Building a Commonsense Knowledge Graph

Filip Ilievski
ilievski@isi.edu
@earthling91

Outline

Commonsense knowledge

Constructing CSKG

Reasoning with CSKG

Implementation in KGTK

Common sense is the common knowledge about the world that is possessed by every schoolchild and the methods for making obvious inferences from this knowledge.

Davis, E. (2014). Representations of commonsense knowledge.

Commonsense knowledge includes the basic facts about events (including actions) and their effects, facts about knowledge and how it is obtained, facts about beliefs and desires. It also includes the basic facts about material objects and their properties.

McCarthy, J. (1989). Artificial intelligence, logic and formalizing common sense.

Common Sense Knowledge Sources

GenericsKB
[Bhakthavatsalam et al., 2020]

COMET
[Bosselut et al., 2019]

Atomic
[Sap et al., 2019]

Quasimodo KB
[Romero et al., 2019]

WebChild
[Tandon et al., 2014]

WebChild 2.0
[Tandon et al., 2017]

Open Mind Common Sense
[Minski, Singh, Havasi,1999]

ConceptNet
[Liu, Singh, 2004]

ConceptNet 5.5
[Speer et al., 2017]

NELL
[Carlson et al., 2010]

NELL
[Mitchell et al., 2015]

Wikidata
[Vrandečić, 2012]

Cyc
[Lenat et al., 1984]

OpenCyc 4.0
[Lenat 2012]

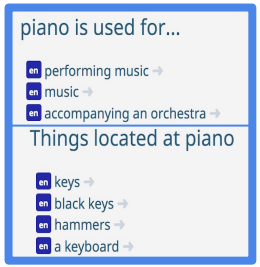
Categories of CSK sources

Category	Source	Relations	Example 1	Example 2
Commonsense KGs	ConceptNet*	34	*food - capable of - go rotten*	*eating - is used for - nourishment*
	ATOMIC	9	*Person X bakes bread - xEffect - eat food*	*PersonX is eating dinner - xEffect - satisfies hunger*
	GLUCOSE	10	*Someone$_A$ makes Something$_A$ (that is food)*	*Causes/Enables Someone$_A$ eats Something$_A$*
	WebChild	4 (groups)	*restaurant food - quality#n#1 - expensive*	*eating - type of - consumption*
	Quasimodo	78,636	*pressure cooker - cook faster - food*	*herbivore - eat - plants*
	SenticNet	4	*cold_food - polarity - negative*	*eating breakfast - polarity - positive*
	HasPartKB	1	*dairy food - has part - vitamin*	*n/a*
Common KGs	Wikidata	6.7k	*food - has quality - mouthfeel*	*eating - subclass of - ingestion*
	YAGO4	116	*banana chip - rdf:type - food*	*eating - rdfs:label - feeding*
	DOLCE*	1	*n/a*	*n/a*
	SUMO*	1,614	*food - hyponym - food_product*	*process - subsumes - eating*
Lexical resources	WordNet	10	*food - hyponym - comfort food*	*eating - part-meronym - chewing*
	Roget	2	*dish - synonym - food*	*eating - synonym - feeding*
	FrameNet	8 (f2f)	*Cooking_creation - has frame element - Produced_food*	*eating - evoke - Ingestion*
	MetaNet	14 (f2f)	*Food - has role - food_consumer*	*consuming_resources - is - eating*
	VerbNet	36 (roles)	*feed.v.01 - Arg1-PPT - food*	*eating - hasPatient - comestible*
Visual sources	Visual Genome	42,374	*food - on - plate*	*boy - is eating - treat*
	Flickr30k	1	*a food buffet - corefers with - a food counter*	*a eating place - corefers with - their kitchen*
Corpora & LMs	GenericsKB	n/a	*Aardvarks search for food.*	*Animals receive nitrogen by eating plants.*
	GPT-2	n/a	*Food causes a person to be hungry and a person to eat.*	*Eating at home will not lead to weight gain.*

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.

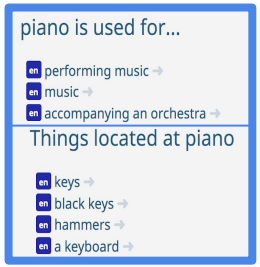
(Zellers et al., 2018)



ConceptNet: pianos have keys, are used to perform music

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.



