Computational approaches to semantic change detection
Day 5, Part 2
Explainable semantic change modeling via definition generation

Andrey Kutuzov, Lidia Pivovarova

University of Oslo, University of Helsinki
ESSLLI'2023

Contextualized definitions as word representations

A retelling of [Giulianelli et al., 2023] (https://aclanthology.org/2023.acl-long.176/)

The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings

Austin C. Kozlowski,[a] Matt Taddy,[b] and James A. Evans[a,c]

The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings

Austin C. Kozlowski,[a] Matt Taddy,[b] and James A. Evans[a,c]

Annual Review of Linguistics

Semantic Structure in Deep Learning

Ellie Pavlick

A computational analysis of crosslinguistic regularity in semantic change

Olivia Fugikawa[1*†], Oliver Hayman[2*†], Raymond Liu[3*†], Lei Yu[4], Thomas Brochhagen[5] and Yang Xu[6*]

The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings

Austin C. Kozlowski,[a] Matt Taddy,[b] and James A. Evans[a,c]

Annual Review of Linguistics

Semantic Structure in Deep Learning

Ellie Pavlick

A computational analysis of crosslinguistic regularity in semantic change

Olivia Fugikawa[1*†], Oliver Hayman[2*†], Raymond Liu[3*†], Lei Yu[4], Thomas Brochhagen[5] and Yang Xu[6*]

Follow the leader: Documents on the leading edge of semantic change get more citations

Sandeep Soni[1] | Kristina Lerman[2] | Jacob Eisenstein[3]

The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings

Austin C. Kozlowski,[a] Matt Taddy,[b] and James A. Evans[a,c]

Annual Review of Linguistics

Semantic Structure in Deep Learning

Ellie Pavlick

A computational analysis of crosslinguistic regularity in semantic change

Olivia Fugikawa[1*†], Oliver Hayman[2*†], Raymond Liu[3*†], Lei Yu[4], Thomas Brochhagen[5] and Yang Xu[6*]

Follow the leader: Documents on the leading edge of semantic change get more citations

Sandeep Soni[1] | Kristina Lerman[2] | Jacob Eisenstein[3]

On the Impact of Temporal Concept Drift on Model Explanations

Zhixue Zhao George Chrysostomou Kalina Bontcheva Nikolaos Aletras

The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings

Austin C. Kozlowski,[a] Matt Taddy,[b] and James A. Evans[a,c]

Word embeddings quantify 100 years of gender and ethnic stereotypes

Nikhil Garg[a,1], Londa Schiebinger[b], Dan Jurafsky[c,d], and James Zou[e,f,1]

Annual Review of Linguistics

Semantic Structure in Deep Learning

Ellie Pavlick

Disentangling the cultural evolution of ancient China: a digital humanities perspective

Siyu Duan[1,2], Jun Wang[1,2,3], Hao Yang[2,3] & Qi Su[2,3,4]

A computational analysis of crosslinguistic regularity in semantic change

Olivia Fugikawa[1*†], Oliver Hayman[2*†], Raymond Liu[3*†], Lei Yu[4], Thomas Brochhagen[5] and Yang Xu[6*]

Characterizing English Variation across Social Media Communities with BERT

Li Lucy and David Bamman

What about Grammar? Using BERT Embeddings to Explore Functional-Semantic Shifts of Semi-Lexical and Grammatical Constructions

Lauren Fonteyn

Slangvolution: A Causal Analysis of Semantic Change and Frequency Dynamics in Slang

Daphna Keidar[*] Andreas Opedal[*] Zhijing Jin Mrinmaya Sachan

Follow the leader: Documents on the leading edge of semantic change get more citations

Sandeep Soni[1] | Kristina Lerman[2] | Jacob Eisenstein[3]

On the Impact of Temporal Concept Drift on Model Explanations

Zhixue Zhao George Chrysostomou Kalina Bontcheva Nikolaos Aletras

The Geometry of Culture:
Analyzing the Mea
of Class throug
Embeddings

Austin C. Kozlowski,[a]
and James A. Evans[a,c]

Annual Review of Linguistics

tic Structure in
earning

Disentangling the cultu
China: a digital humani

Siyu Duan[1,2], Jun Wang[1,2,3], Hao Yang[2,3] & Qi Su

lish Variation across
munities with BERT

avid Bamman

What about Gran
Explore Functiona
Grammatical Con

Lauren Fonteyn

tic Change
g

Mrinmaya Sachan

Follow the leade
semantic change g

Sandeep Soni[1] | Kristina Lerman[2] | Jacob Eisenstein[3]

odel Explanations

Zhixue Zhao George Chrysostomou Kalina Bontcheva Nikolaos Aletras

Historical linguistics

Psycholinguistics

Cognitive science

Sociology

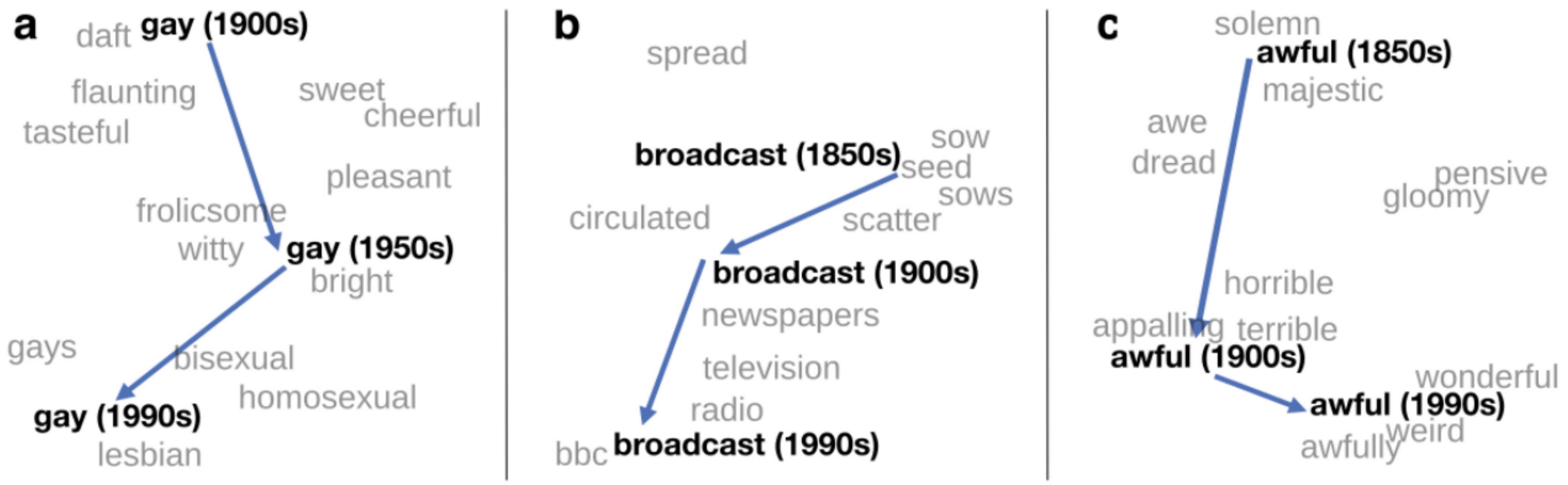
Gender and cultural studies

Political science

Natural language processing

Word embeddings

Static



Hamilton, Leskovec, Jurafsky. ACL 2016.

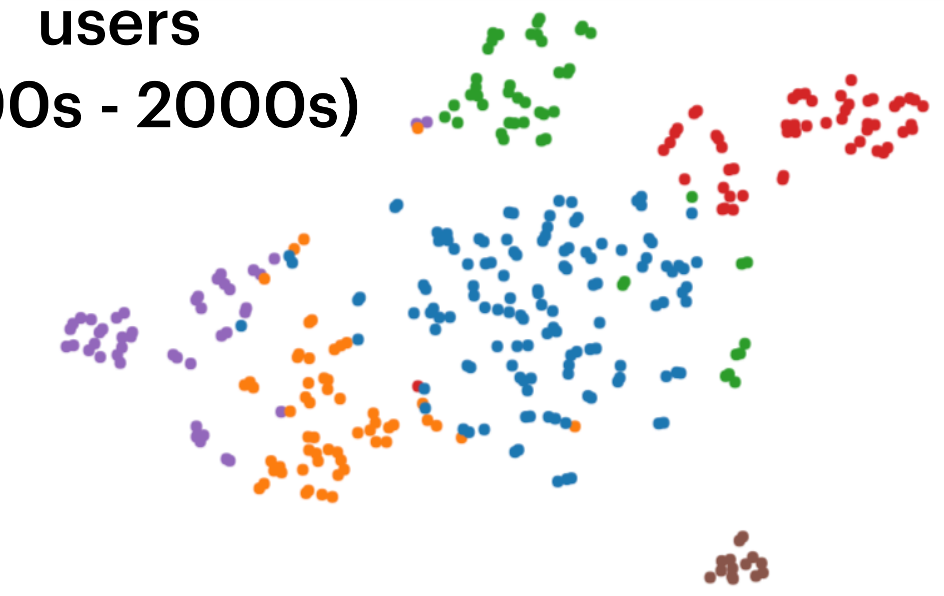
Word embeddings

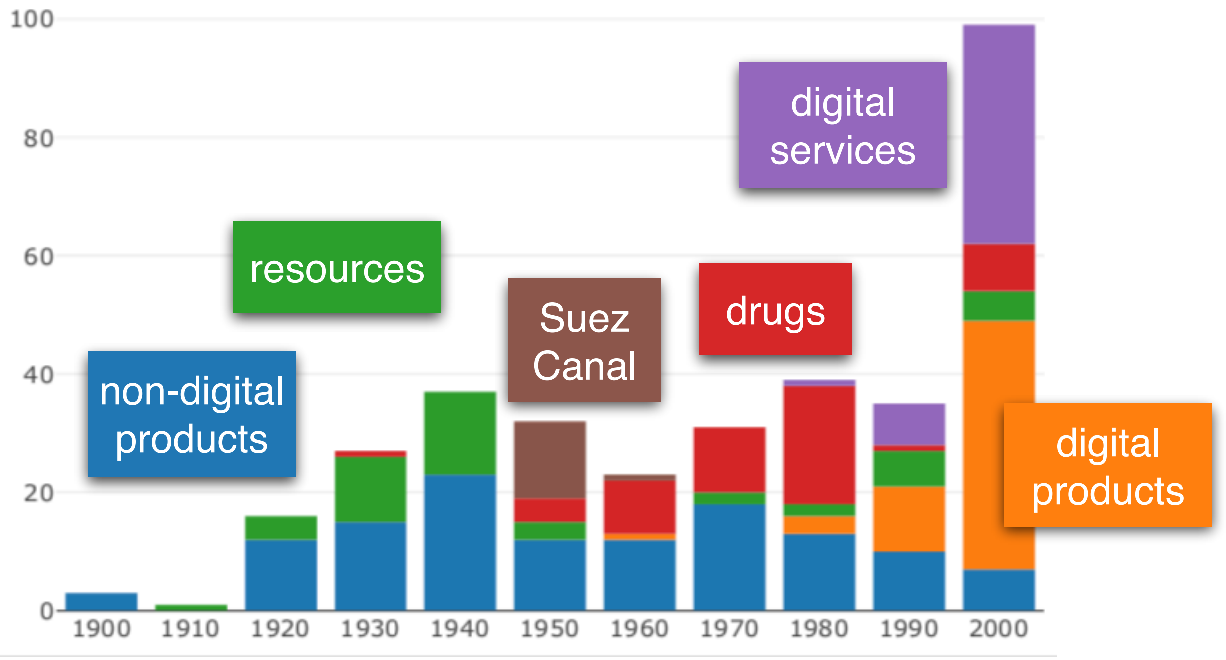
Static

Hamilton, Leskovec, Jurafsky. ACL 2016.

Contextualised

users
(1990s - 2000s)

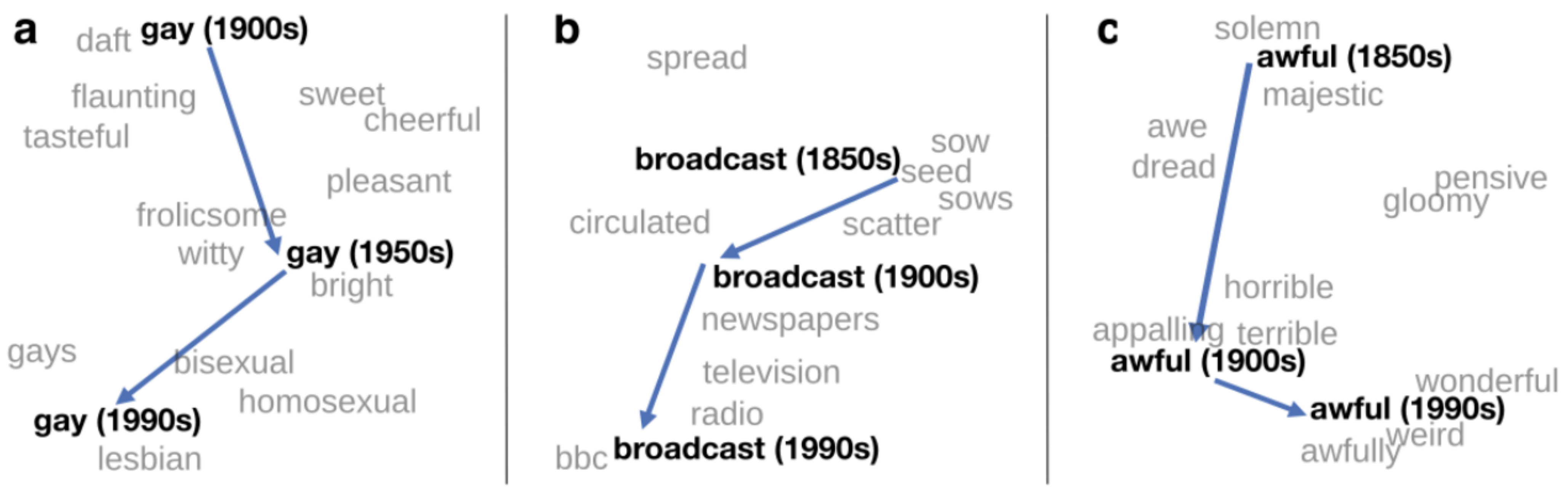




Giulianelli, Del Tredici, Fernández. ACL 2020.

