TempEL: Linking Dynamically Evolving and Newly Emerging Entities

Klim Zaporojets¹, Lucie-Aimée Kaffee², Johannes Deleu¹, Thomas Demeester¹, Chris Develder¹, Isabelle Augenstein²

klimzaporojets@gmail.com





Entity Linking task links anchor mentions to target entities

Entity mentions

The talented soccer player was praised as the "new **Sheva**" by journalists ...

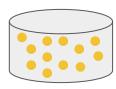
Entity Linking task links anchor mentions to target entities



Entity mentions

Knowledge Base (KB)

The talented soccer player was praised as the "new **Sheva**" by journalists ...



Entity Linking task links anchor mentions to target entities

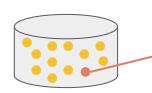


Entity mentions

Knowledge Base (KB)

KB entity description

The talented soccer player was praised as the "new <u>Sheva</u>" by journalists ...



Andriy Shevchenko is a *Ukrainian* footballer who plays for Dynamo Kyiv and the Ukraine national team as a striker [...]

Mentions and entities exist in time

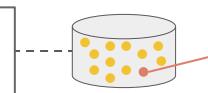


Entity mentions

Knowledge Base (KB)

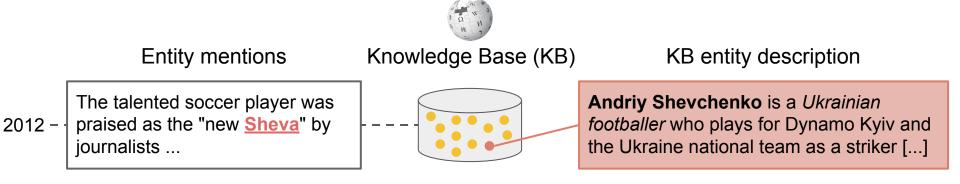
KB entity description

O12 - The talented soccer player was praised as the "new Sheva" by journalists ...



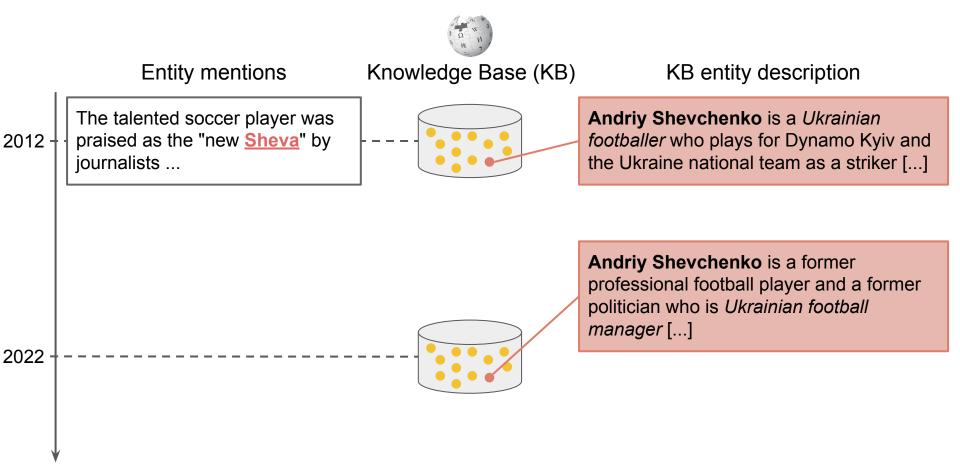
Andriy Shevchenko is a *Ukrainian* footballer who plays for Dynamo Kyiv and the Ukraine national team as a striker [...]