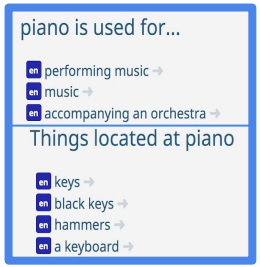
ConceptNet: pianos have keys, are used to perform music

- S: (n) **piano**, pianoforte, forte-piano (a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds)

WordNet: pianos are played by pressing keys

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.



ConceptNet: pianos have keys, are used to perform music

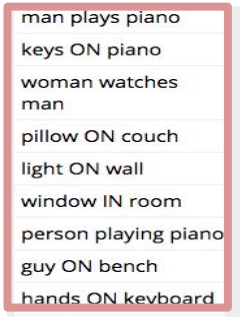
- S: (n) **piano**, pianoforte, forte-piano (a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds)

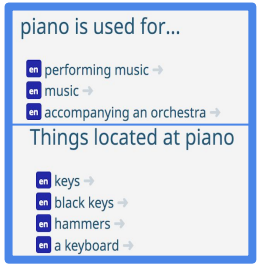
WordNet: pianos are played by pressing keys

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.

Visual Genome: person can play a piano while sitting, his hands are on the keyboard

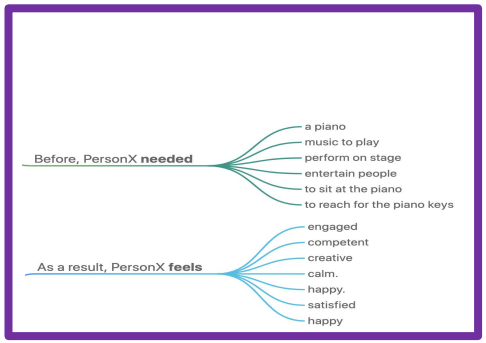




ConceptNet: pianos have keys, are used to perform music

- S: (n) **piano**, pianoforte, forte-piano (a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds)

WordNet: pianos are played by pressing keys

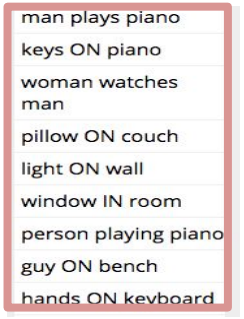


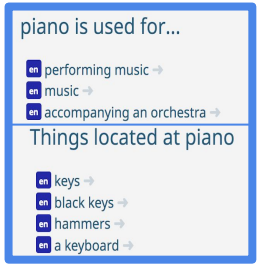
ATOMIC: to play piano, a person needs to sit at it, on stage and reach for the keys; feelings

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.

Visual Genome: person can play a piano while sitting, his hands are on the keyboard

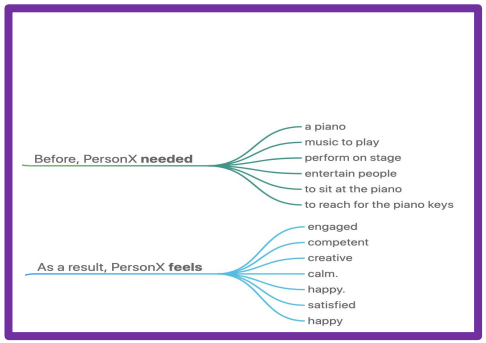




ConceptNet: pianos have keys, are used to perform music

- S: (n) **piano**, pianoforte, forte-piano (a keyboard instrument that is played by depressing keys that cause hammers to strike tuned strings and produce sounds)

WordNet: pianos are played by pressing keys



ATOMIC: to play piano, a person needs to sit at it, on stage and reach for the keys; feelings

On stage, a woman takes a seat at the piano. She

1. sits on a bench as her sister plays with the doll.
2. smiles with someone as the music plays.
3. is in the crowd, watching the dancers.
4. nervously sets her fingers on the keys.

FrameNet: performer entertains audience

Audience [Aud]	The Audience experiences the Performance.
Medium [Medium]	Medium is the physical entity or channel used by the Performer to transmit the Performance to the Audience.
Performance [Perance]	The Performer generates the Performance which the Audience perceives.
Performer [Perfer]	The Performer provides an experience for the Audience.