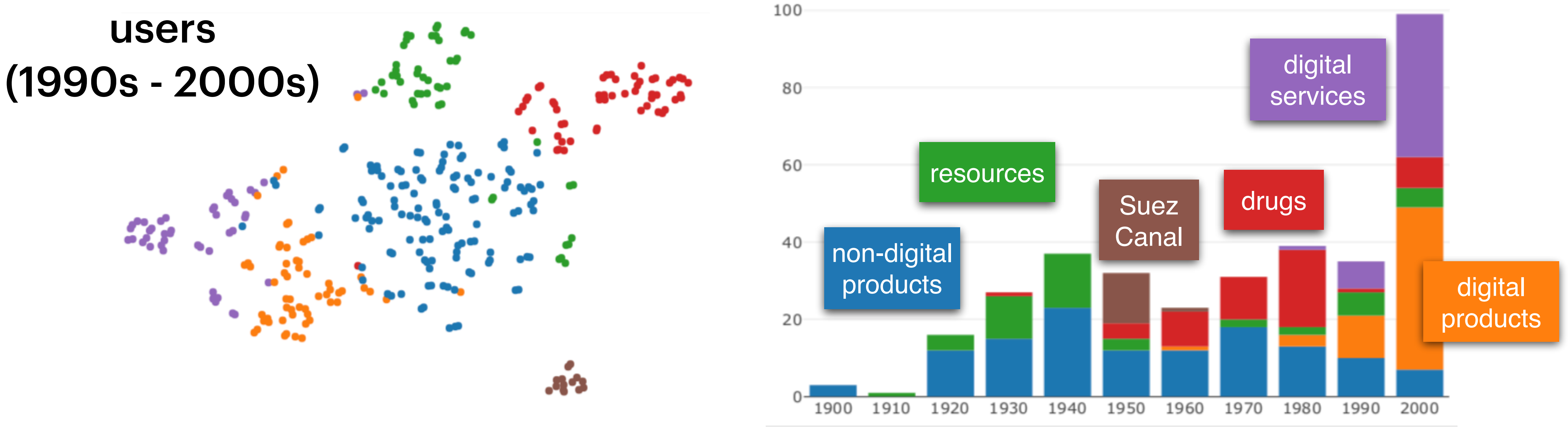
Word embeddings

Static



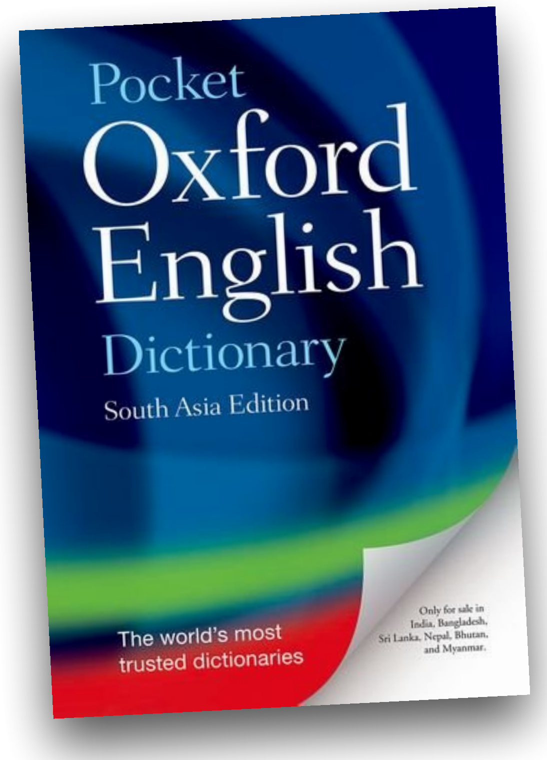
Hamilton, Leskovec, Jurafsky. ACL 2016.

Contextualised



Giulianelli, Del Tredici, Fernández. ACL 2020.

Definitions



record

noun

/ˈrɛkɔːd/

1. a thing constituting a piece of evidence about the past, especially an account kept in writing or some other permanent form.
   "you should keep a written record"
   Similar: account(s) document(s) documentation data file(s) dossier(s)
2. the sum of the past achievements or performance of a person, organization, or thing.
   "the safety record at the airport is first class"
   Similar: previous conduct/performance track record career to date history past

Origin

LATIN	LATIN	OLD FRENCH	OLD FRENCH	
cor cord- heart	recordari remember	recorder bring to remembrance	record remembrance	record *Middle English*

Word embeddings

Definitions

Hamilton,

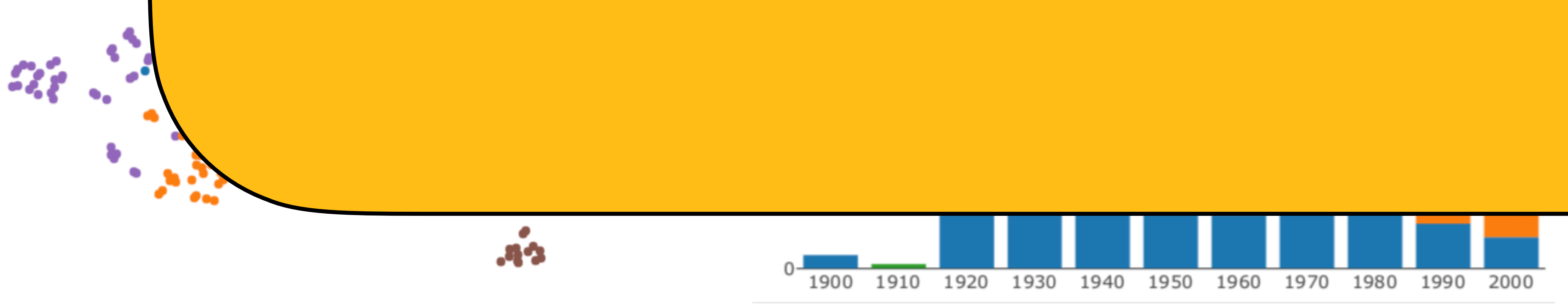
The best of both worlds?

human-readable

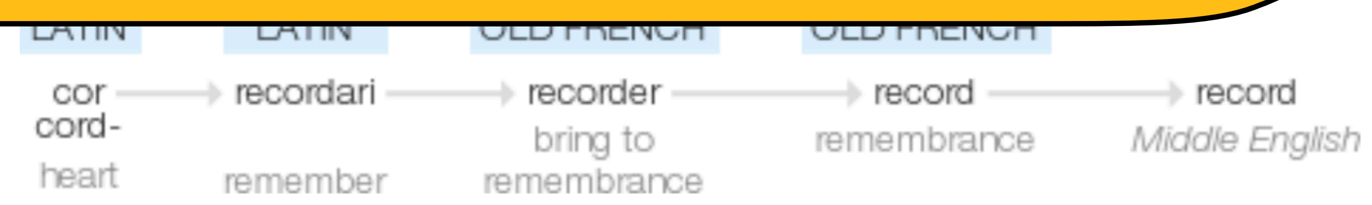
quantitative

data-driven and automatic

users
(1990s - 200



Giulianelli, Del Tredici, Fernández. ACL 2020.



Contextualised definitions as word representations

“you should keep a written ***record****”*

Contextualised definitions as word representations

"you should keep a written ***record****"*

↓

↓

A document or other means of providing information about past events.

Contextualised definitions as word representations

*"you should keep a written **record**"*

A document or other means of providing information about past events.

*"she held the world **record** for ten years"*

Contextualised definitions as word representations

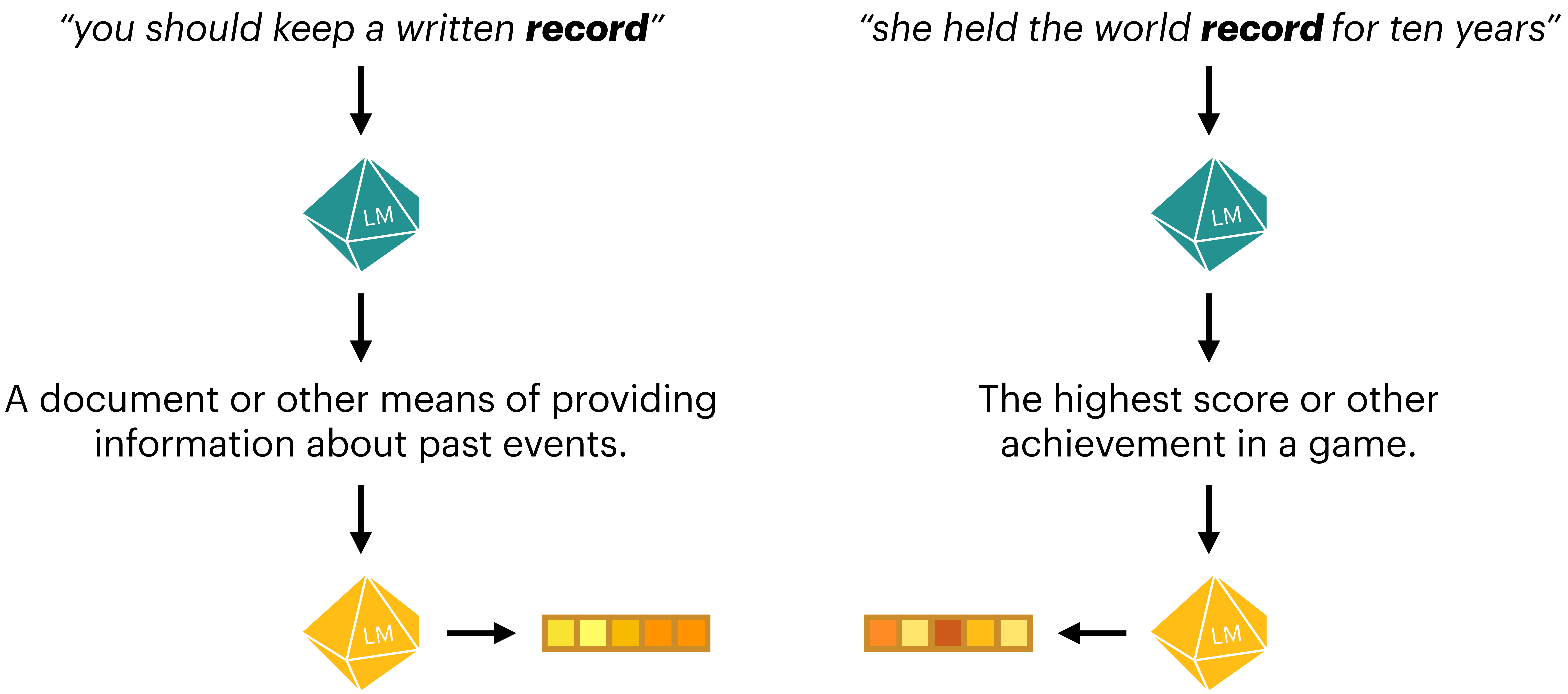
*"you should keep a written **record**"*

A document or other means of providing information about past events.

*"she held the world **record** for ten years"*

The highest score or other achievement in a game.

Contextualised definitions as word representations



Contextualised definitions as word representations

*"you should keep a written **record**"*

FLAN-T5

A document or other means of providing information about past events.

SBERT

*"she held the world **record** for ten years"*

FLAN-T5

Chung et al., 2022

The highest score or other achievement in a game.

SBERT

Reimers and Gurevych, 2019

Definition generation

usage examples

instructions →

teacher forcing →

usage-specific definitions

“she held the world record for ten years”

+ “what is the definition of record?”

The best performance or most remarkable event of its kind.

3 datasets of definitions

WordNet (Ishiwatari et al., 2019)
Oxford (Gadetsky et al., 2018)
CoDWoE (Mickus et al., 2022)

Definition generation

usage examples

instructions →

teacher forcing →

usage-specific definitions

“she held the world record for ten years”

\+ “what is the definition of record?”

The best performance or most remarkable event of its kind.

