Clematide et al. (2016)
Hu et al. (2011)
Negri et al. (2011)
Braslavski (2014)
Amancio and Specia (2014)
Sabou et al. (2014)

Aker et al. (2012)
Bessho et al. (2012)
Cocos et al. (2015)
Felt et al. (2015b)
Gao et al. (2015)
Graham et al. (2013)
Hu et al. (2011)
Ng and Kan (2012)
Lopez Lacalle and Agirre (2012)
Marujo et al. (2013)
Munro et al. (2010)
Nguyen et al. (2014)
Potthast et al. (2013)
Asheghi et al. (2014)
Sayeed et al. (2011)
Tschirsich and Hintz (2013)
Wang et al. (2012)
Zbib et al. (2013)
Al-sabbagh et al. (2014)
Wray et al. (2015)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Wang et al. (2015)
Yan et al. (2014)
Negri et al. (2011)
Takabatake (2015)
Post et al. (2012)
Goldman et al. (2014)
Post et al. (2012)
Post et al. (2012)
Bontcheva et al. (2014b)
Munro et al. (2010)
Wang et al. (2013)

Ambati et al. (2010)
Bontcheva et al. (2014a)
Demartini et al. (2013)
Filatova (2012)
Götze and Boye (2015)
Bonial et al. (2014)
Jamison and Gurevych (2015)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Madnani and Chodorow (2011)
Mehta and Gustafson (2014)
Jurgens (2013)
Novotney and Callison-Burch (2010)
Prabhakaran et al. (2012)
Rumshisky (2011)
Snow et al. (2008)
Vertanen and Kristensson (2011)
Yan et al. (2014)
Zeichner et al. (2012)
Farra et al. (2015)
Zbib et al. (2013)
Kunchukuttan et al. (2013)
Wang et al. (2014a)
Gao et al. (2015)
Zaidan and Callison-Burch (2011)
Schmidt et al. (2015)
Littell et al. (2014)
Liao et al. (2011)
Borg and Gatt (2014)
Post et al. (2012)
Ling et al. (2014)
Felt et al. (2015a)
Sabou et al. (2012)

# By language

**English**
**Arabic**
**Hindi**
**Chinese**
**Urdu**
**German**
**Japanese**

Amancio and Specia (2014)
Aroyo and Welty (2013)
Boyd-Graber and Satinoff (2012)
Feizabadi and Padó (2014)
Finin et al. (2010)
Grady and Lease (2010)
Hovy et al. (2014)
Jha et al. (2010)
Kunchukuttan et al. (2013)
Ling et al. (2014)
Madnani et al. (2013)
Mitchell and Bohus (2014)
Negri et al. (2011)
Post et al. (2012)
Ramanath et al. (2013)
Sayeed et al. (2011)
Talukdar and Cohen (2012)
Voyer et al. (2010)
Zaidan and Callison-Burch (2011)
Hsueh et al. (2009)
Higgins et al. (2014)
Khapra et al. (2011)
Post et al. (2012)
Wang et al. (2014b)
Post et al. (2012)
Clematide et al. (2016)
Kawahara et al. (2014)
Trisedya and Manurung (2012)
Clematide et al. (2016)
Hu et al. (2011)
Negri et al. (2011)
Braslavski (2014)
Amancio and Specia (2014)
Sabou et al. (2014)

Aker et al. (2012)
Bessho et al. (2012)
Cocos et al. (2015)
Felt et al. (2015b)
Gao et al. (2015)
Graham et al. (2013)
Hu et al. (2011)
Ng and Kan (2012)
Lease and Kazai (2011)
Lopez y Lacalle and Agirre (2013)
Marujo et al. (2013)
Muro et al. (2013)
Nguyen et al. (2014)
Potthast et al. (2013)
Asheghi et al. (2014)
Sayeed et al. (2011)
Tschirsich and Hintz (2013)
Wang et al. (2014)
Zbib et al. (2013)
Al-sabbagh et al. (2013)
Wray et al. (2013)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Wang et al. (2014)
Negri et al. (2011)
Takabatake (2015)
Post et al. (2012)
Goldman et al. (2014)
Post et al. (2012)
Post et al. (2012)
Bontcheva et al. (2014b)
Munro et al. (2010)
Wang et al. (2013)