Visual Genome: person can play a piano while sitting, his hands are on the keyboard

- man plays piano
- keys ON piano
- woman watches man
- pillow ON couch
- light ON wall
- window IN room
- person playing piano
- guy ON bench
- hands ON keyboard

Outline

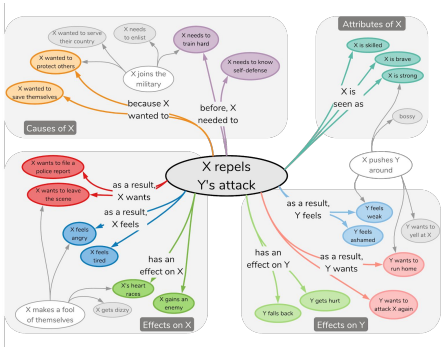
Commonsense knowledge

Constructing CSKG

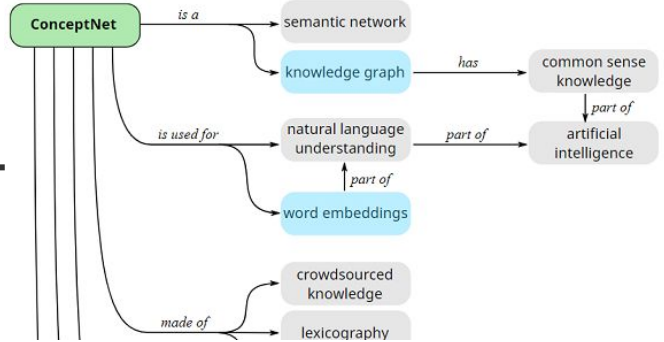
Reasoning with CSKG

Implementation in KGTK

Consolidating Knowledge Graphs



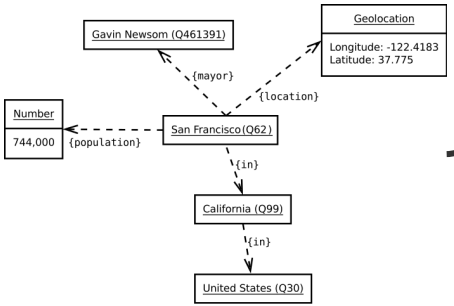
ATOMIC (Sap et al. 2019)



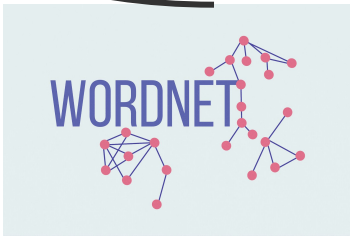
ConceptNet (Speer, Chin and Havasi 2017)

CSKG

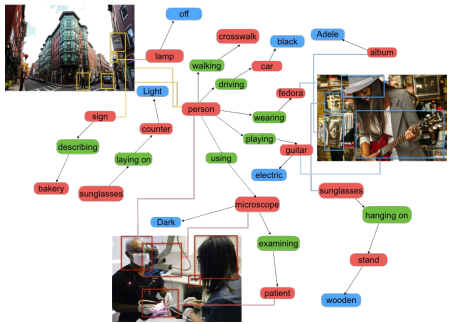
CSKG (Ilievski et al. 2021)



Wikidata (Vrandecic and Krotzsch 2014)



WordNet (Miller 1995)



Visual Genome (Krishna et al. 2017)

F Ilievski, P Szekely, B Zhang (2021). CSKG: The CommonSense Knowledge Graph. ESWC

