The MIT Press

Journals

Books   Journals   Digital   Resources   About   Contact

Home | Computational Linguistics | List Article navigation of Issues | Volume 37 , No. 4 | What Determines Inter-Coder Agreement in Manual Annotations? A Meta-Analytic Investigation

Journal Resources

Editorial Info
Abstracting and Indexing
Release Schedule
Advertising Info

Author Resources

Submission Guidelines

What Determines Inter-Coder Agreement in Manual Annotations? A Meta-Analytic Investigation

Petra Saskia Bayerl and Karsten Ingmar Paul

◎ **Download Options** 〉

Abstract   Authors

Recent discussions of annotator agreement have mostly centered around its calculation and

interpretation, and the correct choice of indices. Although these discussions are important, they only consider the "back-end" of the story, namely, what to do once the data are collected. Just as important in our opinion is to know how agreement is reached in the first place and what factors influence coder agreement as part of the annotation process or setting, as this knowledge can provide concrete guidelines for the planning and set-up of annotation projects. To investigate whether there are factors that consistently impact annotator agreement we conducted a meta-analytic investigation of annotation studies reporting agreement percentages. Our meta-analysis synthesized factors reported in 96 annotation studies from three domains (word-sense disambiguation, prosodic transcriptions, and phonetic transcriptions) and was based on a total of 346 agreement indices. Our analysis identified seven factors that influence reported agreement values: annotation domain, number of categories in a coding scheme, number of annotators in a project, whether annotators received training, the intensity of annotator training, the annotation purpose, and the method used for the calculation of percentage agreements. Based on our results we develop practical recommendations for the assessment, interpretation, calculation, and reporting of coder agreement. We also briefly discuss theoretical implications for the concept of annotation quality.

Forthcoming

Most Read

See More

🔑 Lexicon-Based Methods for Sentiment Analysis (14019 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

🔑 Computational Linguistics and Deep Learning (10513 times)
Christopher D. Manning
Computational Linguistics
Volume: 41, Issue: 4, pp. 701-707

🔑 Near-Synonymy and Lexical Choice (3658 times)
Philip Edmonds et al.
Computational Linguistics
Volume: 28, Issue: 2, pp. 105-144

(Note that the Most Read numbers are based on the number of full text downloads over the last 12 months.)

readers across the globe. All articles are published under a CC BY-NC-ND 4.0 license. For more information on allowed uses, please view the CC license. Support OA at MITP

Most Cited

See More

🔗 **Lexicon-Based Methods for Sentiment Analysis** (436 times)
Maite Taboada et al.
Computational Linguistics
Volume: 37, Issue: 2, pp. 267-307

🔗 **A Systematic Comparison of Various Statistical Alignment Models** (174 times)
Franz Josef Och et al.
Computational Linguistics
Volume: 29, Issue: 1, pp. 19-51

🔗 **Opinion Word Expansion and Target Extraction through Double Propagation** (147 times)
Guang Qiu et al.
Computational Linguistics
Volume: 37, Issue: 1, pp. 9-27

(Note that the Most Cited numbers are based on Crossref's Cited-by service and reflect citation information for the past 24 months.)
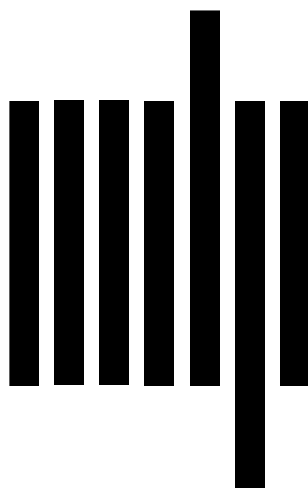
◎ Download ⟩ Options

Favorite ♡

Sign up for Alerts 🔔

Download Citation ↓

RSS TOC 📶

RSS Citation 📶

Submit your article

Support OA at MITP 🔓

Journals

Books

US

One Rogers Street Cambridge MA 02142-1209

Terms & Conditions

UK

Suite 2, 1 Duchess Street London, W1W 6AN, UK

Privacy Statement

Connect

[f] [y] [G+] [p] [◎] [▶]

Contact Us

U.S. Patent and
Trademark Office.
Site Help

U.S. Patent and
Trademark Office.
Site Help