Ambati et al. (2010)
Bontcheva et al. (2014a)
Demartini et al. (2013)
Filatova (2012)
Götze and Boye (2015)
Bonial et al. (2014)
Jamison and Gurevych (2015)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Madnani and Chodorow (2011)
Mohta and Gustafson (2014)
Jurgens (2013)
Novotney and Callison-Burch (2010)
Prabhakaran et al. (2012)
Rumshisky (2011)
Snow et al. (2008)
Vertanen and Kristensson (2011)
Yan et al. (2014)
Teichner et al. (2012)
Parra et al. (2015)
Zbib et al. (2013)
Kunchukuttan et al. (2012)
Wang et al. (2014)
Gao et al. (2015)
Zaidan and Callison-Burch (2011)
Schmidt et al. (2015)
Littell et al. (2014)
Liao et al. (2011)
Borg and Gatt (2014)
Post et al. (2012)
Ling et al. (2014)
Felt et al. (2015a)
Sabou et al. (2014)

# By language

This is a word-cloud style slide with overlapping text in different colored bands.

**English** (large red text, center)

**Arabic** (large blue text)

**Chinese** (orange text)

**Hindi** (large grey text)

**Urdu** (grey text)

**German** (grey text)

**Japanese** (red text)

Names in the left column:
Amancio and Specia (2014)
Aroyo and Welty (2013)
Boyd-Graber and Satinoff (2012)
Feizabadi and Padó (2014)
Finin et al. (2010)
Grady and Lease (2010)
Hovy et al. (2014)
Jha et al. (2010)
Kunchukuttan et al. (2013)
Ling et al. (2014)
Madnani et al. (2011)
Mitchell and Bohus (2014)
Negri et al. (2011)
Post et al. (2012)
Ramanath et al. (2013)
Sayeed et al. (2011)
Talukdar and Cohen (2012)
Voyer et al. (2010)
Zaidan and Callison-Burch (2011)
Hsueh et al. (2009)
Higgins et al.
Khapra et al.
Post et al. (2012)
Wang et al. (2014b)
Post et al. (2012)
Clematide et al. (2016)
Kawahara et al. (2014)
Trisedya and Manurung (2012)
Clematide et al. (2016)
Hu et al. (2011)
Negri et al.
Braslavski
Amancio and Specia (2014)
Sabou et al. (2014)

Names in the middle column:
Aker et al. (2012)
Bessho et al. (2012)
Cocos et al. (2015)
Felt et al. (2015b)
Gao et al. (2015)
Graham et al. (2013)
Hu et al. (2011)
Ng and Kan (2012)
and Lease (2012)
Lopez Lacalle and Agir
Madnani et al. (2013)
Mohit et al.
Nguyen et al. (2014)
Potthast et al.
Asheghi et al. (2014)
Sayeed et al. (2011)
Tschirsich and Hintz (2013)
Wang et al. (2014)
Zbib et al. (2012)
Al-sabbagh et al.
Wray et al. (2013)
Kunchukuttan et al. (2012)
Liao et al. (2011)
Ng et al. (2014)
Negri et al. (2011)
Takabatake et al.
Post et al. (2012)
Goldman et al. (2014)
Post et al. (2012)
Bontcheva et al. (2014b)
Munro et al. (2010)
Wang et al. (2013)

Names in the right column:
Ambati et al. (2010)
Bontcheva et al. (2014a)
Demartini et al. (2013)
Filatova (2012)
Götze and Boye (2015)
Bonial et al. (2014)
Jamison and Gurevych (2015)
Kunchukuttan et al. (2012)
Lia et al. (2011)
Madnani and Chodorow (2011)
Megna and Gustafson (2014)
Junis (2013)
Novotney and Callison-Burch (2010)
Prabhakaran et al. (2012)
Rumshisky (2011)
Snow et al. (2008)
Vertanen and Kristensson (2011)
Yan et al. (2014)
Reichner et al. (2014)
Parra et al. (2015)
Zbib et al. (2013)
Kunchukuttan et al.
Wang et al. (2014)
Gao et al. (2015)
Zaidan and Callison-Burch (2011)
Schmidt et al. (2015)
Littell et al. (2014)
Liao et al. (2011)
Borg and Gatt (2014)
Post et al.
Felt et al. (2015a)
Sabou et al. (2012)