Principles for a modular and useful CSKG

P1. Embrace heterogeneity of nodes

objects, classes, words, actions, frames, states

P2. Reuse edge types across resources

/r/HasProperty from ConceptNet applicable for attributes in Visual Genome

P3. Leverage external links

many sources map to WordNet

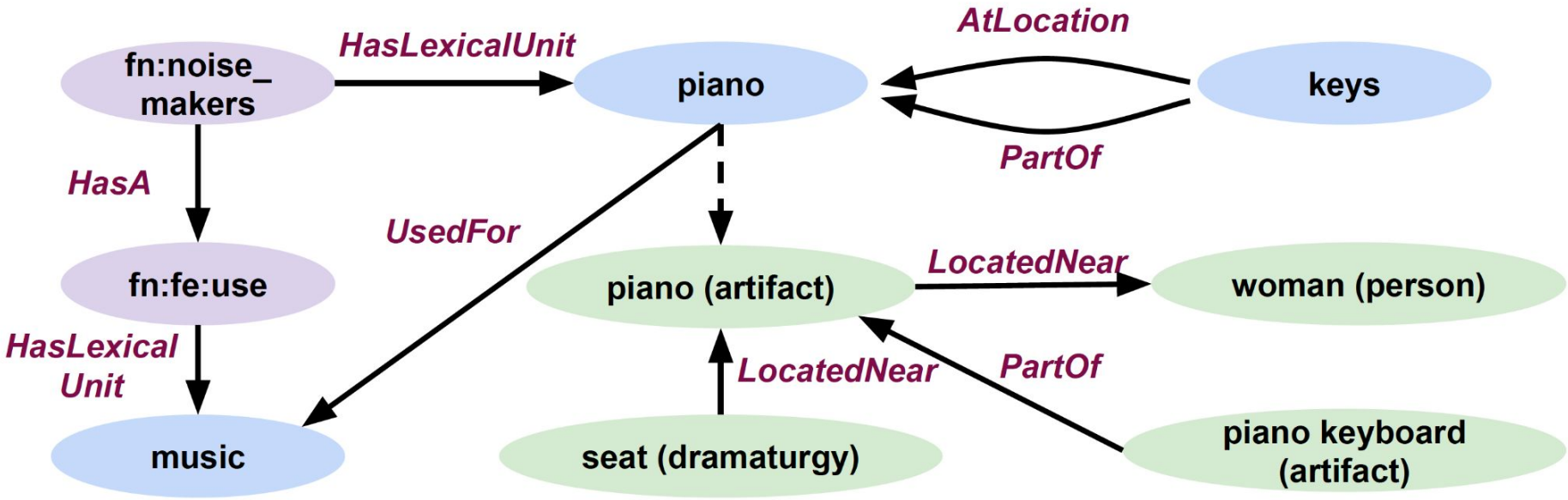
P4. Generate high-quality probabilistic links

many facts not explicitly stated

P5. Enable access to labels

text labels and aliases are the key, in particular for NLP use cases

Consolidated knowledge



CSKG statistics

#nodes	2,160,968
#edges	6,001,531
#relations	58
mean degree	5.55
std degree	0.03

Integration statistics

	AT	CN	FN	RG	VG	WD	WN	CSKG*	CSKG
#nodes	304,909	1,787,373	15,652	71,804	11,264	91,294	71,243	2,414,813	**2,160,968**
#edges	732,723	3,423,004	29,873	1,403,955	2,587,623	111,276	101,771	6,349,731	**6,001,531**
#relations	9	47	9 (23)	2	3 (42k)	3	15 (45)	59	**58**
avg degree	4.81	3.83	3.82	39.1	459.45	2.44	2.86	5.26	**5.55**
std degree	0.07	0.02	0.13	0.34	35.81	0.02	0.05	0.02	**0.03**

Top PageRank nodes

1. /c/en/chromatic/a/wn
2. /c/en/organic_compound
3. /c/en/chemical_compound/n
4. /c/en/change/n/wn/artifact
5. /c/en/natural_science/n/wn/cognition

