

# WWW 2020

## *Domain-Specific Automatic Scholar Profiling Based on Wikipedia*

Ziang Chuai, Qian Geng, Jian Jin

DEPARTMENT OF INFORMATION MANAGEMENT, BEIJING NORMAL UNIVERSITY



# Fine-Grained Entity Typing

- In June 1903, supervised by Gabriel Lippmann, **Marie Curie** was awarded her doctorate from the **University of Paris**.
- She was also the first woman to become a professor at the **University of Paris**, and in 1995 became the first woman to be entombed on her own merits in the Panthéon in Paris.



c. 1920

**Born** Maria Salomea Skłodowska  
7 November 1867  
Warsaw, Congress Poland,  
Russian Empire<sup>[1]</sup>

**Died** 4 July 1934 (aged 66)  
Passy, Haute-Savoie, France

**Cause of death** [Aplastic anemia](#) from exposure to radiation

**Citizenship** Poland (by birth)  
France (by marriage)

**Alma mater** [University of Paris](#)  
[ESPCI](#)<sup>[2]</sup>

**Known for** [Radioactivity](#)  
[Polonium](#)

CRF

TransP

k-means

## TransP:

- An embedding method based on TransE;
- A new penalty scheme based on local context.

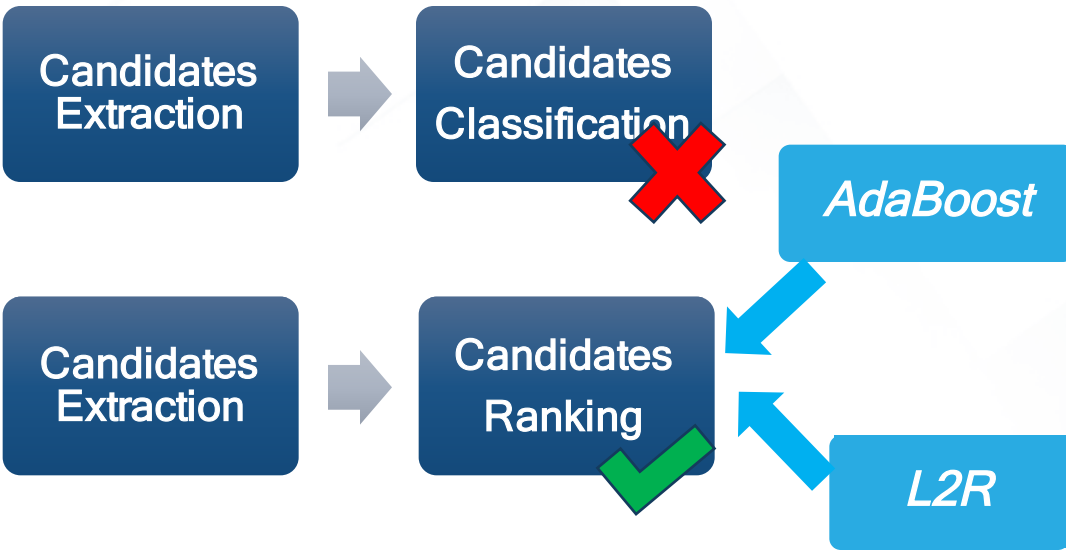
Models	Precision	Recall	F1
<b>TransP</b>	<b>0.85</b>	<b>0.73</b>	<b>0.78</b>
Rescal	0.49	0.49	0.49
TransE	0.56	0.57	0.56
TransH	0.45	0.44	0.44
TransR	0.49	0.48	0.48
TransD	0.56	0.56	0.56

# Keyword Extraction

## Bibliography [\[edit\]](#)

- Salton, [Automatic Information Organization and Retrieval](#), 1968.
- Gerard Salton (1975). *A Theory of Indexing*. Society for Industrial and Applied Mathematics. p. 56.
- --- and Michael J. McGill, *Introduction to modern Information Retrieval*, 1983. ISBN 0-07-054484-0
- Gerard Salton (1989). *Automatic Text Processing*. Addison-Wesley Publishing Company. p. 530. ISBN 978-0-201-12227-5.
- [Gerard Salton](#) at DBLP Bibliography Server
- G. Salton, A. Wong, and C. S. Yang (1975), "[A Vector Space Model for Automatic Indexing](#)," *Communications of the ACM*, vol. 18, nr. 11, pages 613–620. (Article in which a vector space model was presented)

Concepts to be extracted



Models	Precision	Recall	F1	MAP
<b>Our method</b>	<b>0.52</b>	<b>0.89</b>	<b>0.65</b>	<b>0.37</b>
He's method	0.31	0.51	0.39	0.20
KEA	0.18	0.70	0.29	0.17
RAKE	0.46	0.78	0.58	0.33
TextRank	0.20	0.33	0.25	0.15