**singletons – but not Slowenian!**

**Challenging another belief**

*"Crowdsourcing's greatest contribution to language studies might be the ability to generate **new kinds** of data"* (Munro et al., 2010)
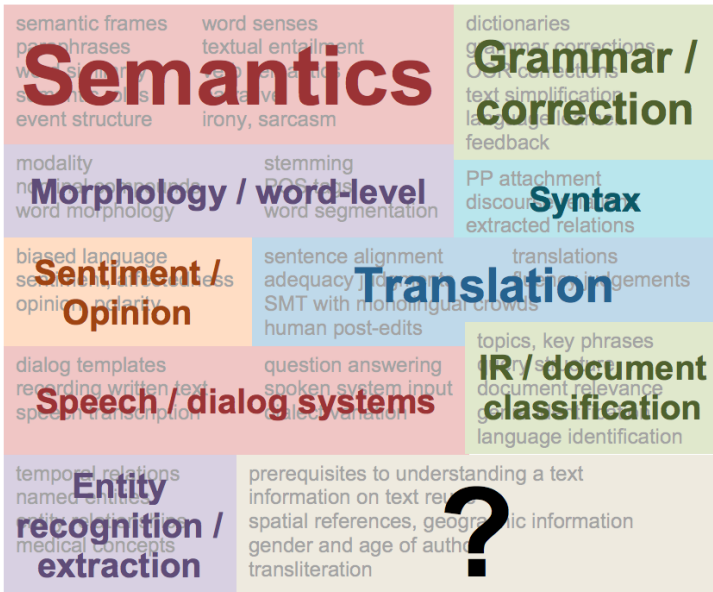
# What is crowdsourcing used for?

semantic frames
paraphrases
word similarity
semantic roles
event structure

word senses
textual entailment
verb semantics
narrative
irony, sarcasm

dictionaries
grammar corrections
OCR corrections
text simplification
language learner
feedback

modality
nominal compounds
word morphology

stemming
POS tags
word segmentation

PP attachment
discourse relations
extracted relations

biased language
sentiment, affectedness
opinion, polarity

sentence alignment
adequacy judgments
SMT with monolingual crowds
human post-edits

translations
fluency judgements

dialog templates
recording written text
speech transcription

question answering
spoken system input
dialect variation

topics, key phrases
query structure
document relevance
genre identification
language identification

temporal relations
named entities
entity relationships
medical concepts

prerequisites to understanding a text
information on text reuse
spatial references, geographic information
gender and age of author
transliteration

# What is crowdsourcing used for?

semantic frames    word senses
paraphrases    textual entailment
word similarity    verb semantics
semantic roles    narrative
event structure    irony, sarcasm

dictionaries
grammar corrections
OCR corrections
text simplification
language learner
feedback

modality    stemming
nominal compounds    POS tags
word morphology    word segmentation

PP attachment
discourse relations
extracted relations

biased language    sentence alignment    translations
sentiment, affectedness    adequacy judgments    fluency judgements
opinion, polarity    SMT with monolingual crowds
human post-edits

dialog templates    question answering
recording written text    spoken system input
speech transcription    dialect variation

topics, key phrases
query structure
document relevance
genre identification
language identification

temporal relations    prerequisites to understanding a text
named entities    information on text reuse
entity relationships    spatial references, geographic information
medical concepts    gender and age of author
transliteration