3 datasets of definitions

WordNet (Ishiwatari et al., 2019)
Oxford (Gadetsky et al., 2018)
CoDWoE (Mickus et al., 2022)

8 task verbalisations

“define w”
“define the word w”
“give the definition of w”
“what is the definition of w?”
…

Definition generation

usage examples

“she held the world record for ten years”

+ “what is the definition of record?”

instructions →

teacher forcing →

usage-specific definitions

The best performance or most remarkable event of its kind.

3 datasets of definitions

WordNet (Ishiwatari et al., 2019)
Oxford (Gadetsky et al., 2018)
CoDWoE (Mickus et al., 2022)

8 task verbalisations

“define w”
“define the word w”
“give the definition of w”
“what is the definition of w?”
...

3 variants of Flan-T5

Base
Large
XL
+ vanilla T5 (Raffel et al., 2020)

Definition generation

usage examples

usage-specific definitions

"she held the world record for ten years"

+ "what is the definition of record?"

The best performance or most remarkable event of its kind.

		WordNet			***Oxford***		
Model	Test	BLEU	ROUGE-L	BERT-F1	BLEU	ROUGE-L	BERT-F1
[Huang et al., 2021]	*Unknown*	32.72	-	-	**26.52**	-	-
Flan-T5 XL	Zero-shot (task shift)	2.70	12.72	86.72	2.88	16.20	86.52
Flan-T5 XL	In-distribution	11.49	28.96	88.90	16.61	36.27	89.40
Flan-T5 XL	Hard domain shift	29.55	48.17	91.39	8.37	25.06	87.56
Flan-T5 XL	Soft domain shift	**32.81**	**52.21**	**92.16**	18.69	**38.72**	**89.75**

Definition generation

usage examples

usage-specific definitions

"she held the world record for ten years"

+ "what is the definition of record?"

The best performance or most remarkable event of its kind.

		WordNet			Oxford		
Model	Test	BLEU	ROUGE-L	BERT-F1	BLEU	ROUGE-L	BERT-F1
[Huang et al., 2021]	*Unknown*	32.72	-	-	**26.52**	-	-
Flan-T5 XL	Zero-shot (task shift)	2.70	12.72	86.72	2.88	16.20	86.52
Flan-T5 XL	In-distribution	11.49	28.96	88.90	16.61	36.27	89.40
Flan-T5 XL	Hard domain shift	29.55	48.17	91.39	8.37	25.06	87.56
Flan-T5 XL	Soft domain shift	**32.81**	**52.21**	**92.16**	18.69	**38.72**	**89.75**

Definition generation

“she held the world record for ten years”

+ “what is the definition of record?”

The best performance or most remarkable event of its kind.

		WordNet			*Oxford*		
Model	Test	BLEU	ROUGE-L	BERT-F1	BLEU	ROUGE-L	BERT-F1
[Huang et al., 2021]	*Unknown*	32.72	-	-	**26.52**	-	-
Flan-T5 XL	Zero-shot (task shift)	2.70	12.72	86.72	2.88	16.20	86.52
Flan-T5 XL	In-distribution	11.49	28.96	88.90	16.61	36.27	89.40
Flan-T5 XL	Hard domain shift	29.55	48.17	91.39	8.37	25.06	87.56
Flan-T5 XL	Soft domain shift	**32.81**	**52.21**	**92.16**	18.69	**38.72**	**89.75**

Definition generation

But...

Are definitions good

meaning representations?

“she he

+ “wl

ons

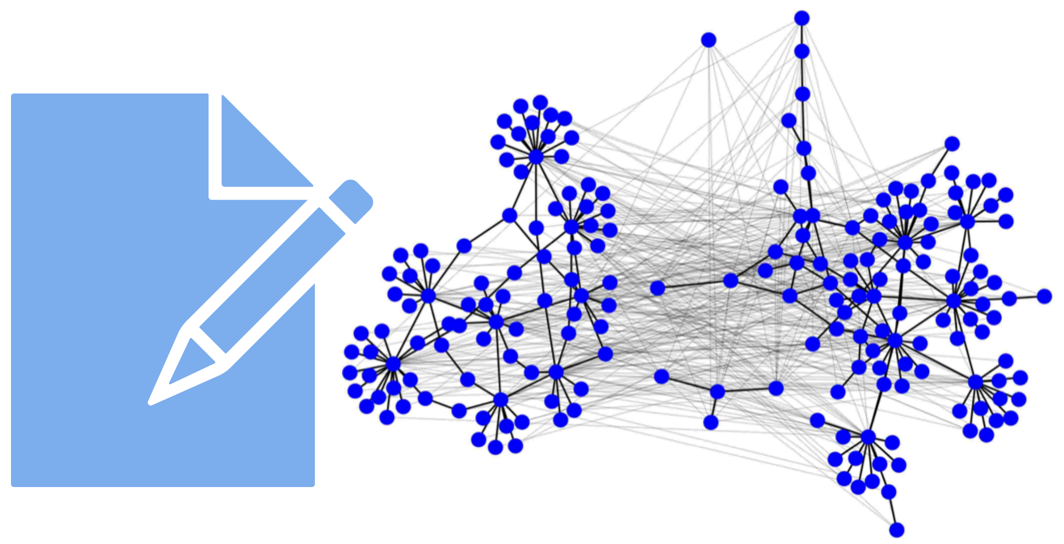
ost

d.

Flan-T5 XL	Hard domain shift	29.55	48.17	91.39	8.37	25.06	87.56
Flan-T5 XL	Soft domain shift	**32.81**	**52.21**	**92.16**	18.69	**38.72**	**89.75**

Word-in-context similarity

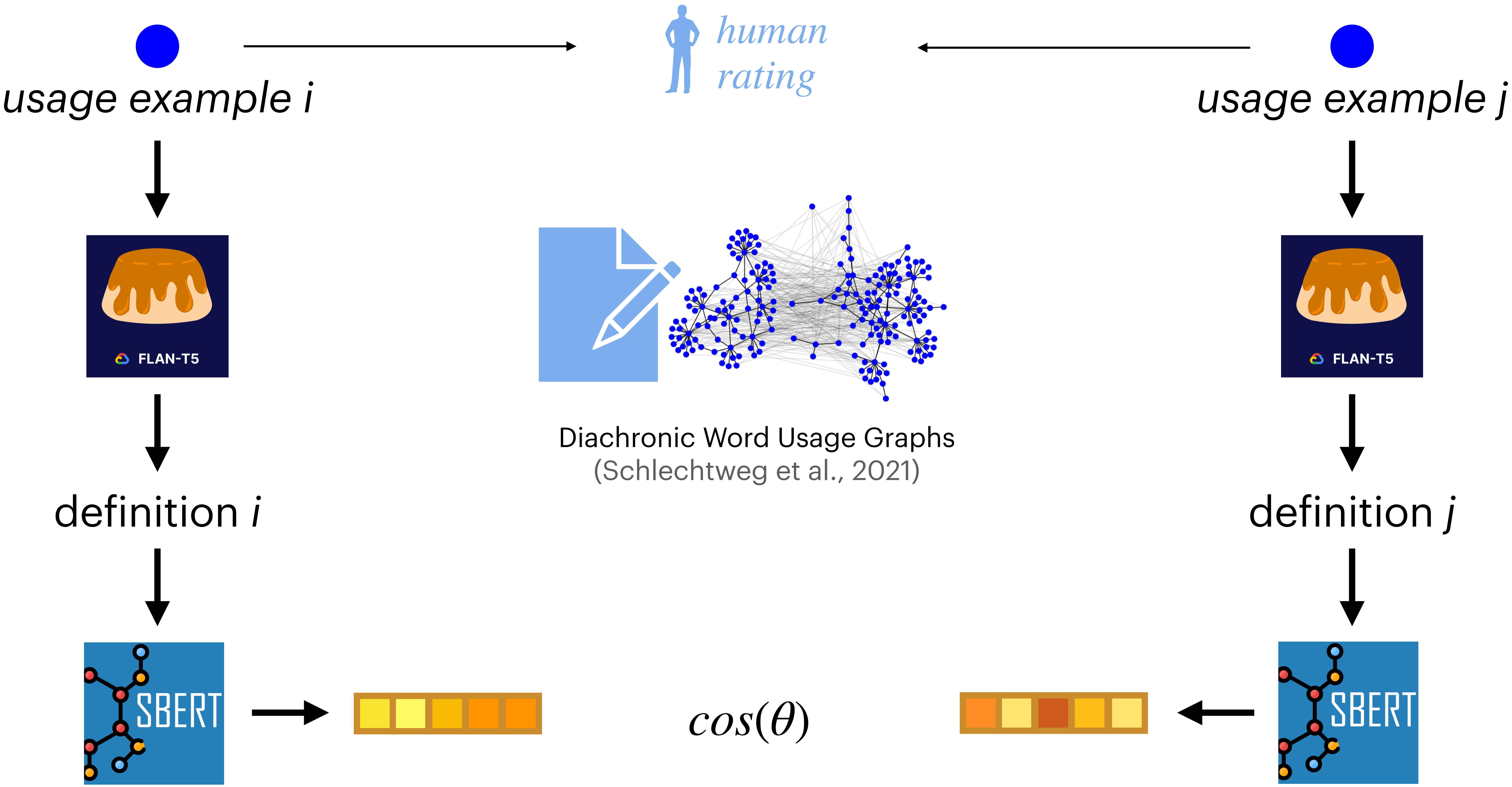
Correlation with human pairwise similarity judgements



Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)

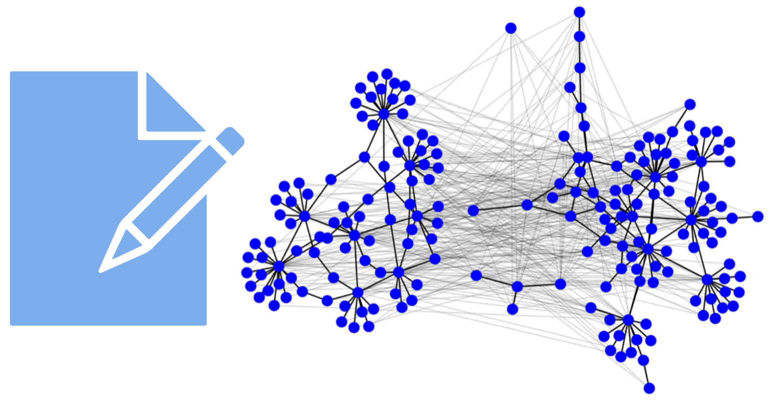
Word-in-context similarity

Correlation with human pairwise similarity judgements

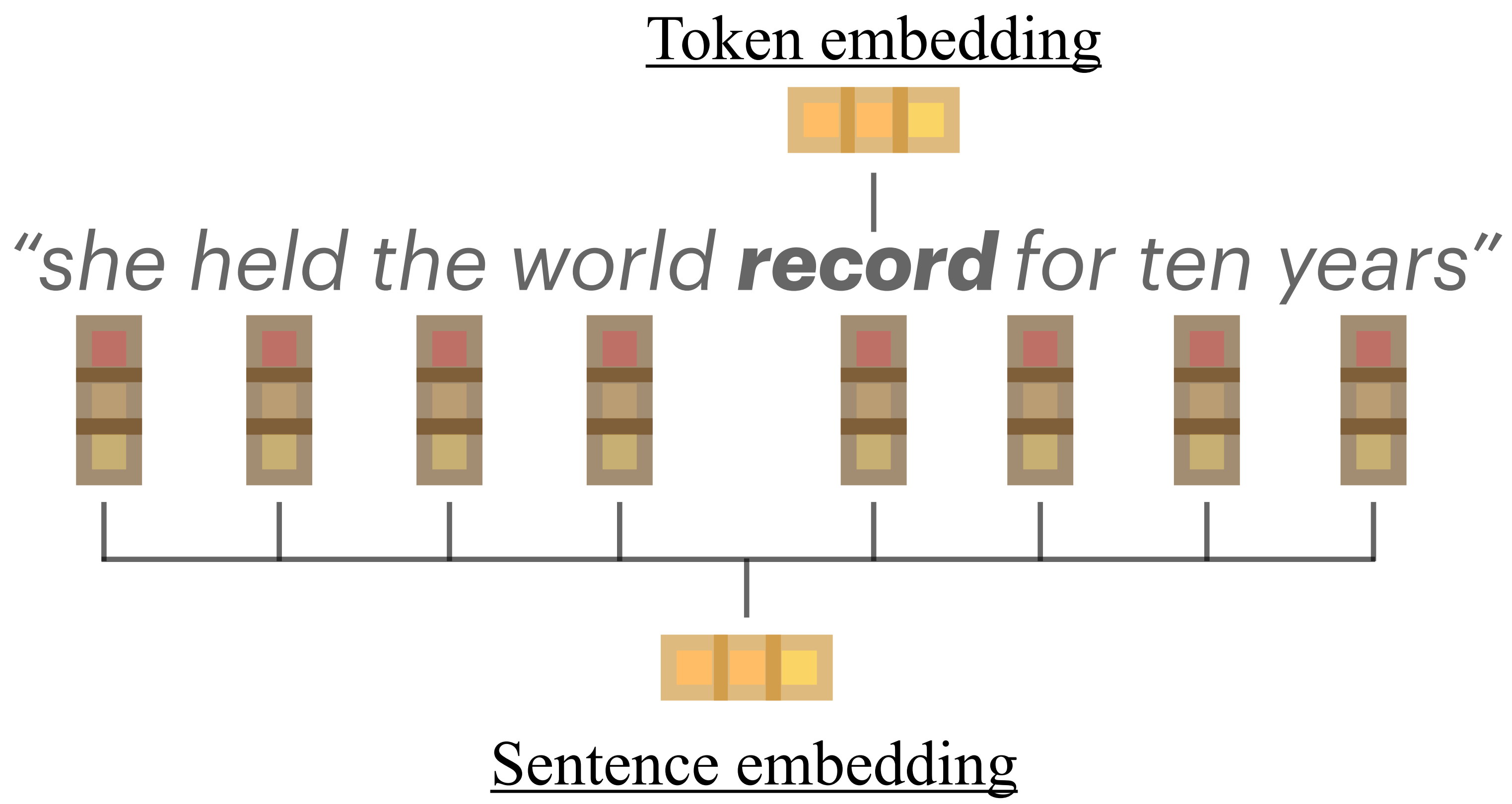


Word-in-context similarity

Correlation with human pairwise similarity judgements



Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)



