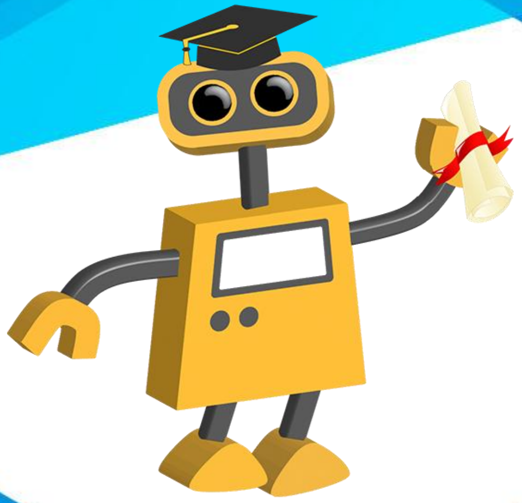
On the Role of Knowledge Graphs and Language Models in Machine Understanding of Scientific Documents

Jose Manuel Gomez-Perez
Director, Language Technology Research
expert.ai



Tackling Increasingly Complex Scientific Phenomena

Evolution of the scientific enterprise from [Barabasi, 2005] extended with the ATLAS Detector Project at the Large Hadron Collider [The ATLAS Collaboration, 2012].

The Imperative for AI in Science

Future AI Systems as Partners for Discovery
[Gil DSJ 2017]

Thoughtful AI: Principles for Partnership	
Rationality	Behavior is governed by explicit knowledge structures
Context	Seek to understand the purpose and scope of tasks
Initiative	Proactively learn new knowledge relevant to their task
Networking	Access external sources of knowledge and capabilities
Articulation	Respond with persuasive justifications and arguments
Systems	Facilitate integration & collaboration with humans/systems
Ethics	Behavior that conveys scope and uncertainty

These are important research challenges for AI

Yolanda Gil. *Will AI Write the Scientific Papers of the Future?* AAAI 2020 presidential address