# By overall theme



**Semantics** — semantic frames, paraphrases, word senses, textual entailment, event structure, irony, sarcasm

**Grammar / correction** — dictionaries, grammar corrections, OCR corrections, text simplification, language feedback

**Morphology / word-level** — modality, stemming, POS tagging, word morphology, word segmentation

**Syntax** — PP attachment, discourse, extracted relations

**Sentiment / Opinion** — biased language, sentiment analysis, opinion polarity

**Translation** — sentence alignment, adequacy, SMT with monolingual crowds, human post-edits, translations

**Speech / dialog systems** — dialog templates, recording written text, speech transcription, question answering, spoken system input, dialect variation

**IR / document classification** — topics, key phrases, document relevance, language identification

**Entity recognition / extraction** — temporal relations, named entities, entity relations, medical concepts

**?** — prerequisites to understanding a text, information on text, spatial references, geographic information, gender and age of author, transliteration

**New kinds of data?**

prerequisites to understanding a text
information on text reuse
spatial references, geographic information
gender and age of author
transliteration

Prerequisite structure (Talukdar and Cohen, 2012) as an example:

- let crowd define the prerequisites necessary to understand Wikipedia articles
- *Is document A a prerequisite of document B?*

## New kinds of data?

> prerequisites to understanding a text
> information on text reuse
> spatial references, geographic information
> gender and age of author
> transliteration

Text reuse (Potthast et al., 2013):

- ▸ document the genesis of a text
- ▸ interactions of authors with sources

**Wrapping up**

- crowdsourced annotations are mostly English
- new kinds of data have emerged from crowdsourcing

One could also meta review:

- detailed cost analysis, validating the claim that translations can be crowdsourced at 1/10 of the price of traditional annotation (Zaidan and Callison-Burch, 2011; Zbib et al., 2013)
- how many people are in crowds?

# Bibliography I

Aker, A., El-haj, M., Albakour, M.-d., and Kruschwitz, U. (2012). Assessing Crowdsourcing Quality through Objective Tasks. Number 2007, pages 1456–1461.

Al-sabbagh, R., Girju, R., and Diesner, J. (2014). 3arif : A Corpus of Modern Standard and Egyptian Arabic Tweets Annotated for Epistemic Modality Using Interactive Crowdsourcing. *Coling*, pages 1521–1532.

Amancio, M. A. and Specia, L. (2014). An Analysis of Crowdsourced Text Simplifications. *EACL (Workshop - PITR)*, (2011):123–130.

Ambati, V., Vogel, S., and Carbonell, J. (2010). Active Learning and Crowd-Sourcing for Machine Translation. *Lrec*, pages 2169–2174.

Aroyo, L. and Welty, C. (2013). Crowd Truth: Harnessing disagreement in crowdsourcing a relation extraction gold standard. *Web Science'13*, pages 1–6.

Asheghi, N. R., Sharoff, S., and Markert, K. (2014). Designing and evaluating a reliable corpus of web genres via crowd-sourcing. In *LREC*, pages 1339–1346.

Bessho, F., Harada, T., and Kuniyoshi, Y. (2012). Dialog system using real-time crowdsourcing and Twitter large-scale corpus. *Proceedings of the 13th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, (July):227–231.

Bonial, C., Palmer, M., and Hartshorne, J. K. (2014). The VerbCorner Project : Findings from Phase 1 of Crowd-Sourcing a Semantic Decomposition of Verbs. *Acl*, (1989):397–402.

Bontcheva, K., Derczynski, L., and Roberts, I. (2014a). Crowdsourcing named entity recognition and entity linking corpora. *The Handbook of Linguistic . . .* , pages 1–18.

# Bibliography II

Bontcheva, K., Roberts, I., Derczynski, L., and Rout, D. (2014b). The GATE Crowdsourcing Plugin: Crowdsourcing Annotated Corpora Made Easy. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pages 1–4.

Borg, C. and Gatt, A. (2014). Crowd-sourcing evaluation of automatically acquired, morphologically related word groupings. In *LREC*, pages 3325–3332.

Boyd-Graber, J. and Satinoff, B. (2012). Besting the quiz master: crowdsourcing incremental classification games. *Emnlp*, (July):1290–1301.

Braslavski, P. (2014). A Spinning Wheel for YARN : User Interface for a Crowdsourced Thesaurus. *Eacl2014*, pages 101–104.