Mentions and entities exist in time

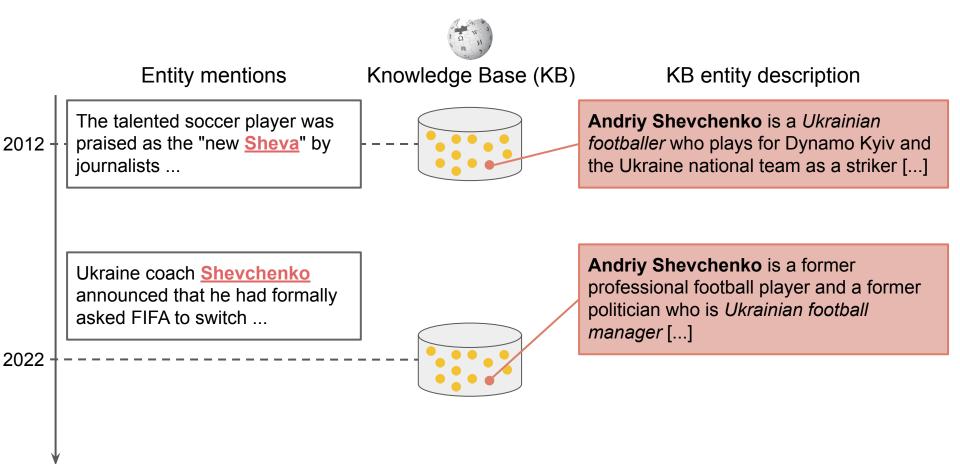


Key research question: How temporal evolution in KB entities and entity mentions affects entity linking task?