Definition

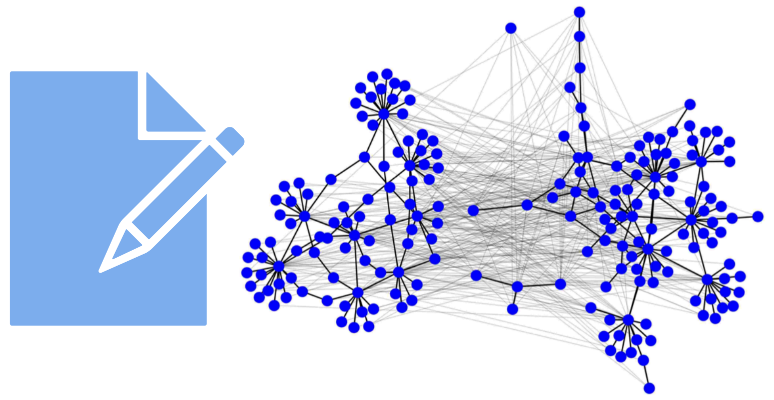
*"**record**": the best performance or most remarkable event of its kind*

Spearman's correlation.

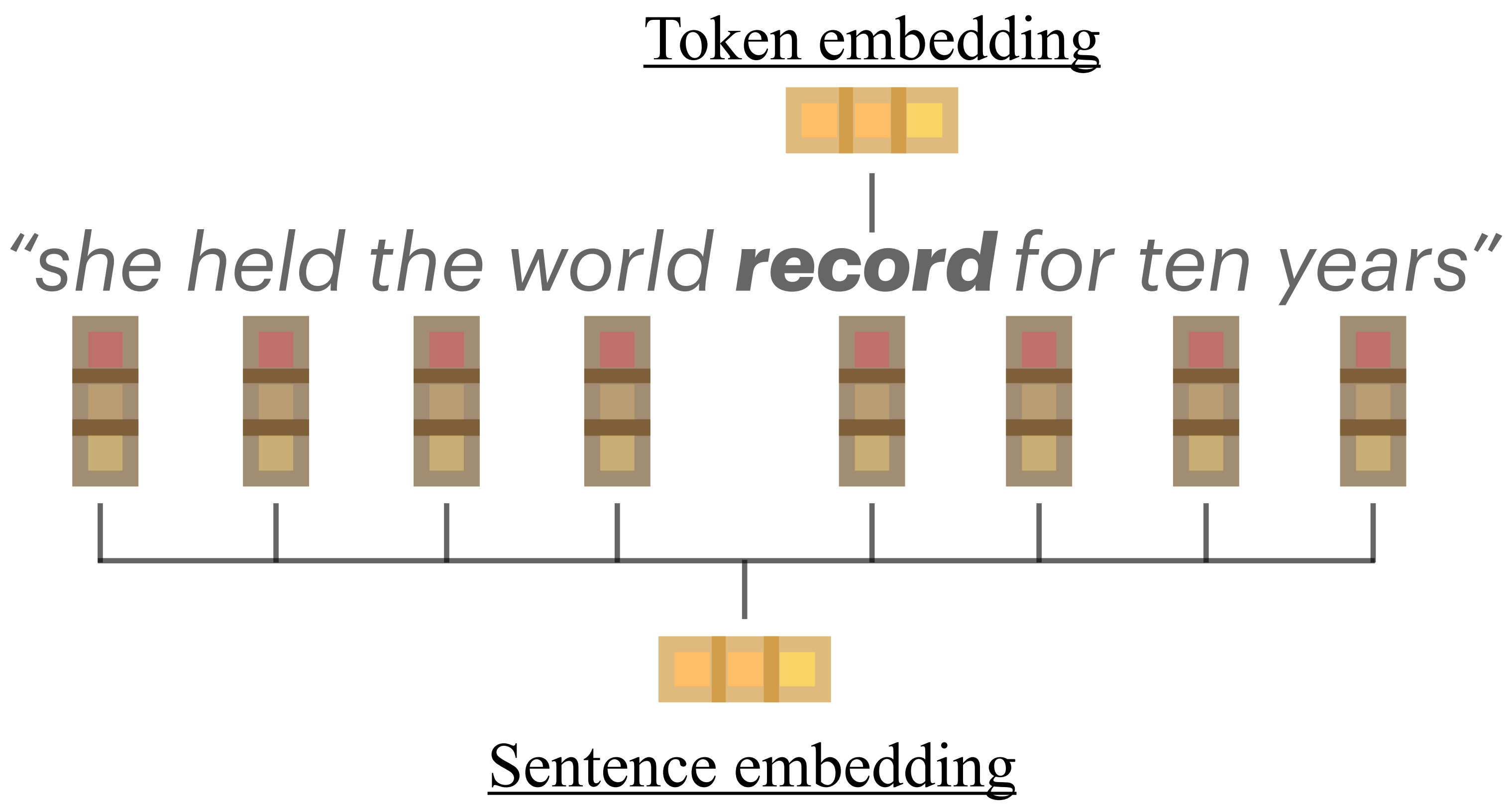
Method	Cosine	SacreBLEU	METEOR
Token embeddings	0.141		
Sentence embeddings	0.114		
FLAN-T5 XL Zero-shot	0.188		
FLAN-T5 XXL Zero-shot	0.206		
FLAN-T5 base FT	0.221		
FLAN-T5 XL FT	**0.264**		

Word-in-context similarity

Correlation with human pairwise similarity judgements



Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)



Definition

"**record**": the best performance or most remarkable event of its kind

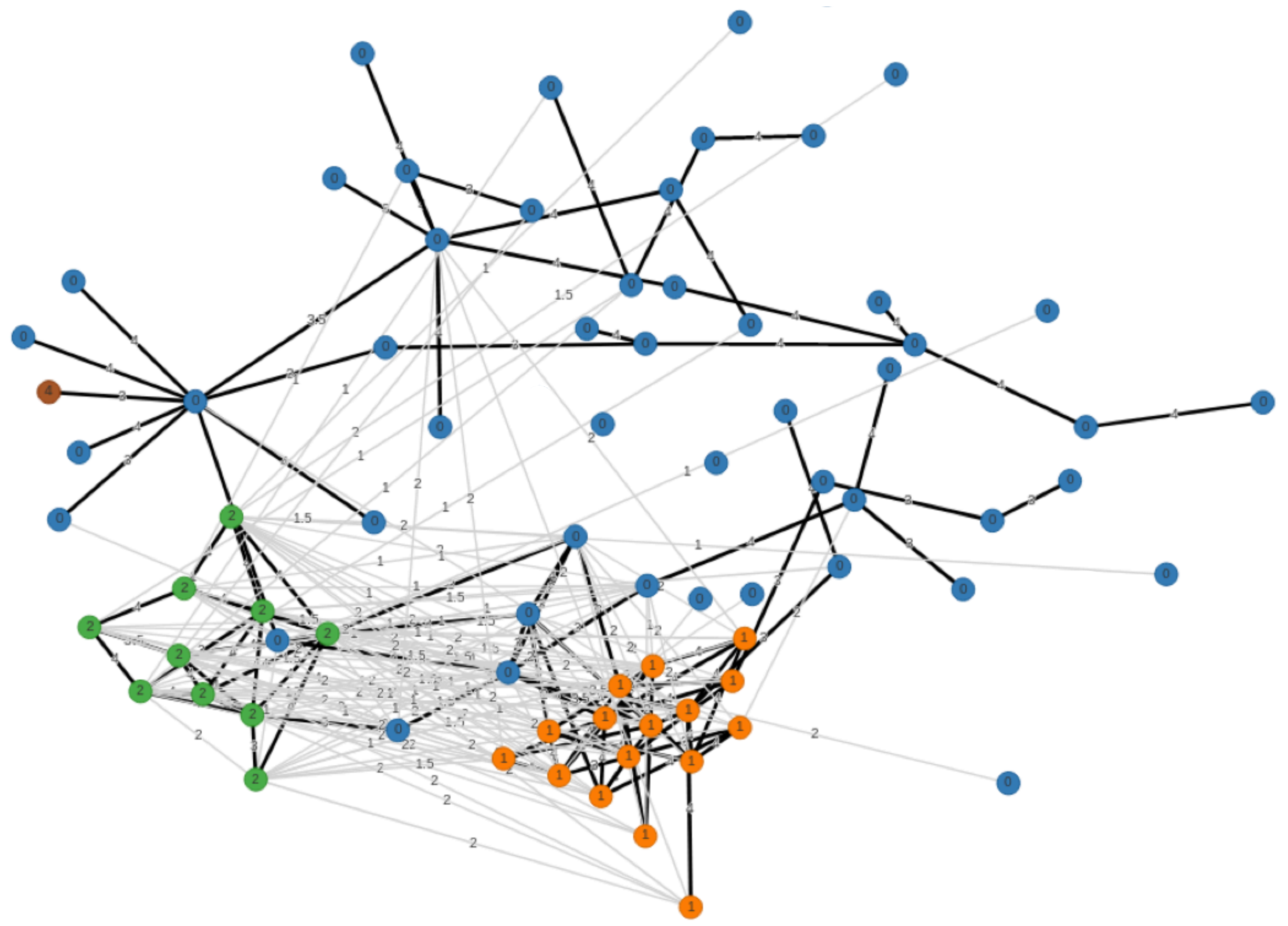
Spearman's correlation.

Method	Cosine	SacreBLEU	METEOR
Token embeddings	0.141	-	-
Sentence embeddings	0.114	-	-
FLAN-T5 XL Zero-shot	0.188	0.041	0.083
FLAN-T5 XXL Zero-shot	0.206	0.045	0.092
FLAN-T5 base FT	0.221	0.078	0.077
FLAN-T5 XL FT	**0.264**	**0.108**	**0.117**

Case study:
Semantic change analysis

Labelling word sense clusters with definitions

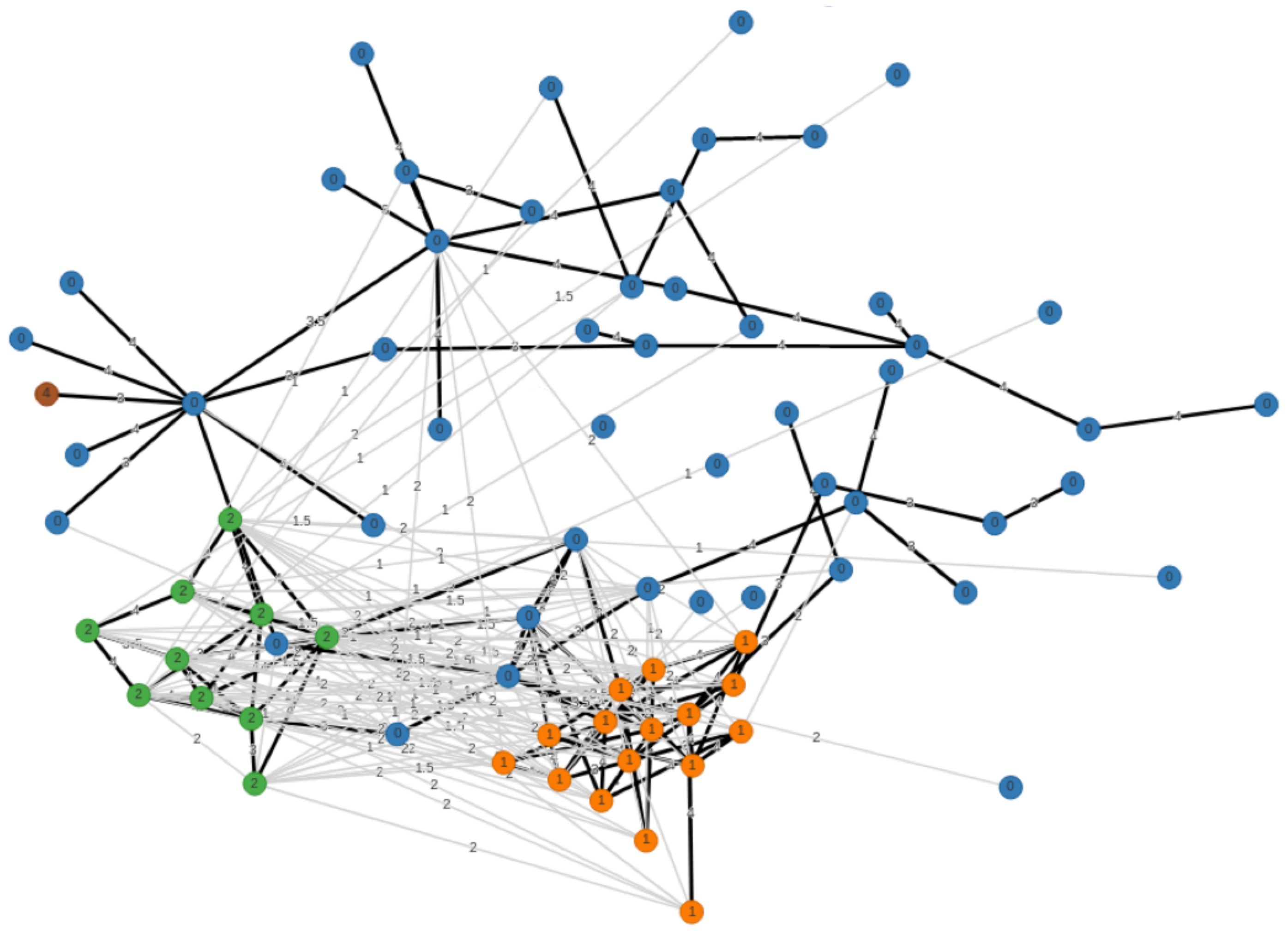
"record"
(1960-2010)