Callison-Burch, C. and Dredze, M. (2010). Creating speech and language data with amazon's mechanical turk. In *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, pages 1–12. Association for Computational Linguistics.

Clematide, S., Furrer, L., and Volk, M. (2016). Crowdsourcing an OCR Gold Standard for a German and French Heritage Corpus. *LREC*, page forthcoming.

Cocos, A., Masino, A., Qian, T., Pavlick, E., and Callison-Burch, C. (2015). Effectively Crowdsourcing Radiology Report Annotations. *Proceedings of the Sixth International Workshop on Health Text Mining and Information Analysis*, (September):109–114.

Demartini, G., Kraska, B., and Franklin, M. (2013). CrowdQ: Crowdsourced Query Understanding. *Conference on Innovative Data Systems Research (CIDR)*, page 4.

# Bibliography III

Farra, N., McKeown, K., and Habash, N. (2015). Annotating Targets of Opinions in Arabic using Crowdsourcing. *Proceedings of the Second Workshop on Arabic Natural Language Processing*, pages 89–98.

Feizabadi, P. S. and Padó, S. (2014). Crowdsourcing Annotation of Non-Local Semantic Roles. *Proceedings of the 14th Conference of the European Chapter of the Association for Computational Linguistics, volume 2: Short Papers*, pages 226–230.

Felt, P., Ringger, E., Boyd-Graber, J., and Seppi, K. (2015a). Making the Most of Crowdsourced Document Annotations: Confused Supervised LDA. *Conference on Computational Natural Language Learning*, pages 194–203.

Felt, P., Ringger, E., Seppi, K., Black, K., and Haertel, R. (2015b). Early Gains Matter : A Case for Preferring Generative over Discriminative Crowdsourcing Models. *North American Chapter of the Association for Computational Linguistics - Human Language Technologies Conference (NAACL-HLT)*, (2001):882–891.

Filatova, E. (2012). Irony and Sarcasm: Corpus Generation and Analysis Using Crowdsourcing. *Lrec*, pages 392–398.

Finin, T., Murnane, W., Karandikar, A., Keller, N., Martineau, J., and Dredze, M. (2010). Annotating Named Entities in Twitter Data with Crowdsourcing. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, 2010(January):80–88.

Gao, M., Xu, W., and Callison-burch, C. (2015). Cost Optimization for Crowdsourcing Translation. *Proc. NAACL2015*, (2008):705–713.

## Bibliography IV

Goldman, J. P., Leemann, A., Kolly, M. J., Hove, I., Almajai, I., Dellwo, V., and Moran, S. (2014). A Crowdsourcing Smartphone Application for Swiss German: Putting language documentation in the hands of the users. *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, pages 3444–3447.

Götze, J. and Boye, J. (2015). Resolving spatial references using crowdsourced geographical data. In *20th Nordic Conference on Computational Linguistics, Wilna, Lithuania, 2015*, pages 61–68. Linköping University Electronic Press.

Grady, C. and Lease, M. (2010). Crowdsourcing document relevance assessment with Mechanical Turk. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, (June):172–179.

Graham, Y., Baldwin, T., Moffat, A., and Zobel, J. (2013). Crowd-Sourcing of Human Judgments of Machine Translation Fluency. *Proceedings of the Australasian Language Technology Association Workshop 2013 (ALTA 2013)*, pages 16–24.

Higgins, C., McGrath, E., and Moretto, L. (2010). MTurk Crowdsourcing: A Viable Method for Rapid Discovery of Arabic Nicknames? *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, (June):89–92.

Hovy, D., Plank, B., and Søgaard, A. (2014). Experiments with crowdsourced re-annotation of a POS tagging data set. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 377–382.

## Bibliography V

Hsueh, P.-y., Melville, P., and Sindhwani, V. (2009). Data quality from crowdsourcing: a study of annotation selection criteria. *Proceedings of the NAACL HLT 2009 Workshop on Active Learning for Natural Language Processing*, (June):27–35.

Hu, C., Resnik, P., and Kronrod, Y. (2011). The value of monolingual crowdsourcing in a real-world translation scenario: Simulation using Haitian Creole emergency SMS messages. *. . . Machine Translation*, pages 399–404.