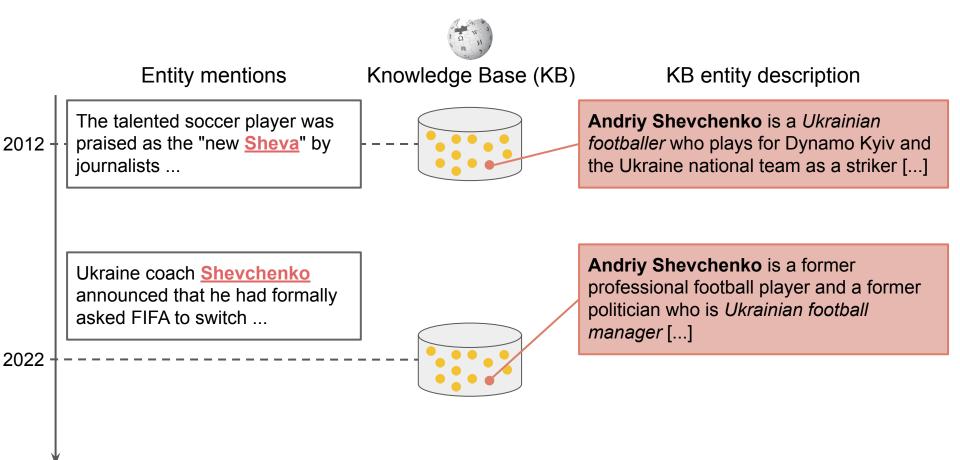
Continual entities evolve over time



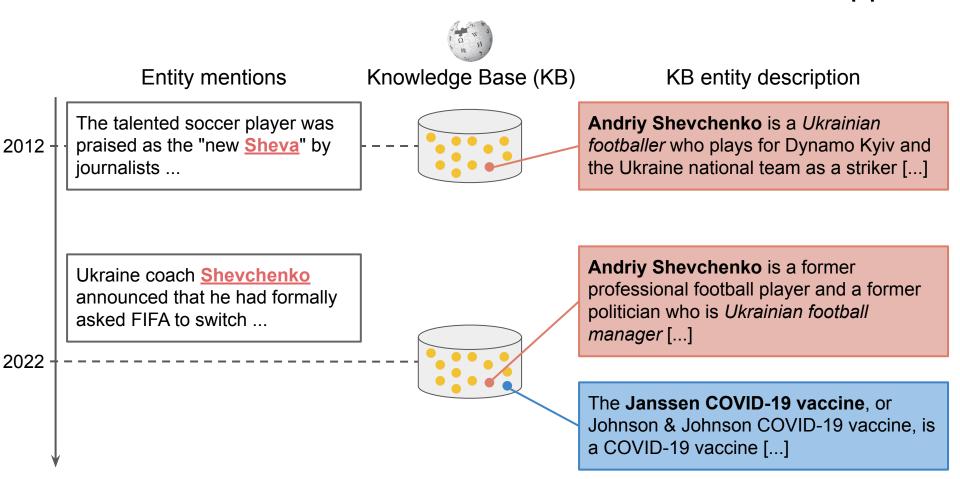
Continual entities evolve over time



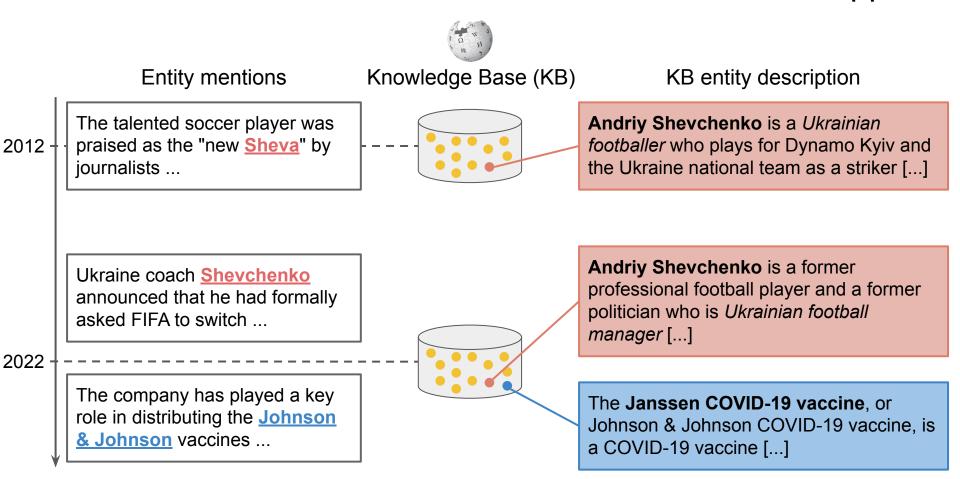
Continual entities evolve over time



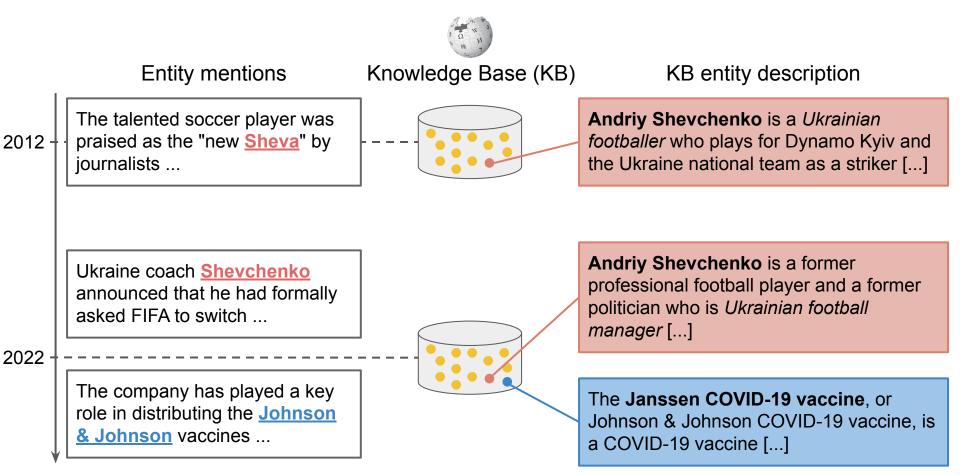
Continual entities evolve over time and new entities appear



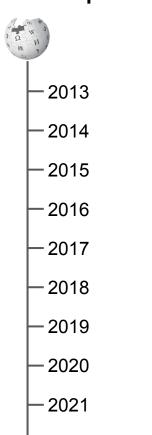
Continual entities evolve over time and new entities appear



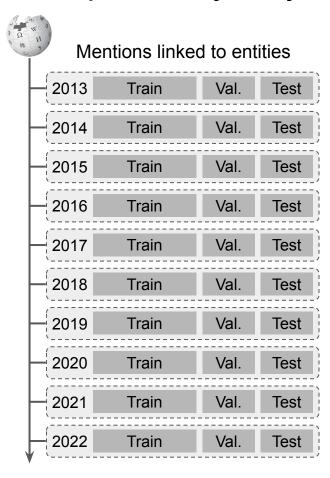
How time evolution affects entity linking?



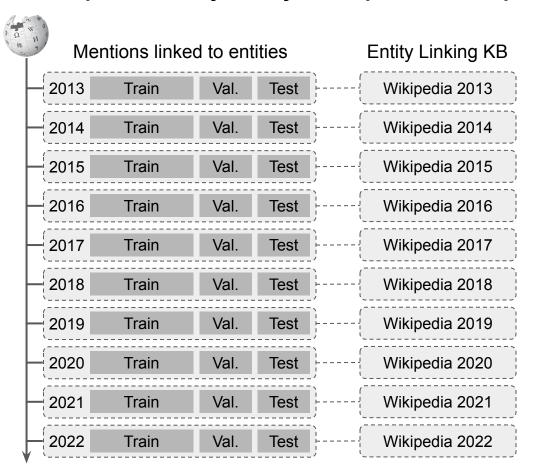
TempEL: 10 yearly temporal snapshots from Wikipedia