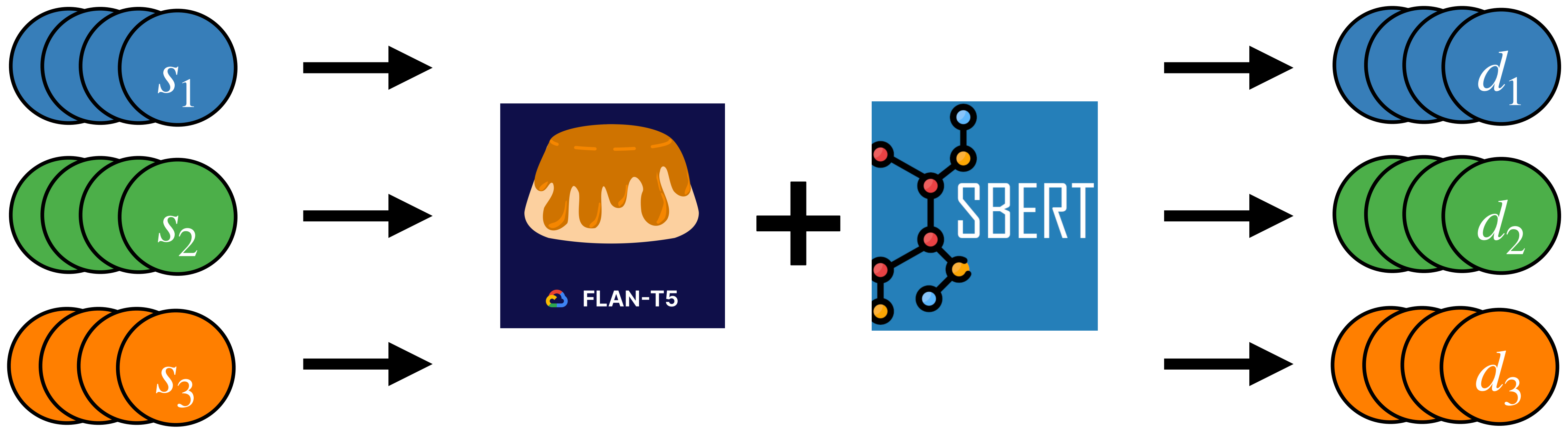
Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)

Labelling word sense clusters with definitions

"record"
(1960-2010)

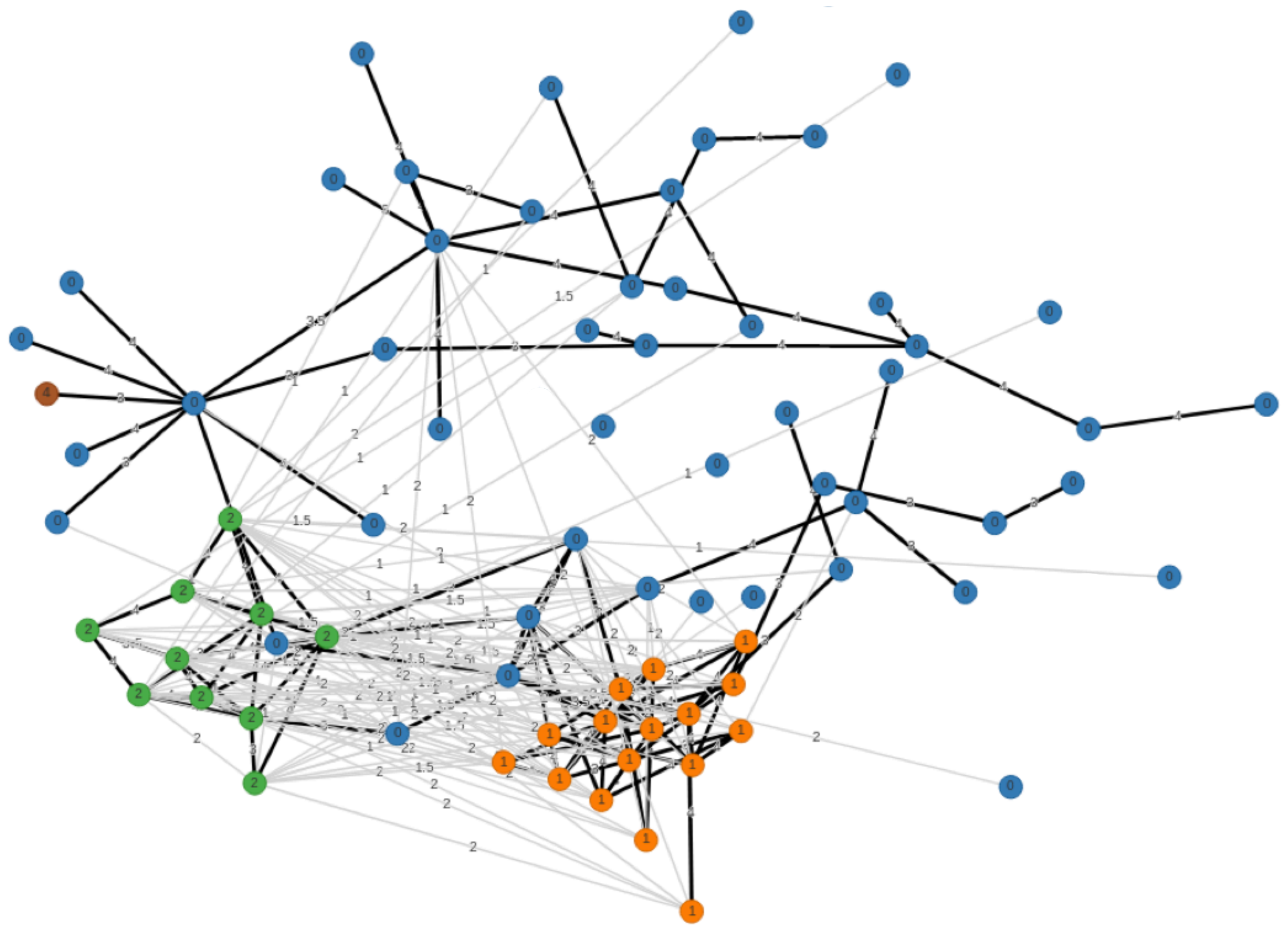


Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)

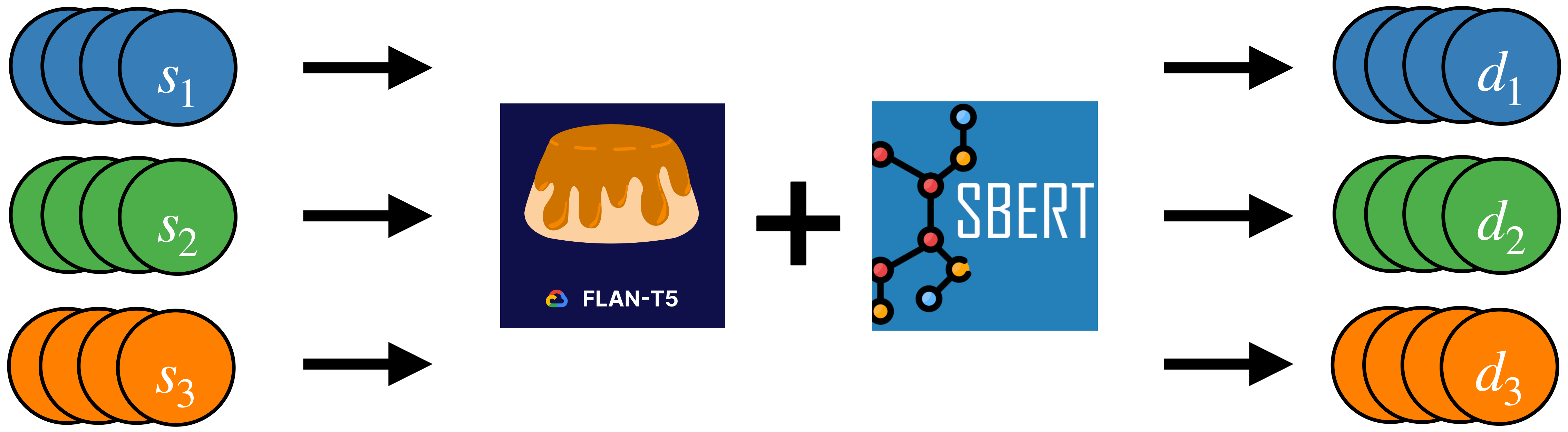


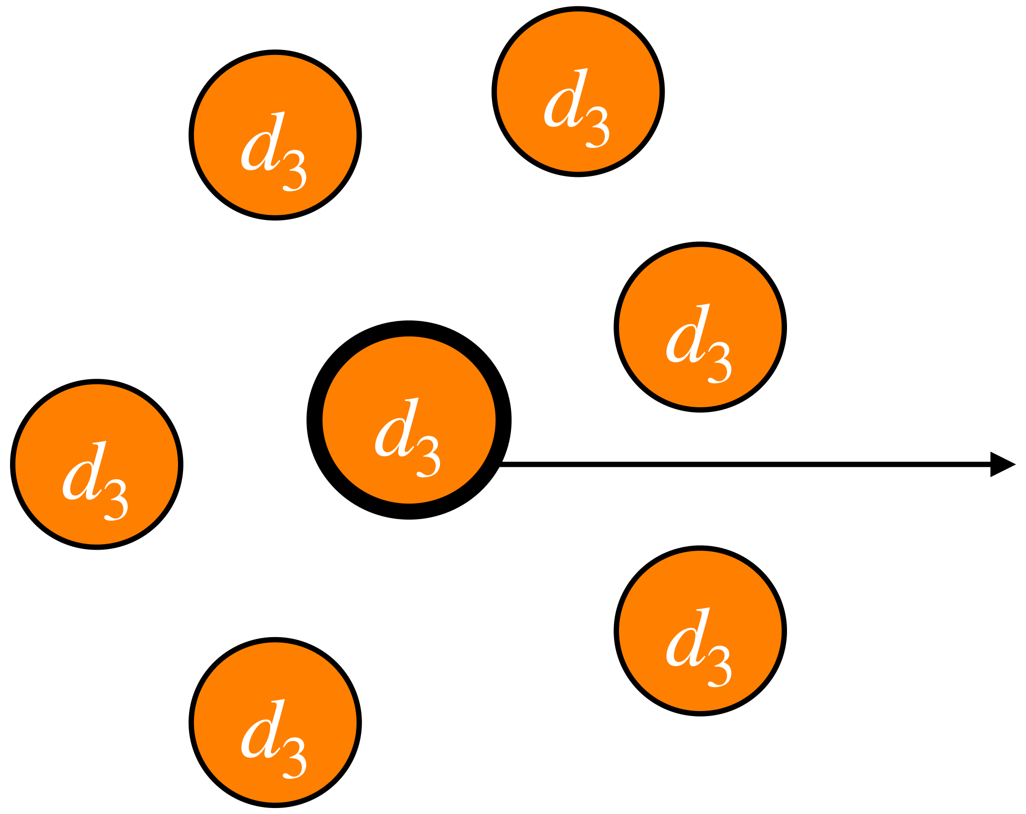
Labelling word sense clusters with definitions

"record"
(1960-2010)



Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)