Jamison, E. K. and Gurevych, I. (2015). Noise or additional information ? Leveraging crowdsource annotation item agreement for natural language tasks . *Emnlp*, (September):291–297.

Jha, M., Andreas, J., Thadani, K., Rosenthal, S., and Mckeown, K. (2010). Corpus Creation for New Genres: A Crowdsourced Approach to PP Attachment. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, (June):13–20.

Jurgens, D. (2013). Embracing ambiguity: A comparison of annotation methodologies for crowdsourcing word sense labels. In *HLT-NAACL*, pages 556–562.

Kawahara, D., Machida, Y., Shibata, T., Kurohashi, S., Kobayashi, H., and Sassano, M. (2014). Rapid Development of a Corpus with Discourse Annotations using Two-stage Crowdsourcing. *{COLING} 2014, 25th International Conference on Computational Linguistics, Proceedings of the Conference: Technical Papers, August 23-29, 2014, Dublin, Ireland*, pages 269–278.

## Bibliography VI

Khapra, M. M., Ramanathan, A., Kunchukuttan, A., Visweswariah, K., and Bhattacharyya, P. (2014). When transliteration met crowdsourcing: An empirical study of transliteration via crowdsourcing using efficient, non-redundant and fair quality control. In *LREC*, pages 196–202.

Kunchukuttan, A., Chatterjee, R., Roy, S., Mishra, A., and Bhattacharyya, P. (2013). TransDoop: A Map-Reduce based Crowdsourced Translation for Complex Domain. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 175–180.

Kunchukuttan, A., Roy, S., Patel, P., Ladha, K., Gupta, S., Khapra, M., and Bhattacharyya, P. (2012). Experiences in Resource Generation for Machine Translation through Crowdsourcing. *LREC 2012: Eighth international conference on Language Resources and Evaluation, 21-27 May 2012, Istanbul, Turkey*, pages 384–391.

Li, B. and Lee-Urban, S. (2012). Crowdsourcing narrative intelligence. *Advances in Cognitive . . .* , 1:1–18.

Liao, S., Wu, C., and Huerta, J. M. (2011). Evaluating human correction quality for machine translation from crowdsourcing. In *RANLP*, pages 598–603.

Ling, W., Marujo, L., Dyer, C., Black, A. W., and Trancoso, I. (2014). Crowdsourcing High-Quality Parallel Data Extraction from Twitter. *Proceedings of the Ninth Workshop on Statistical Machine Translation*, 13(1):426–436.

Littell, P., Price, K., and Levin, L. (2014). Morphological parsing of Swahili using crowdsourced lexical resources. *Proc. 47th Annual Meeting of the ACL and the 4th IJCNLP of the AFNLP. Suntec*, pages 3333–3339.

## Bibliography VII

Lopez de Lacalle, O. and Agirre, E. (2015). Crowdsourced Word Sense Annotations and Difficult Words and Examples. *Proceedings of the 11th International Conference on Computational Semantics (IWCS 2015)*, (2002):94–100.

Madnani, N. and Chodorow, M. (2011). They Can Help : Using Crowdsourcing to Improve the Evaluation of Grammatical Error Detection Systems. *Computational Linguistics*, pages 508–513.

Madnani, N., Chodorow, M., Cahill, A., Lopez, M., Futagi, Y., and Attali, Y. (2015). Preliminary Experiments on Crowdsourced Evaluation of Feedback Granularity. *Proceedings of the Tenth Workshop on Innovative Use of NLP for Building Educational Applications, 2015*, pages 162–171.

Marujo, L., Gershman, A., Carbonell, J., and Frederking, R. (2013). Supervised Topical Key Phrase Extraction of News Stories using Crowdsourcing, Light Filtering and Co-reference Normalization. *Lrec 2012*, pages 399–403.

Meena, R. and Gustafson, J. (2014). Crowdsourcing Street-level Geographic Information Using a Spoken Dialogue System. *Csc.Kth.Se*, (June):2–11.