TempEL: 10 yearly temporal snapshots from Wikipedia



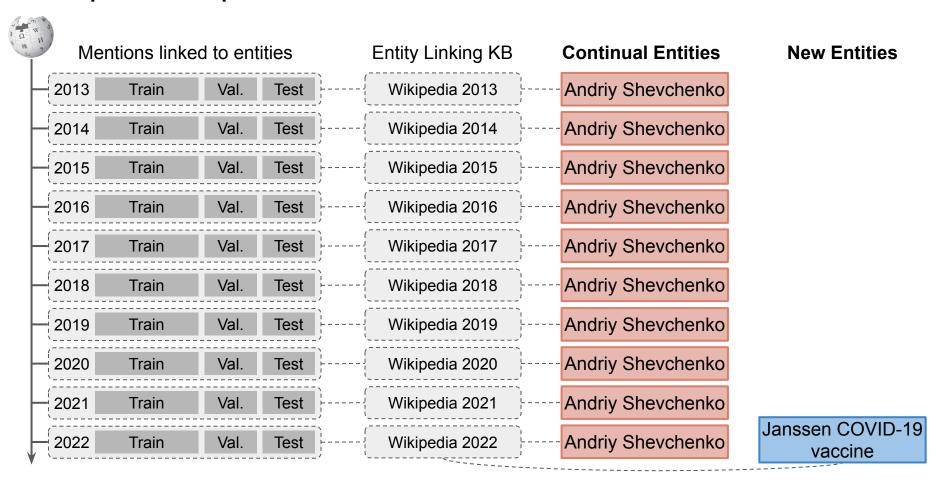
TempEL: 10 yearly temporal snapshots from Wikipedia



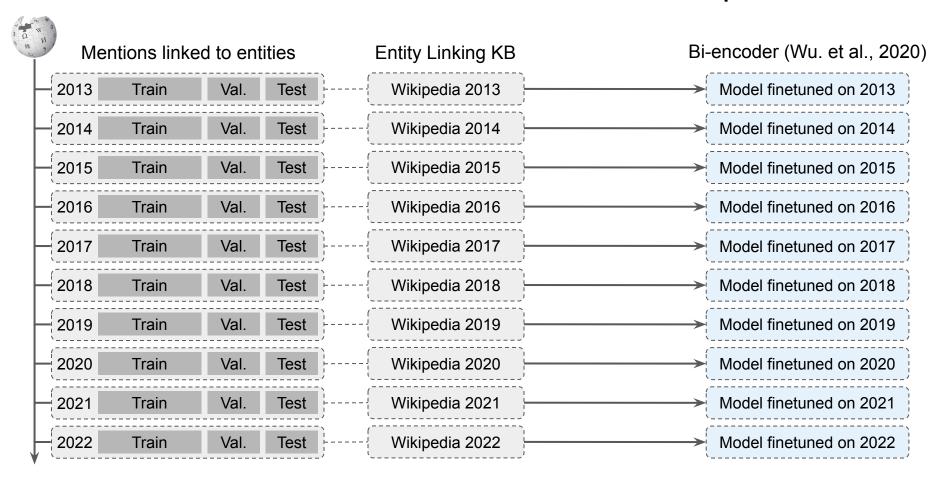
TempEL composed of *continual* and *new* entities

Mer	ntions linke	ities		Entity Linking KB		Continual Entities	
2013	Train	Val.	Test	}	Wikipedia 2013	}	Andriy Shevchenko
	Train	Val.	Test	}	Wikipedia 2014	}	Andriy Shevchenko
2015	Train	Val.	Test	}	Wikipedia 2015	}	Andriy Shevchenko
— [2016	Train	Val.	Test	}	Wikipedia 2016	}	Andriy Shevchenko
2017	Train	Val.	Test	}	Wikipedia 2017	}	Andriy Shevchenko
2018	Train	Val.	Test	}	Wikipedia 2018	}	Andriy Shevchenko
2019	Train	Val.	Test	}	Wikipedia 2019	}	Andriy Shevchenko
_[2020	Train	Val.	Test	}	Wikipedia 2020	}	Andriy Shevchenko
2021	Train	Val.	Test	}	Wikipedia 2021	}	Andriy Shevchenko
2022	Train	Val.	Test	}	Wikipedia 2022	}	Andriy Shevchenko

