MinIE: Minimizing Facts in Open Information Extraction
Kiril Gashteovski¹, Rainer Gemulla¹, Luciano del Corro²
¹University of Mannheim, Mannheim, Germany ²Max-Planck-Institut für Informatik, Saarbrücken, Germany

MinIE: Annotation and Minimization of Facts
Input sentence: “The Joker believes that Batman was not actually born in foggy Gotham City.”
Output triple: (“Batman”; “was not actually born in”; “the foggy Gotham City”)

MinIE’s Semantic Annotations: Factuality
Polarity: is the fact positive (+) or negative (–)?
Input: “Superman does live in Metropolis.” … (1)
“Superman does not live in Metropolis.” … (2)
“Superman probably does live in Metropolis.” … (3)
“Superman probably does not live in Metropolis.” … (4)
Output: (“Superman”; “does live in”; “Metropolis”)
(+, CT) (–, CT) (+, PS) (–, PS)

MinIE’s Semantic Annotations: Attribution
Attribution (A) consists of:
• attribution phrase: the supplier of information
• attribution factuality: polarity + modality about the supplier of information
Input sentences:
“Orly Taitz believes that B. Obama was not born in U.S.” … (5)
“Joshua Wisch said that Barack Obama was born in U.S.” … (6)
Output facts:
(“Barack Obama”; “was born in”; “U.S.”)

MinIE’s Semantic Annotations: Quantities
A quantity is a phrase that expresses an amount of something.
Example:
9,890 habitats, all habitats, almost about 10,000 habitats
(Q=9,890 / Q=all / Q=almost about 10,000)

Common Problems with OIE
Relations can be uninformative
• E.g. “make” has 49 meanings in WordNet
• v.s. the more informative “make deal with”

Arguments and relations can be overly specific
• overly specific relation: “make a very good deal with”
• overly specific argument: “the extraordinary James Watt”

Lack of context of a triple
• E.g. (“North Korea”; “attack”; “Guam”)
• Is this a certainty or merely a possibility?
• According to whom?

MinIE: OIE System Providing Useful Extractions
• Represents contextual information with semantic annotations
• Identifies and removing words that are considered overly specific
• High precision/recall
• Shorter, semantically enriched extractions

MinIE’s OIE System Providing Useful Extractions
Experiments
Input sentence: “The big celebration on the campus lasted for 2 days.”
Output: (“The big celebration on the campus”; “lasted for”; “Q₁ days”)
MinIE-C (Complete) 8.3 ± 4.9 0.75/0.75 727/635
MinIE-S (Safe) 7.2 ± 4.2 0.75/0.74 600/602
MinIE-D (Dictionary) 7.0 ± 4.1 0.74/0.73 681/593
MinIE-A (Aggressive) 4.7 ± 1.9 0.59/0.61 505/474

N = New York Times (sample of 200 sentences); W = Wikipedia (sample of 200 sentences)

Resources
All data, labels, and source code are available at:
http://dws.informatik.uni-mannheim.de/en/resources/software/minie/