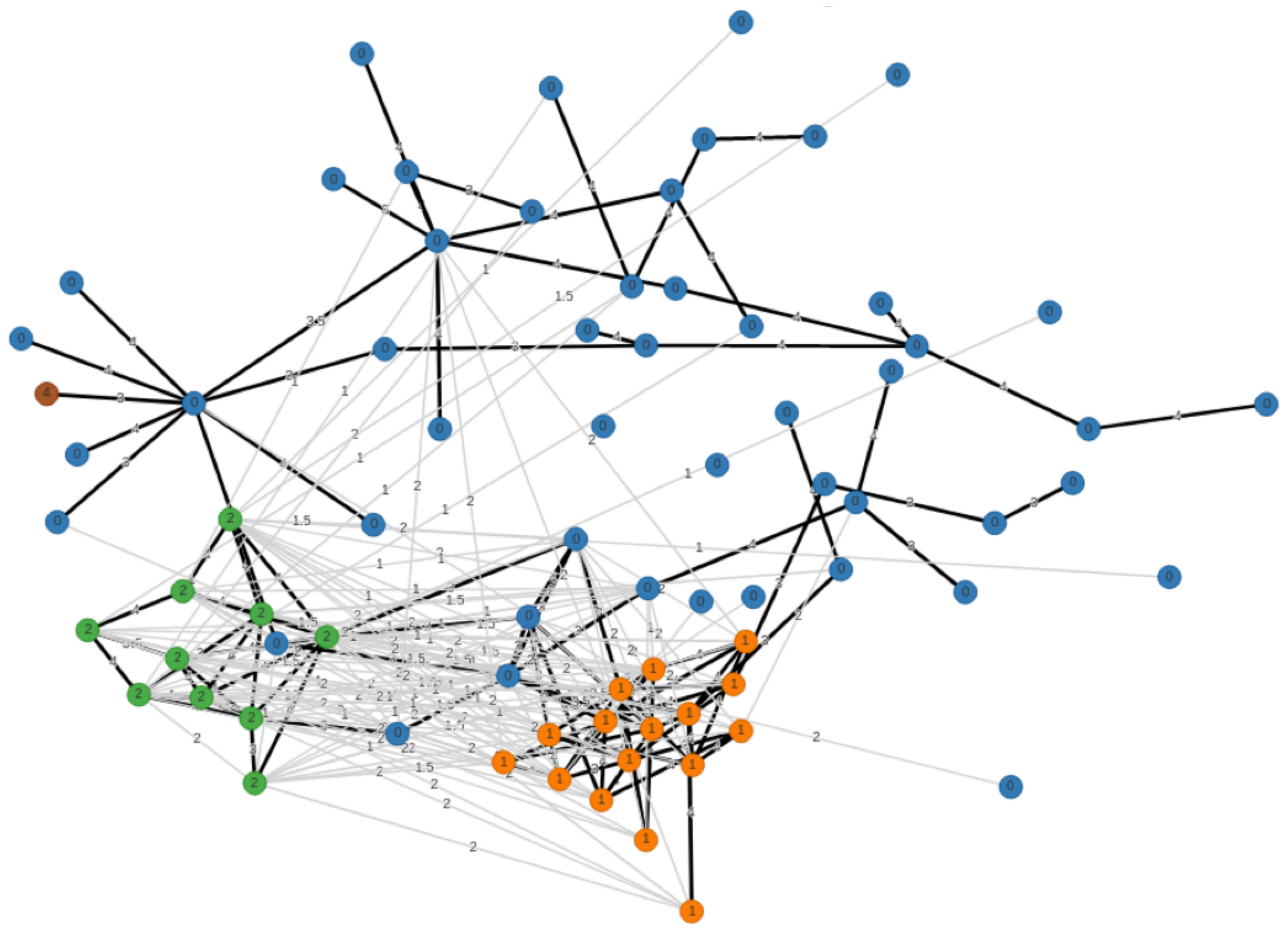


Sense label

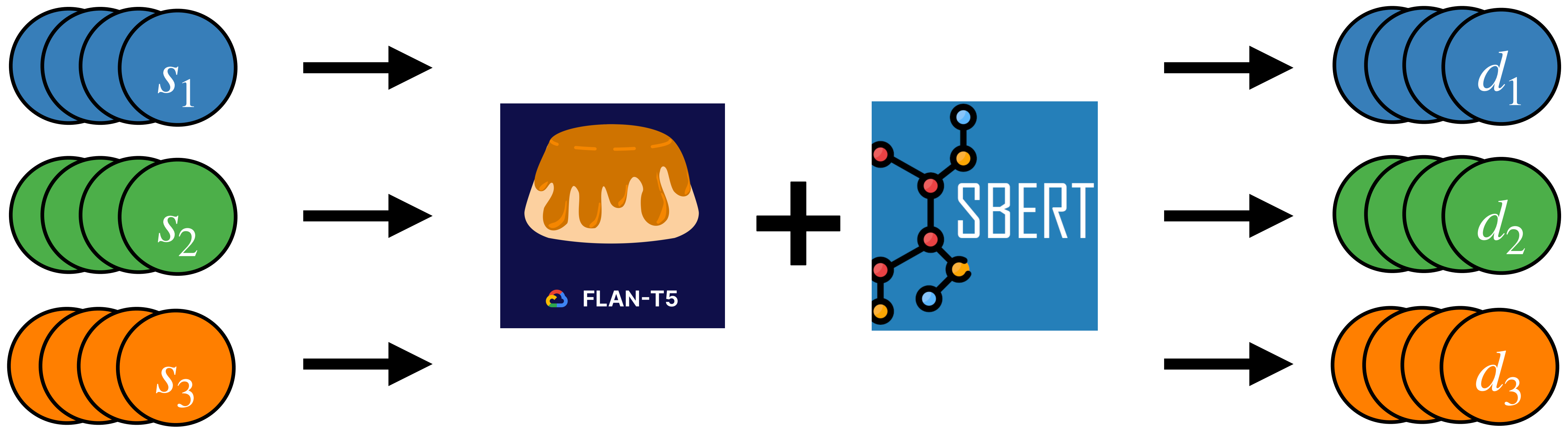
The highest score or other achievement in a game.

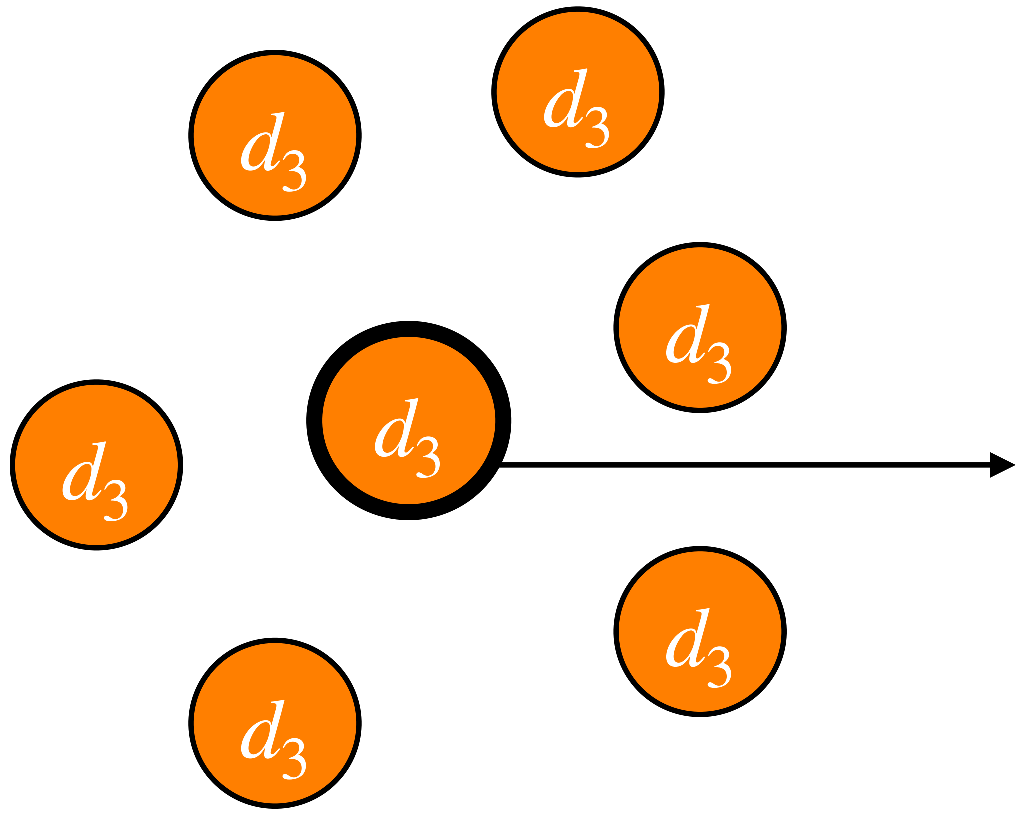
Labelling word sense clusters with definitions

“record”
(1960-2010)



Diachronic Word Usage Graphs
(Schlechtweg et al., 2021)





Sense label

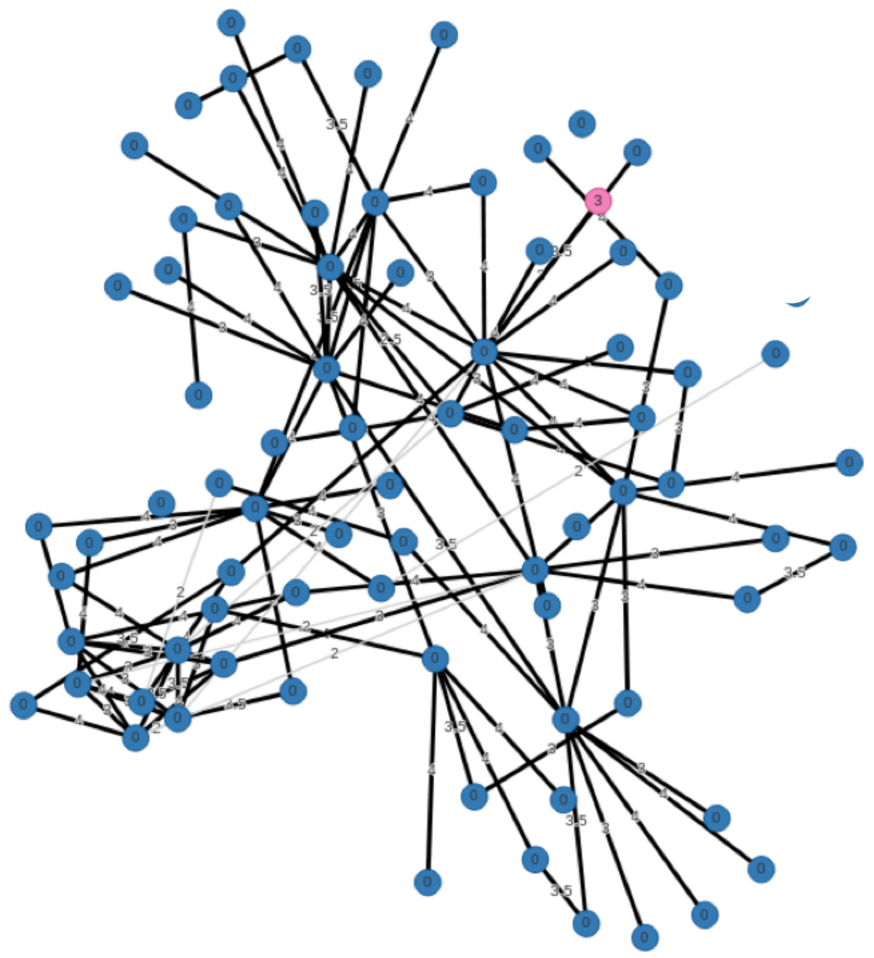
The highest score or other achievement in a game.

Human evaluation: 80% sufficient quality, 31% better than usage-based labels.