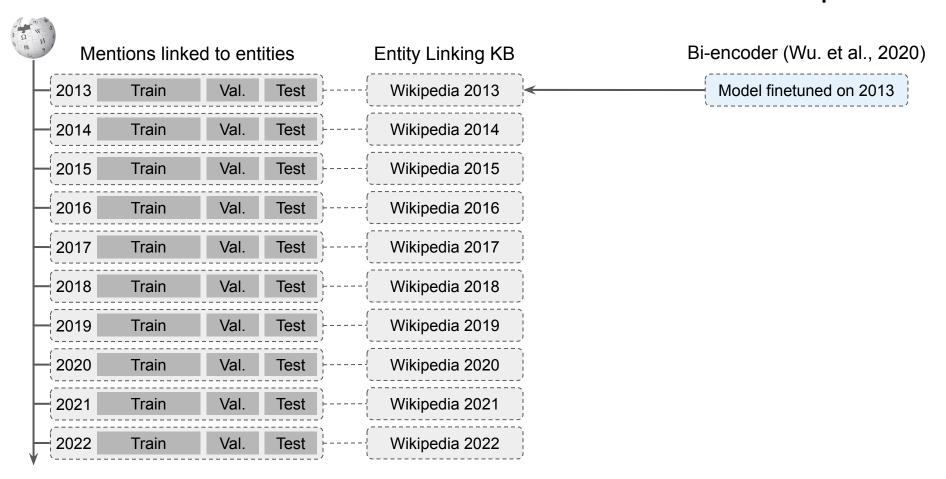
TempEL composed of *continual* and *new* entities



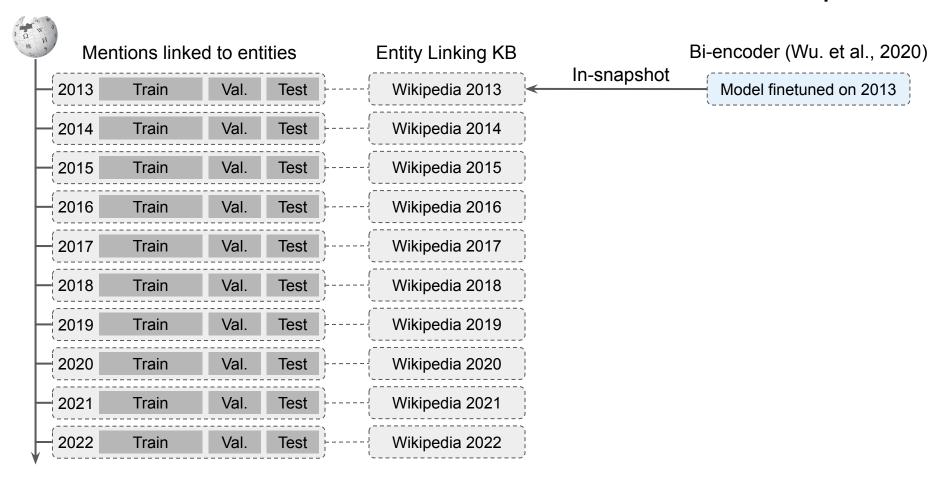
Different model is finetuned on each of the snapshots



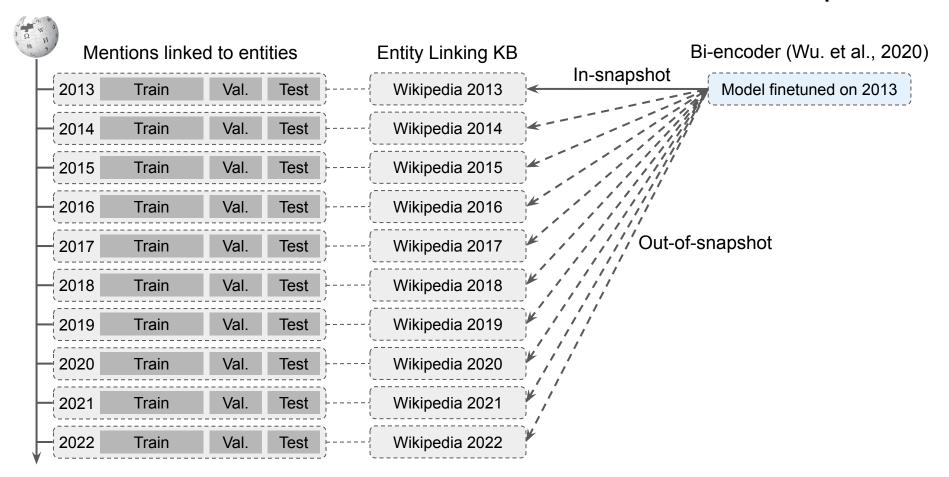
Each model is evaluated on the same or different snapshot