Outline

Commonsense knowledge

Constructing CSKG

Reasoning with CSKG

Implementation in KGTK

CSKG embeddings

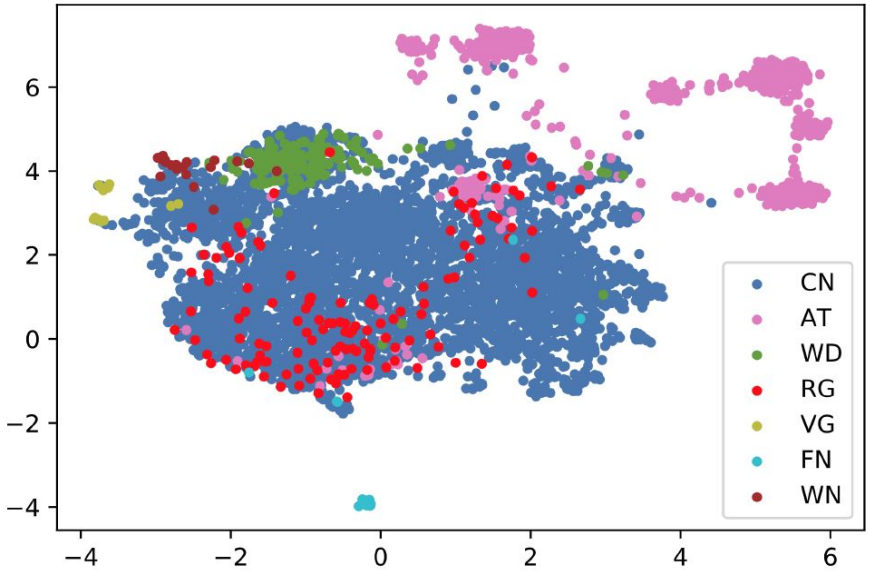
Graph embeddings

a. TransE
b. ComplEx
c. RESCAL
d. DistMult

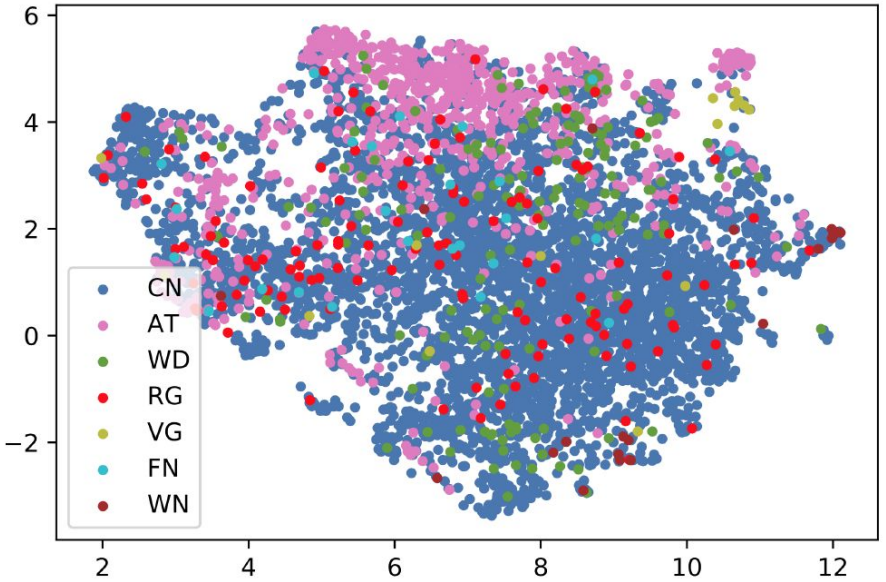
Text embeddings

a. BERT-large of lexicalized nodes

What do the embeddings capture?