AI that **assists** scientists

AI that **understands** scientific content

AI that **does** science

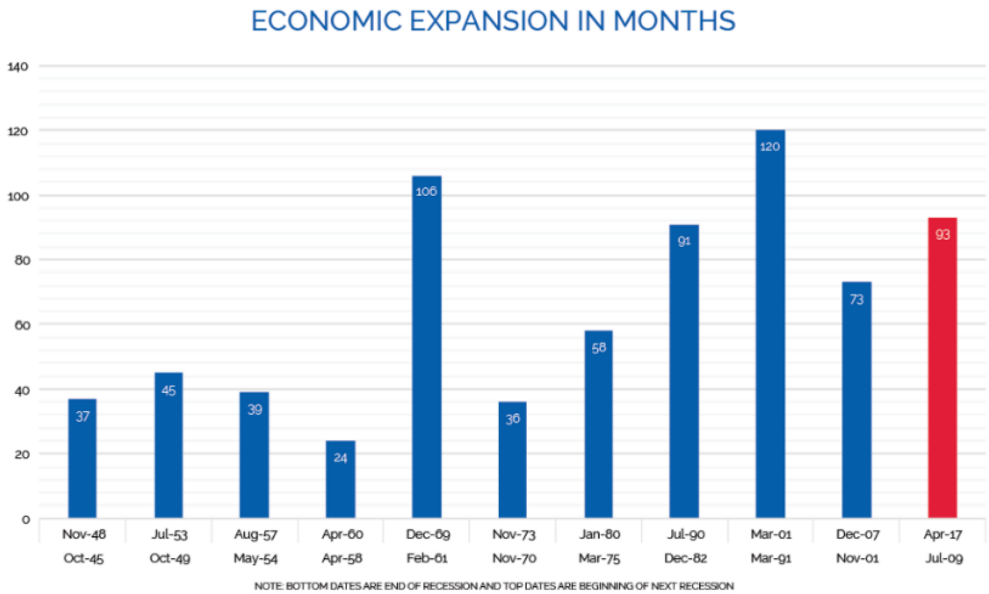
Machine Reading Comprehension

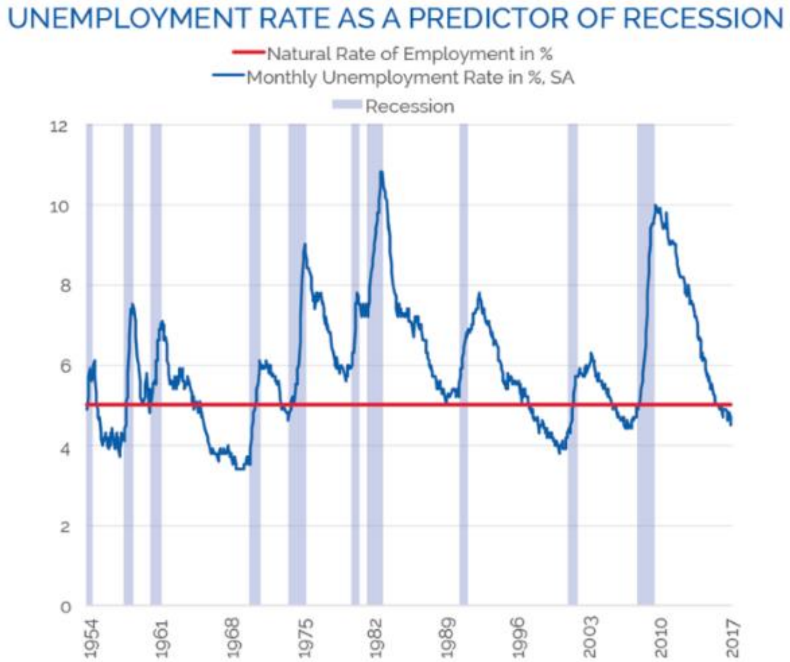
”Reading a chapter in a college freshman text (say physics or accounting) and answering the questions at the end of the chapter is a hard AI problem that requires advances in vision, language, problem-solving, and learning theory.”

Raj Reddy. *Foundations and Grand Challenges of Artificial Intelligence.* AAAI 1988 presidential address.

How do humans reason?

Is Another Recession Coming?





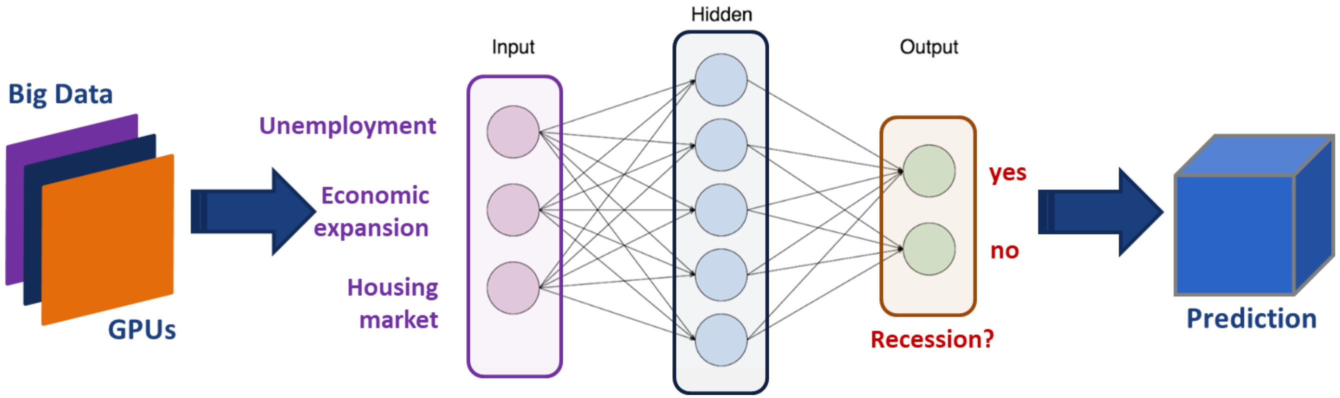
[..] the length of the current recovery – 93 months as of April 2017 – is the third longest of the 11 expansion periods since the end of World War II. Should the recovery last past May 2018, it would surpass the 106-month expansion of 1961-1969. It would match the longest period, 120 months, in July 2019. **Looking simply at the cycle, one could say the due date is near.**

[..] **when the unemployment rate falls below 5% [..] a recession usually follows within the next two or three years.** For example, the unemployment rate started sinking below 5% in December 2005, and the Great Recession started in January 2008 — 25 months later.