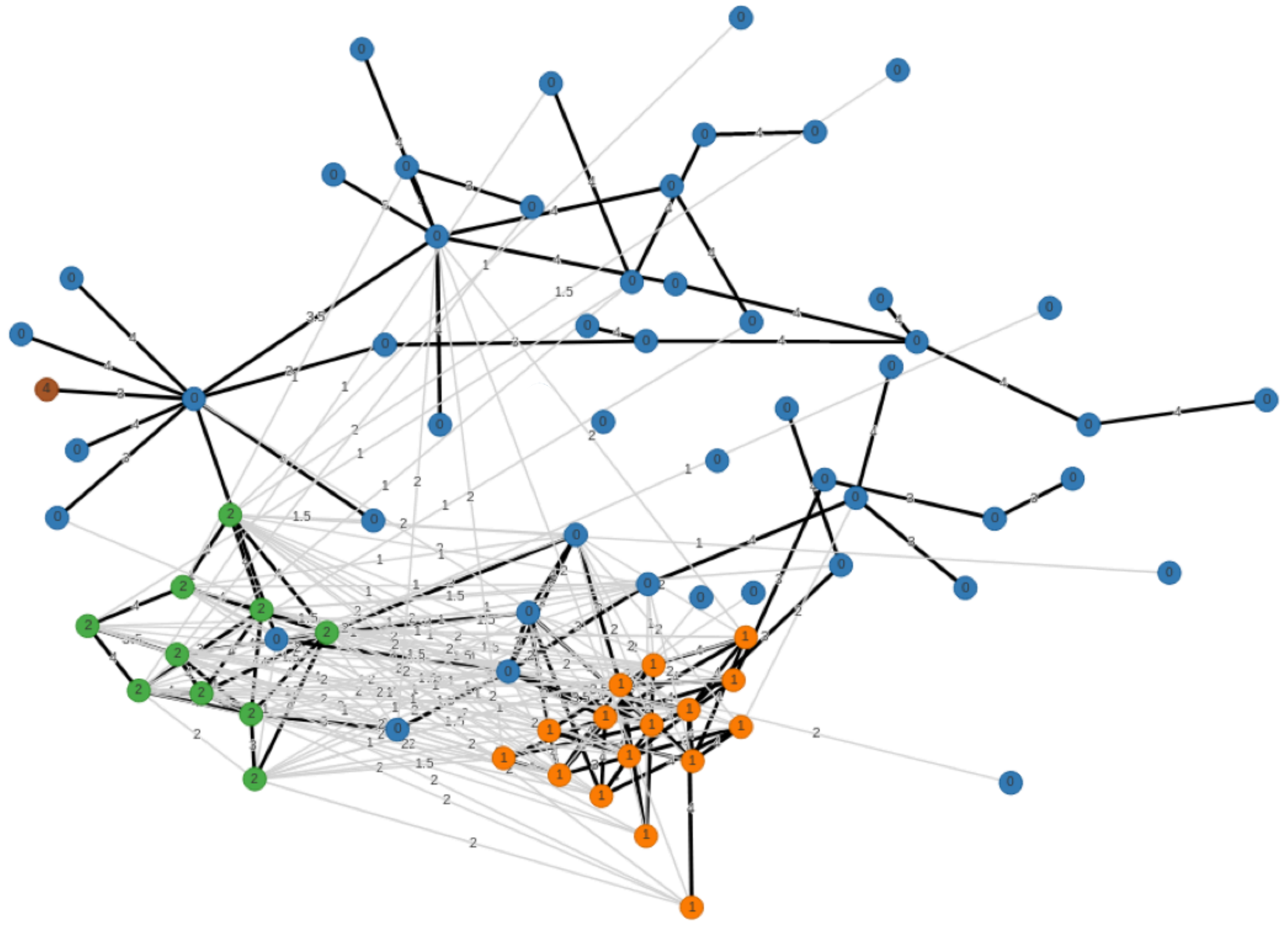
Explainable semantic change detection

Sense dynamic maps by measuring similarity between sense labels

"record"
(1810-1860)



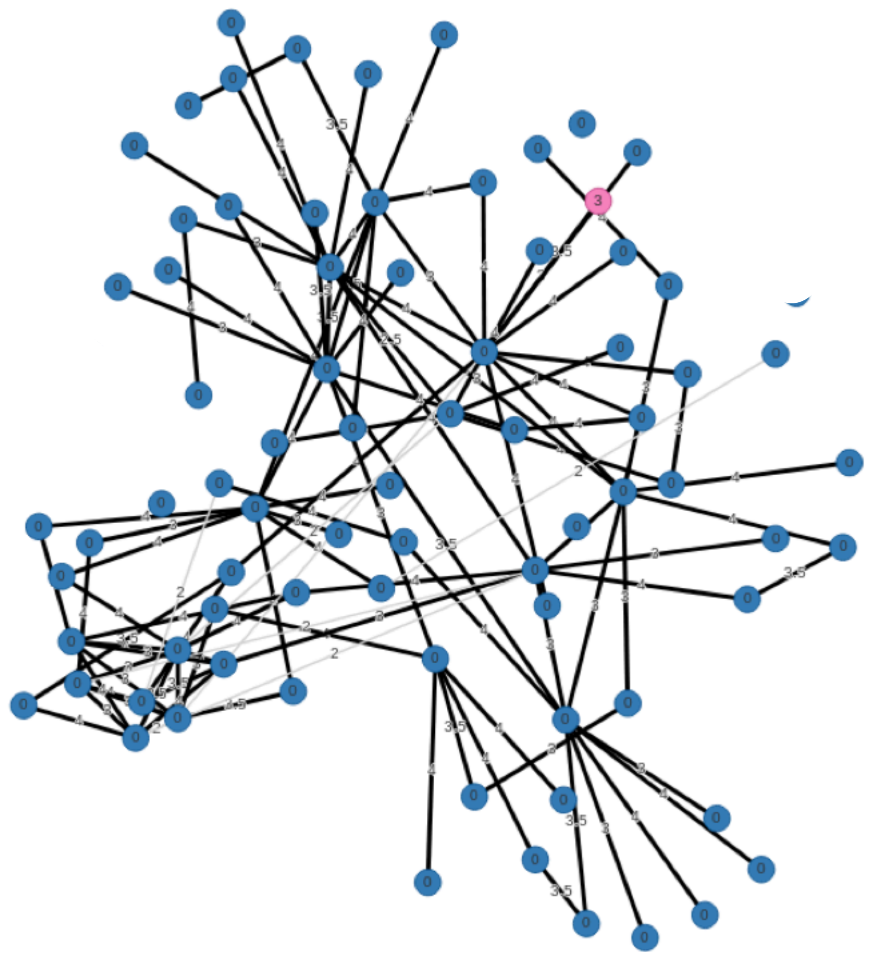
"record"
(1960-2010)



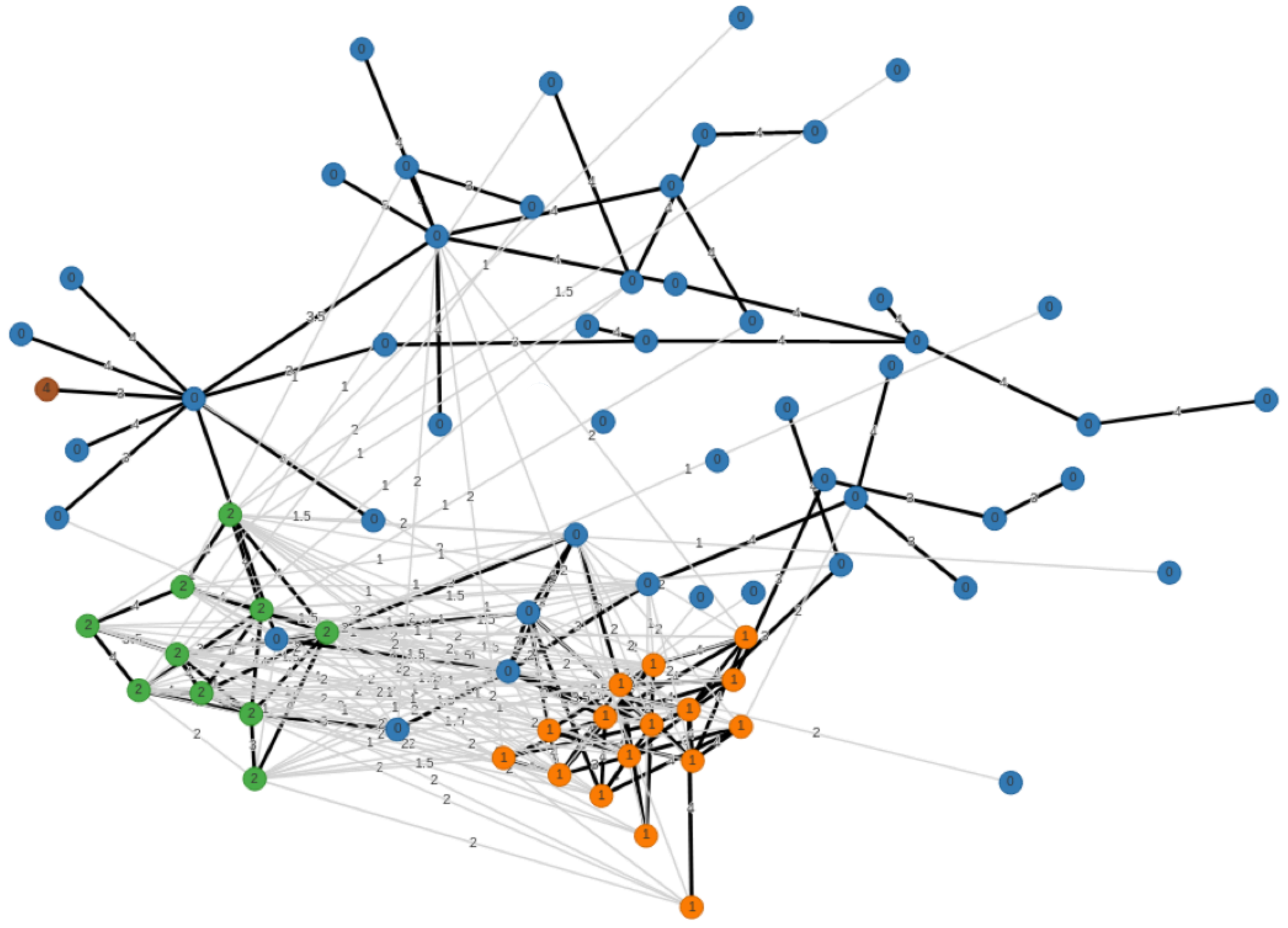
Explainable semantic change detection

Sense dynamic maps by measuring similarity between sense labels

“record”
(1810-1860)



“record”
(1960-2010)



d_1 A document or other means of providing information about past events.

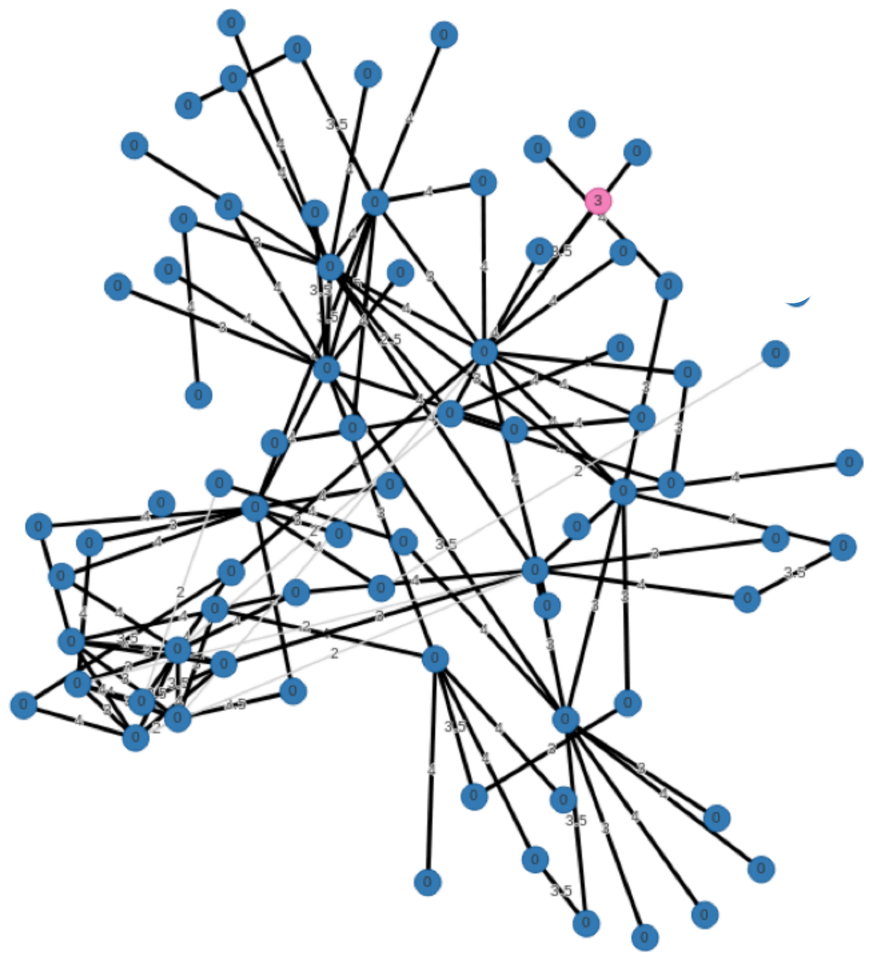
d_2 A phonograph or gramophone cylinder containing an audio recording.

d_3 The highest score or other achievement in a game.