Graph embeddings - structural similarity



Text embeddings - lexical similarity

Most similar nodes

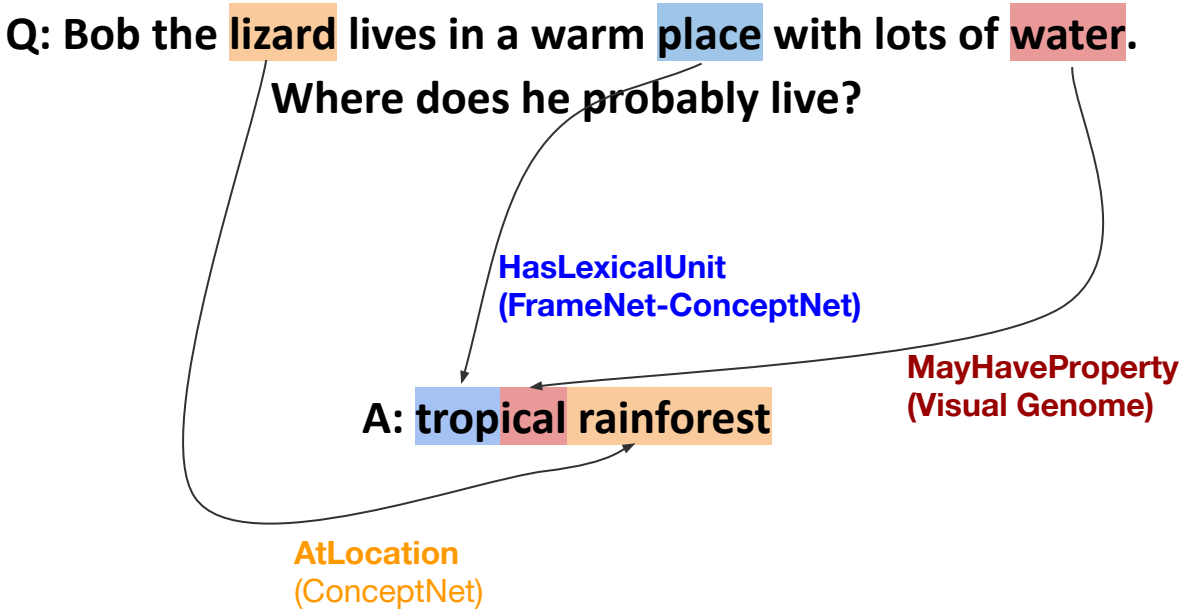
TransE		BERT
/c/en/chelonian/n/wn/animal	/c/en/turtle/n/wn/animal	/c/en/glyptemys/n
/c/en/mud_turtle/n/wn/animal		/c/en/pelocomastes/n
/c/en/cooter/n/wn/animal		/c/en/staurotypus/n
/c/en/common_snapping_turtle/n/wn/animal		/c/en/parahydraspis/n
/c/en/sea_turtle/n/wn/animal		/c/en/trachemys/n
/c/en/excited	/c/en/happy	/c/en/bring_happiness
/c/en/satisfied		/c/en/new_happiness
/c/en/smile_mood		at:like_a_party_is_a_good_way_to_...
/c/en/pleased		/c/en/encouraging_person's_talent
/c/en/joyful		at:happy_that_they_went_to_the_party

Finding evidence in CSKG

Q: Bob the lizard lives in a warm place with lots of water. Where does he probably live?

A: tropical rainforest

Finding evidence in CSKG



Pre-training LMs with CSKG questions

CSKG

Question: losing weight is for?
A: being healthier **B**. embedded software **C.** buying things at store