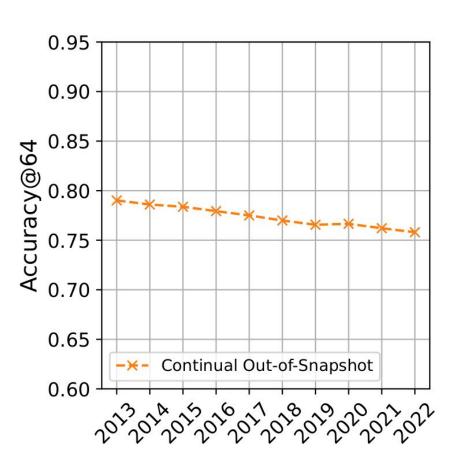
Each model is evaluated on the same or different snapshot



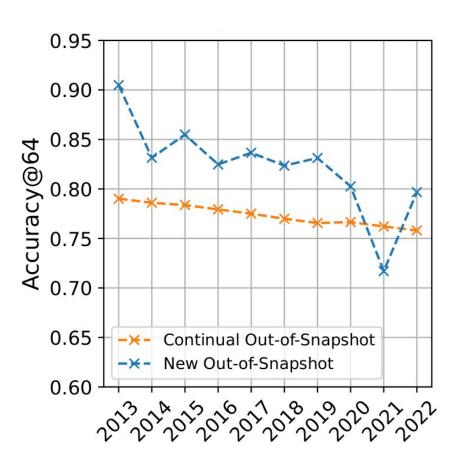
Each model is evaluated on the same or different snapshot



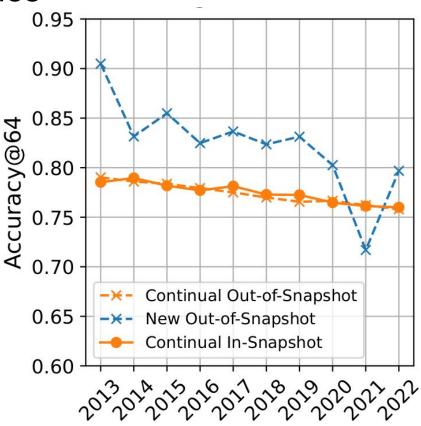
General decrease in performance for continual entities



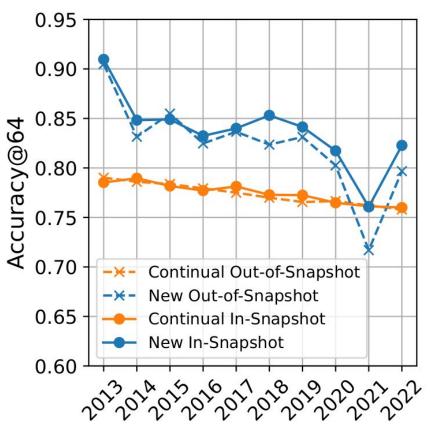
Irregular decreasing trend in performance on new entities



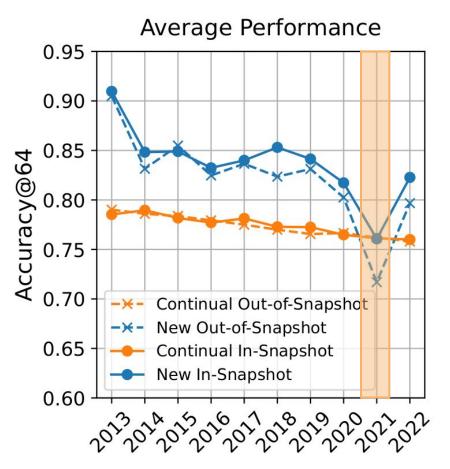
Marginal difference between in and out-of-snapshot for continual entities