Mitchell, M. and Bohus, D. (2014). Crowdsourcing Language Generation Templates for Dialogue Systems. *Inlg*, (June):172–180.

Munro, R., Bethard, S., Kuperman, V., Lai, V. T., Melnick, R., Potts, C., Schnoebelen, T., and Tily, H. (2010). Crowdsourcing and language studies : the new generation of linguistic data. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, (June):122–130.

## Bibliography VIII

Negri, M., Bentivogli, L., and Marchetti, A. (2011). Divide and Conquer : Crowdsourcing the Creation of Cross-Lingual Textual Entailment Corpora. *Computational Linguistics*, pages 670–679.

Ng, J.-P. and Kan, M.-Y. (2012). Improved Temporal Relation Classification using Dependency Parses and Selective Crowdsourced Annotations. *Proceedings of COLING 2012*, (December 2012):2109–2124.

Nguyen, D., Trieschnigg, D., Doğruöz, A. S., Gravel, R., Theune, M., Meder, T., and de Jong, F. (2014). Why Gender and Age Prediction from Tweets is Hard: Lessons from a Crowdsourcing Experiment. *Proceedings of the 25th International Conference on Computational Linguistics, COLING 2014, Dublin, Ireland*, pages 1950–1961.

Novotney, S. and Callison-Burch, C. (2010). Shared Task: Crowdsourced Accessibility Elicitation of Wikipedia Articles. *Proceedings of the NAACL HLT 2010 Workshop on Creating Speech and Language Data with Amazon's Mechanical Turk*, (June):41–44.

Post, M., Callison-Burch, C., and Osborne, M. (2012). Constructing parallel corpora for six indian languages via crowdsourcing. *Wmt-2012*, pages 401–409.

Potthast, M., Hagen, M., Völske, M., and Stein, B. (2013). Crowdsourcing Interaction Logs to Understand Text Reuse from the Web. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, 1(August):1212–1221.

# Bibliography IX

Prabhakaran, V., Bloodgood, M., Diab, M., Dorr, B., Levin, L., Piatko, C. D., Rambow, O., and Van Durme, B. (2012). Statistical modality tagging from rule-based annotations and crowdsourcing. *Proceedings of the Workshop on Extra-Propositional Aspects of Meaning in Computational Linguistics*, (July):57–64.

Ramanath, R., Choudhury, M., Bali, K., and Saha Roy, R. (2013). Crowd Prefers the Middle Path: A New IAA Metric for Crowdsourcing Reveals Turker Biases in Query Segmentation. *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1713–1722.

Rumshisky, A. (2011). Crowdsourcing Word Sense Definition. *Computational Linguistics*, (June):74–81.

Sabou, M., Bontcheva, K., Derczynski, L., and Scharl, A. (2014). Corpus Annotation through Crowdsourcing : Towards Best Practice Guidelines. *Proceedings of the Ninth International Conference on Language Resources and Evaluation*, (2010):859–866.

Sabou, M., Bontcheva, K., and Scharl, A. (2012). Crowdsourcing research opportunities. *Proceedings of the 12th International Conference on Knowledge Management and Knowledge Technologies - i-KNOW '12*, page 1.

Sayeed, A. B., Meyer, T. J., Nguyen, H. C., Buzek, O., and Weinberg, A. (2010). Crowdsourcing the evaluation of a domain-adapted named entity recognition system. *Computational Linguistics*, (June):345–348.

# Bibliography X

Sayeed, A. B., Rusk, B., Petrov, M., Nguyen, H. C., Meyer, T. J., and Weinberg, A. (2011). Crowdsourcing syntactic relatedness judgements for opinion mining in the study of information technology adoption. In *Proceedings of the 5th ACL-HLT Workshop on Language Technology for Cultural Heritage, Social Sciences, and Humanities*, pages 69–77. Association for Computational Linguistics.

Schmidt, M., Müller, M., Wagner, M., Stüker, S., Waibel, A., Hofmann, H., and Werner, S. (2015). Evaluation of crowdsourced user input data for spoken dialog systems. In *16th Annual Meeting of the Special Interest Group on Discourse and Dialogue*, page 427.