SocialIQA test data

CSQA test data

PhysicalIQA test data

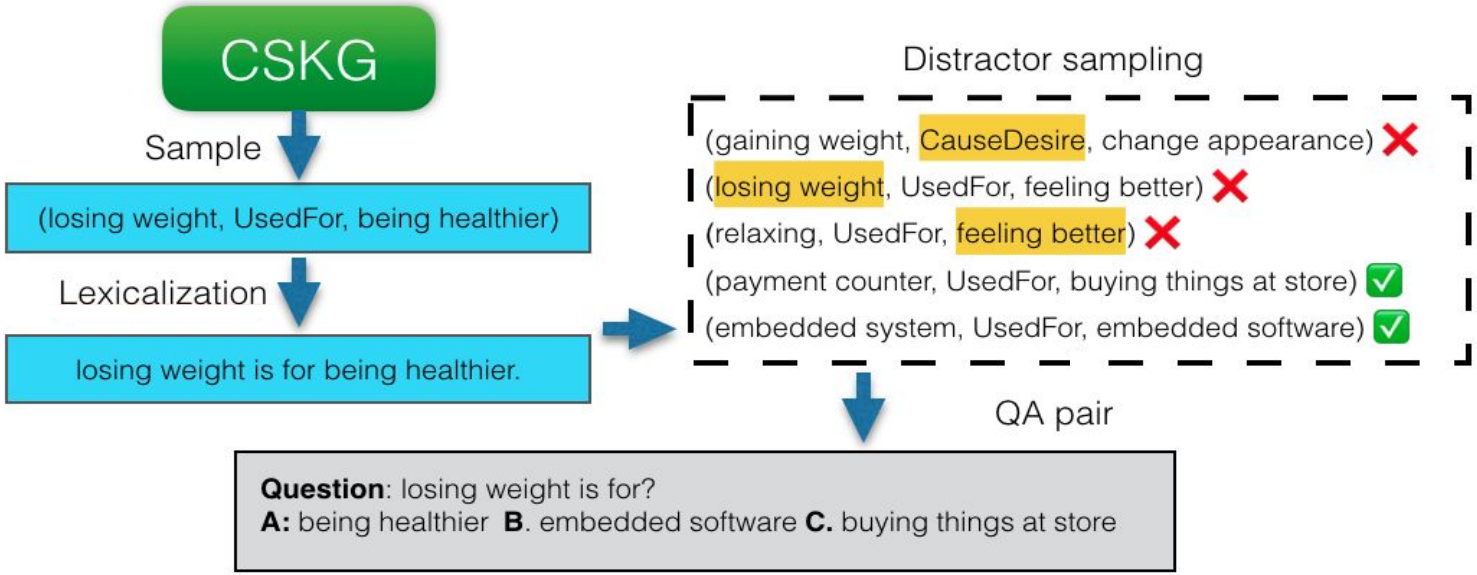
Winogrande test data

aNLI test data

Generating questions with CSKG

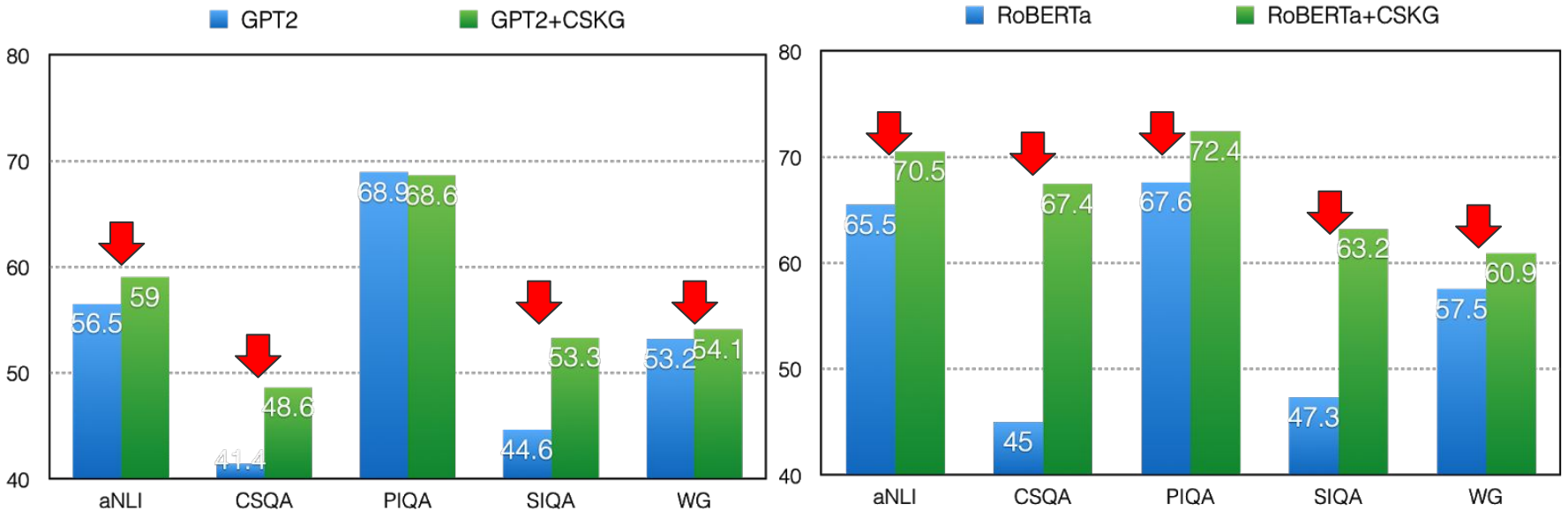
Pretrain LMs with synthetic QA sets generated from CSKG

Answer commonsense questions on unseen datasets (zero-shot QA)



K Ma, F Ilievski, J Francis, Y Bisk, E Nyberg, A Oltramari (2021).
Knowledge-driven Data Construction for Zero-shot Evaluation in Commonsense Question Answering. In AAAI

Pretraining on CSKG questions helps accuracy



K Ma, F Ilievski, J Francis, Y Bisk, E Nyberg, A Oltramari (2021).
Knowledge-driven Data Construction for Zero-shot Evaluation in Commonsense Question Answering. In AAAI