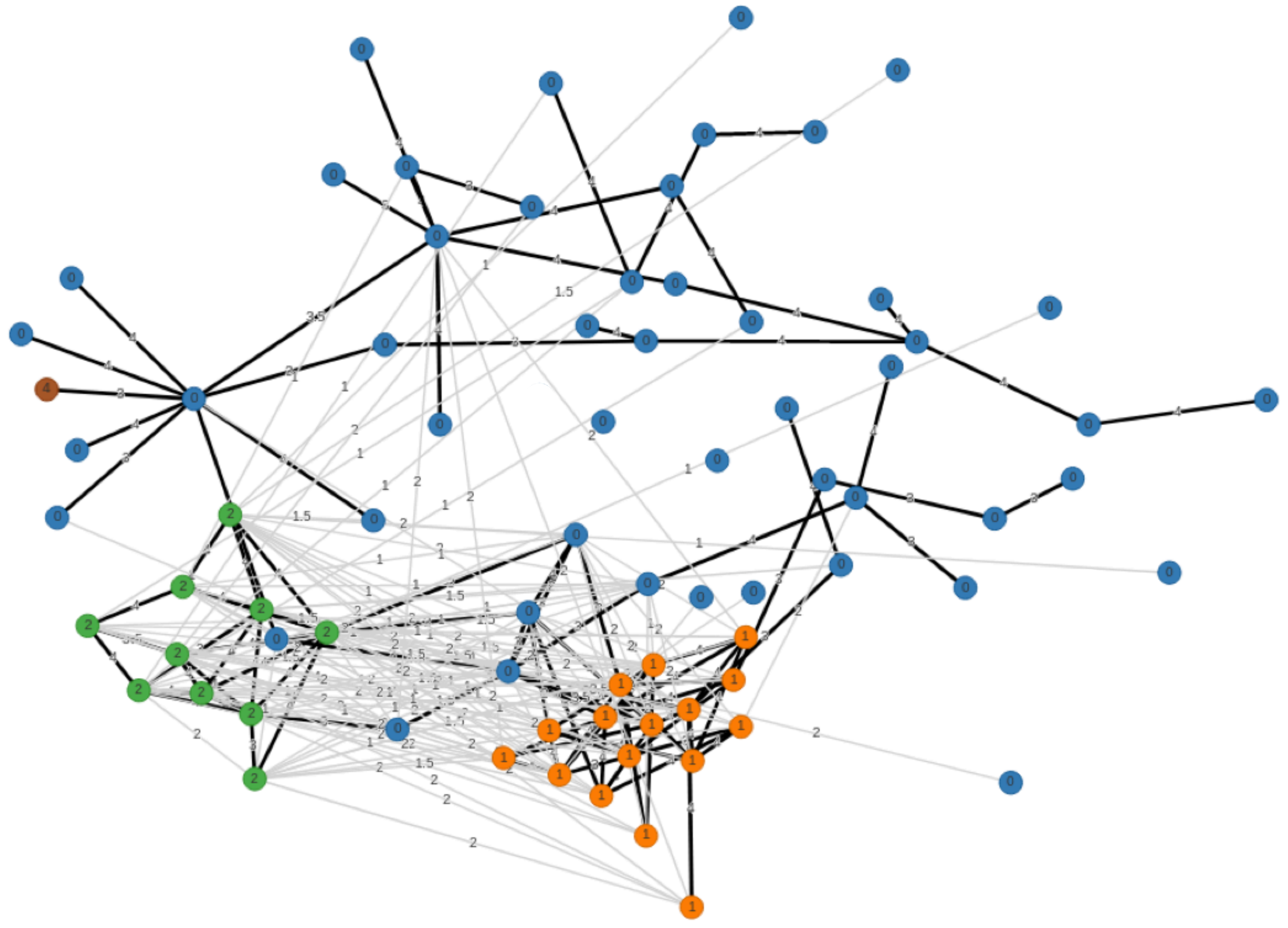
Explainable semantic change detection

Sense dynamic maps by measuring similarity between sense labels

"record"
(1810-1860)



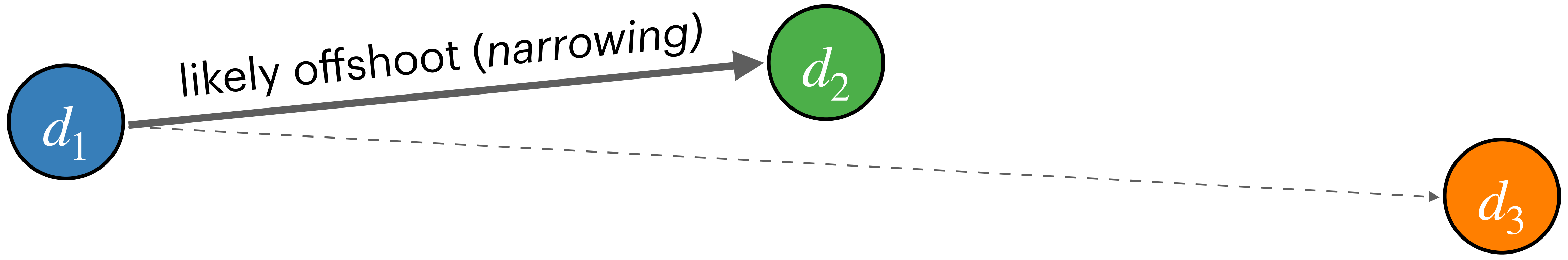
"record"
(1960-2010)



d_1 **A document or other means of providing information about past events.**

d_2 **A phonograph or gramophone cylinder containing an audio recording.**

d_3 The highest score or other achievement in a game.



Explainable semantic change detection

Fixing DWUGs

- trace incorrect or inconsistent DWUG clustering
- Two sense clusters have the same label? Likely, they are one cluster/sense.

Explainable semantic change detection

Fixing DWUGs

- trace incorrect or inconsistent DWUG clustering
- Two sense clusters have the same label? Likely, they are one cluster/sense.

'Ball' example

- Sense similarities are non-transitive:
  - Ball 0: 'A SPHERE OR OTHER OBJECT USED AS THE OBJECT OF A HIT'
  - Ball 2: 'A ROUND SOLID PROJECTILE, SUCH AS IS USED IN SHOOTING'
  - Ball 3: 'A BULLET'
- c_0 to c_2: 0.70
- c_2 to c_3: 0.53
- c_0 to c_3: 0.50 (below the outlier threshold)

Inconsistent clustering, but also interesting insights about meaning trajectory of '*ball*'.

Explainable semantic change detection

- Semantic change modelling with definitions as lexical representations

Explainable semantic change detection

- Semantic change modelling with definitions as lexical representations
- Benefits:
  - human-readable representations
  - more abstract and robust to noise
  - outperforms ‘standard’ embeddings in word-in-context similarity judgements
  - for humanities, it’s easier to operate in the space of the definitions.

More in the paper: [Giulianelli et al., 2023]

Full results

		WordNet test set			Oxford test set		
Model	Generalization test	BLEU	ROUGE-L	BERT-F1	BLEU	ROUGE-L	BERT-F1
[Huang et al., 2021]	*Unknown*	32.72	-	-	**26.52**	-	-
T5 base	Zero-shot (task shift)	2.01	8.24	82.98	1.72	7.48	78.79
T5 base	Soft domain shift	9.21	25.71	86.44	7.28	24.13	86.03
Flan-T5 base	Zero-shot (task shift)	4.08	15.32	87.00	3.71	17.25	86.44
Flan-T5 base	In-distribution	8.80	23.19	87.49	6.15	20.84	86.48
Flan-T5 base	Hard domain shift	6.89	20.53	87.16	4.32	17.00	85.88
Flan-T5 base	Soft domain shift	10.38	27.17	88.22	7.18	23.04	86.90
Flan-T5 large	Soft domain shift	14.37	33.74	88.21	10.90	30.05	87.44
T5 XL	Zero-shot (task shift)	2.05	8.28	81.90	2.28	9.73	80.37
T5 XL	Soft domain shift	**34.14**	**53.55**	91.40	18.82	38.26	88.81
Flan-T5 XL	Zero-shot (task shift)	2.70	12.72	86.72	2.88	16.20	86.52
Flan-T5 XL	In-distribution	11.49	28.96	88.90	16.61	36.27	89.40
Flan-T5 XL	Hard domain shift	29.55	48.17	91.39	8.37	25.06	87.56
Flan-T5 XL	Soft domain shift	**32.81**	**52.21**	**92.16**	18.69	**38.72**	**89.75**

Contents

Future directions and open problems

Still much to be done in the field of semantic change modeling

Future directions and open problems

Still much to be done in the field of semantic change modeling

- Better methods (of course).

Future directions and open problems

Still much to be done in the field of semantic change modeling

- Better methods (of course).
- Wider scope of languages required.

Future directions and open problems

Still much to be done in the field of semantic change modeling

- Better methods (of course).
- Wider scope of languages required.
- More gold standard test sets for different languages and domains are needed (especially beyond Indo-European).

Future directions and open problems

Still much to be done in the field of semantic change modeling

- Better methods (of course).
- Wider scope of languages required.
- More gold standard test sets for different languages and domains are needed (especially beyond Indo-European).
- Explainable semantic change modeling:
  1. Sub-classification of types of semantic shifts.
  2. Identifying the source of a shift (linguistic or extra-linguistic causes).
  3. Quantifying the weight of senses acquired over time.
  4. Identifying groups of words that change their meaning together in correlated ways (co-evolution).
- Is it possible to to multi-modal semantic change detection?

Future directions and open problems

- How to infer the temporal dynamics of semantic relations between persons and organizations, ideas and technologies, etc.

Future directions and open problems

- How to infer the temporal dynamics of semantic relations between persons and organizations, ideas and technologies, etc.
- How to tell words used in different lexicographic senses from words used in contextually varied surroundings, but in the same sense? Does this difference even exist?

Future directions and open problems

- How to infer the temporal dynamics of semantic relations between persons and organizations, ideas and technologies, etc.
- How to tell words used in different lexicographic senses from words used in contextually varied surroundings, but in the same sense? Does this difference even exist?
- The task of binary classification is important, not only ranking: did the word change its meaning or not?
  - unavoidably linked with the issues of senses and their emergence or dying out.

Future directions and open problems

- How to infer the temporal dynamics of semantic relations between persons and organizations, ideas and technologies, etc.
- How to tell words used in different lexicographic senses from words used in contextually varied surroundings, but in the same sense? Does this difference even exist?
- The task of binary classification is important, not only ranking: did the word change its meaning or not?
  - unavoidably linked with the issues of senses and their emergence or dying out.
- How to model onomasiological change?
  - *'How do you call a hand-held peripheral device designed to provide input to a computer or a gaming console? And how was this object called in the 60's?'*

Future directions and open problems

- How to infer the temporal dynamics of semantic relations between persons and organizations, ideas and technologies, etc.
- How to tell words used in different lexicographic senses from words used in contextually varied surroundings, but in the same sense? Does this difference even exist?
- The task of binary classification is important, not only ranking: did the word change its meaning or not?
  - unavoidably linked with the issues of senses and their emergence or dying out.
- How to model onomasiological change?
  - *'How do you call a hand-held peripheral device designed to provide input to a computer or a gaming console? And how was this object called in the 60's?'*

Despite the challenges, significant results in semantic change modeling are already achieved. You are welcome to try these methods in your own work!

This is it!

Thanks for coming to our course!
Any questions?

References I

Chung, H. W., Hou, L., Longpre, S., Zoph, B., Tay, Y., Fedus, W., Li, E., Wang, X., Dehghani, M., Brahma, S., et al. (2022).
Scaling instruction-finetuned language models.
arXiv preprint arXiv:2210.11416.

Gadetsky, A., Yakubovskiy, I., and Vetrov, D. (2018).
Conditional generators of words definitions.
In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*, pages 266–271, Melbourne, Australia. Association for Computational Linguistics.

References II

Giulianelli, M., Luden, I., Fernandez, R., and Kutuzov, A. (2023).
Interpretable word sense representations via definition generation: The case of semantic change analysis.
In *Proceedings of the 61st Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)*, pages 3130–3148, Toronto, Canada. Association for Computational Linguistics.

Huang, H., Kajiwara, T., and Arase, Y. (2021).
Definition modelling for appropriate specificity.
In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 2499–2509, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.

References III

Hupkes, D., Giulianelli, M., Dankers, V., Artetxe, M., Elazar, Y., Pimentel, T., Christodoulopoulos, C., Lasri, K., Saphra, N., Sinclair, A., et al. (2022). State-of-the-art generalisation research in NLP: A taxonomy and review. *arXiv preprint arXiv:2210.03050*.

Ishiwatari, S., Hayashi, H., Yoshinaga, N., Neubig, G., Sato, S., Toyoda, M., and Kitsuregawa, M. (2019). Learning to describe unknown phrases with local and global contexts. In *Proceedings of the 2019 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, Volume 1 (Long and Short Papers)*, pages 3467–3476, Minneapolis, Minnesota. Association for Computational Linguistics.

References IV

Mickus, T., Van Deemter, K., Constant, M., and Paperno, D. (2022).
Semeval-2022 task 1: CODWOE – comparing dictionaries and word embeddings.
In *Proceedings of the 16th International Workshop on Semantic Evaluation (SemEval-2022)*, pages 1–14, Seattle, United States. Association for Computational Linguistics.

Reimers, N. and Gurevych, I. (2019).
Sentence-BERT: Sentence embeddings using Siamese BERT-networks.
In *Proceedings of the 2019 Conference on Empirical Methods in Natural Language Processing and the 9th International Joint Conference on Natural Language Processing (EMNLP-IJCNLP)*, pages 3982–3992, Hong Kong, China. Association for Computational Linguistics.

References V

Schlechtweg, D., Tahmasebi, N., Hengchen, S., Dubossarsky, H., and McGillivray, B. (2021).
DWUG: A large resource of diachronic word usage graphs in four languages.
In *Proceedings of the 2021 Conference on Empirical Methods in Natural Language Processing*, pages 7079–7091, Online and Punta Cana, Dominican Republic. Association for Computational Linguistics.