Snow, R., O'Connor, B., Jurafsky, D., and Ng, A. Y. (2008). Cheap and fast—but is it good?: evaluating non-expert annotations for natural language tasks. In *Proceedings of the conference on empirical methods in natural language processing*, pages 254–263. Association for Computational Linguistics.

Søgaard, A., Martinez, H., Elming, J., and Johannsen, A. (2013). Using crowdsourcing to get representations based on regular expressions. *Emnlp*, (October):1476–1480.

Takabatake, Y. (2015). Classification and Acquisition of Contradictory Event Pairs using Crowdsourcing. pages 99–107.

Talukdar, P. and Cohen, W. (2012). Crowdsourced Comprehension: Predicting Prerequisite Structure in Wikipedia. *Proceedings of the Second Workshop on Building Educational Applications Using NLP*, pages 307–315.

## Bibliography XI

Trisedya, B. D. and Manurung, R. (2012). A graf-compliant indonesian speech recognition web service on the language grid for transcription crowdsourcing. In *Proceedings of the Sixth Linguistic Annotation Workshop*, pages 67–74. Association for Computational Linguistics.

Tschirsich, M. and Hintz, G. (2013). Leveraging Crowdsourcing for Paraphrase Recognition. *Law Vii & Id*, pages 205–213.

Vertanen, K. and Kristensson, P. O. (2011). The Imagination of Crowds: Conversational AAC Language Modeling using crowdsourcing and large data sets. *Computational Linguistics*, pages 700–711.

Voyer, R., Nygaard, V., Fitzgerald, W., and Copperman, H. (2010). A Hybrid Model for Annotating Named Entity Training Corpora. *Proceedings of the Fourth Linguistic Annotation Workshop*, (July):243–246.

Wang, A., Hoang, C. D. V., and Kan, M. Y. (2013). Perspectives on crowdsourcing annotations for natural language processing. *Language Resources and Evaluation*, 47(1):9–31.

Wang, S., Huang, C.-R., Yao, Y., and Chan, A. (2014a). Building a Semantic Transparency Dataset of Chinese Nominal Compounds: A Practice of Crowdsourcing Methodology. *Proceedings of Workshop on Lexical and Grammatical Resources for Language Processing*, pages 147–156.

Wang, S., Huang, C.-R., Yao, Y., and Chan, A. (2014b). Exploring Mental Lexicon in an Efficient and EconomicWay: Crowdsourcing Method for Linguistic Experiments. *Coling 2014*, pages 7–14.

# Bibliography XII

Wang, S., Huang, C.-R., Yao, Y., and Chan, A. (2015). Create a Manual Chinese Word Segmentation Dataset Using Crowdsourcing Method. *Acl-Ijcnlp 2015*, pages 7–14.

Wang, W. Y., Bohus, D., Kamar, E., and Horvitz, E. (2012). Crowdsourcing the acquisition of natural language corpora: Methods and observations. In *Spoken Language Technology Workshop (SLT), 2012 IEEE*, pages 73–78. IEEE.

Wray, S., Mubarak, H., and Ali, A. (2015). Best Practices for Crowdsourcing Dialectal Arabic Speech Transcription. pages 99–107.

Yan, R., Gao, M., Pavlick, E., and Callison-Burch, C. (2014). Are Two Heads Better than One? Crowdsourced Translation via a Two-Step Collaboration of Non-Professional Translators and Editors. *Proceedings of the 52nd Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 1134–1144.

Zaidan, O. F. and Callison-Burch, C. (2011). Crowdsourcing Translation: Professional Quality from Non-Professionals. *Acl*, pages 1220–1229.

Zbib, R., Markiewicz, G., Matsoukas, S., Schwartz, R., and Makhoul, J. (2013). Systematic Comparison of Professional and Crowdsourced Reference Translations for Machine Translation. *Proceedings of the 2013 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies*, (June):612–616.

Zeichner, N., Berant, J., and Dagan, I. (2012). Crowdsourcing Inference-Rule Evaluation. (July):156–160.