Outline

Commonsense knowledge

Constructing CSKG

Reasoning with CSKG

Implementation in KGTK

Implementation in KGTK

Notebook

CSKG Resources

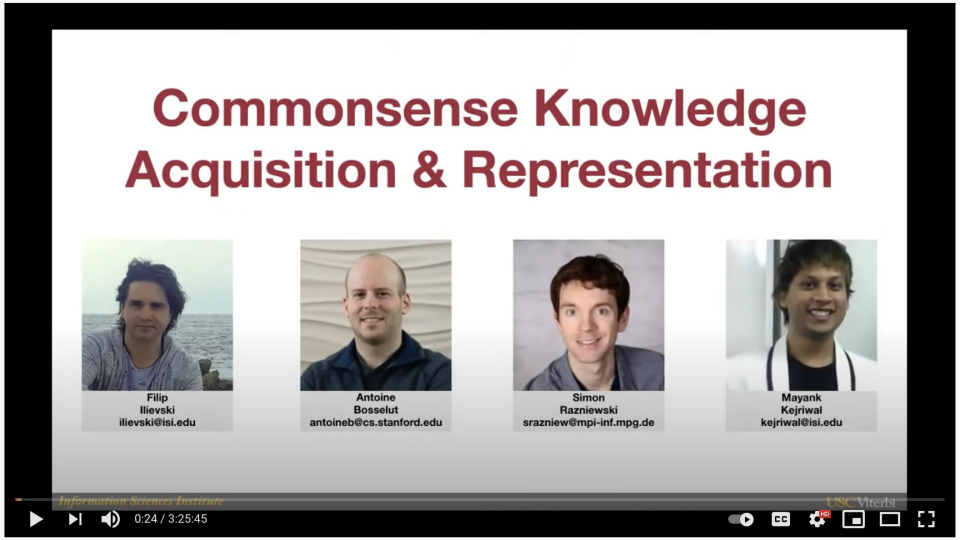
Graph on Zenodo

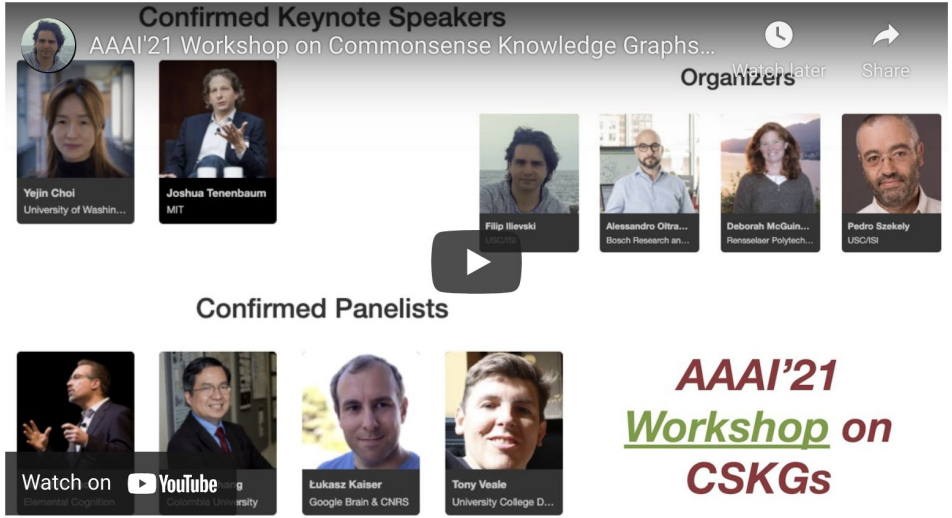
Documentation

Graph & text embeddings

Code on GitHub

Recent events on machine common sense





Special SWJ Issue on Commonsense Knowledge and Reasoning

Guest editors:

- **Filip Ilievski, USC Information Sciences Institute, CA, USA**
- **Antoine Bosselut, EPFL, Switzerland**
- **Kenneth Forbus, Northwestern University, IL, USA**
- **Simon Razniewski, Max Planck Institute for Informatics, Germany**
- **Vered Shwartz, Allen Institute for AI and University of Washington, WA, USA**

Deadline: October 30th, 2021

Contact: commonsense-swj@googlegroups.com