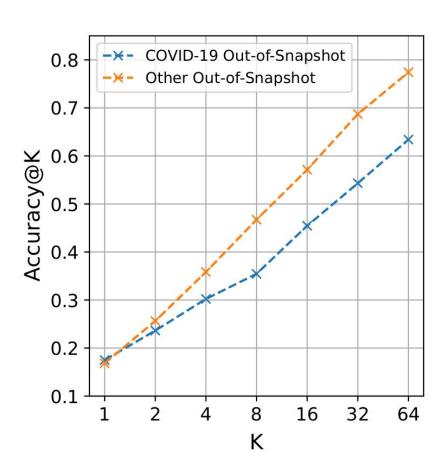
In-snapshot significantly superior to out-of-snapshot for new entities



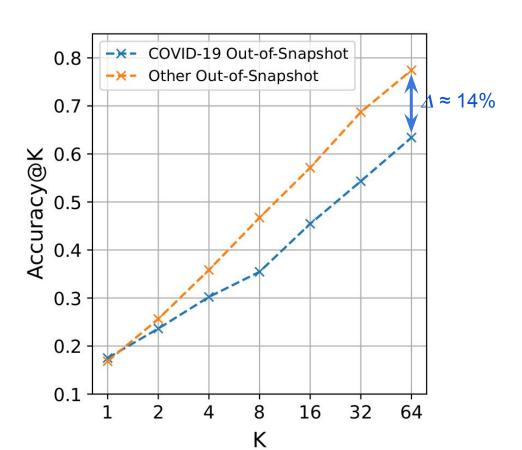
Sharp decrease in performance for new entities of 2021



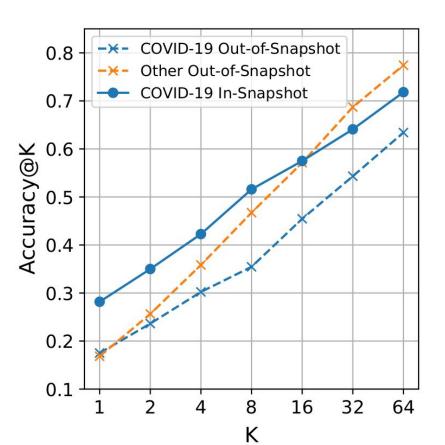
Low performance on COVID-19 new entities



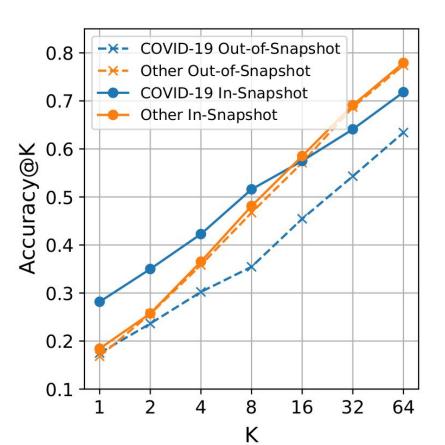
Low performance on COVID-19 new entities



Superior impact of in-snapshot finetuning on COVID-19 new entities



Superior impact of in-snapshot finetuning on COVID-19 new entities



TempEL, a new large-scale temporal entity linking dataset composed of 10 yearly snapshots.

- TempEL, a new large-scale temporal entity linking dataset composed of 10 yearly snapshots.
- TempEL enables to measure temporal drift in performance on continual and previously non-existent (new) entities.

- TempEL, a new large-scale temporal entity linking dataset composed of 10 yearly snapshots.
- TempEL enables to measure temporal drift in performance on continual and previously non-existent (new) entities.
- General temporal decrease in performance.

- TempEL, a new large-scale temporal entity linking dataset composed of 10 yearly snapshots.
- TempEL enables to measure temporal drift in performance on continual and previously non-existent (new) entities.
- General temporal decrease in performance.
- Finetuning has major effect on new entities requiring previously non-existent knowledge to be disambiguated (e.g., COVID-19).

- TempEL, a new large-scale temporal entity linking dataset composed of 10 yearly snapshots.
- TempEL enables to measure temporal drift in performance on continual and previously non-existent (new) entities.
- General temporal decrease in performance.
- Finetuning has major effect on new entities requiring previously non-existent knowledge to be disambiguated (e.g., COVID-19).
- Data https://github.com/klimzaporojets/TempEL.

TempEL: Linking Dynamically Evolving and Newly Emerging Entities

Klim Zaporojets¹, Lucie-Aimée Kaffee², Johannes Deleu¹, Thomas Demeester¹, Chris Develder¹, Isabelle Augenstein²

klimzaporojets@gmail.com