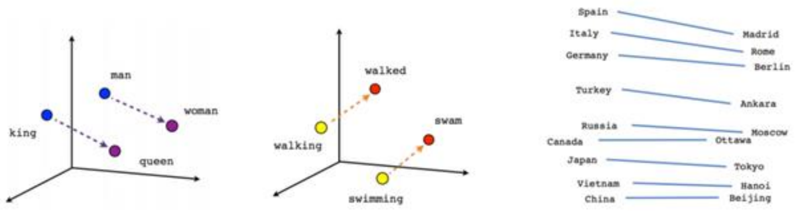
Forbes

The statistic/neural approach

Is Another Recession Coming?

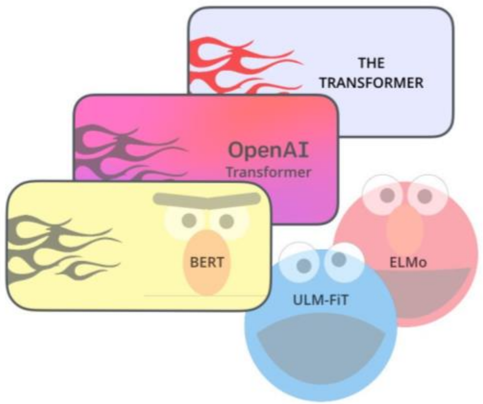


Word embeddings



Male-Female Verb tense Country-Capital

Language models



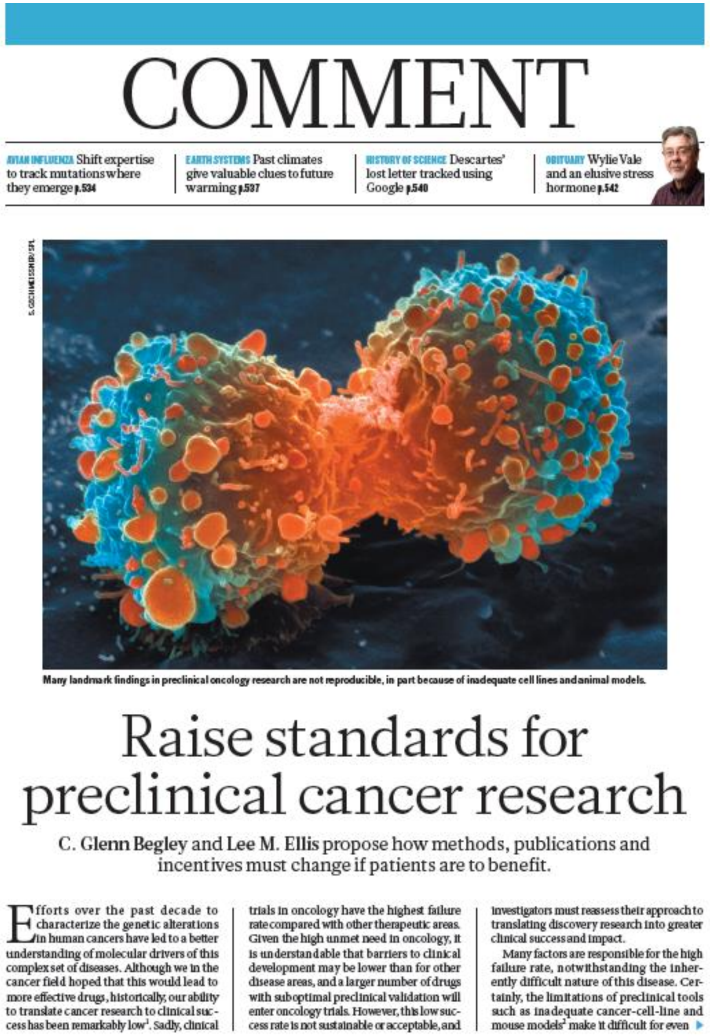
- **Pros**
  - **Grounded on the data**
  - **Broad, flexible, scalable**
  - **SotA** in most NLP/NLU benchmarks
- **Cons**
  - **Black box:** Induction, not logical explanation
  - **Lack of true understanding** of real-world semantics and pragmatics
  - **Risk of bias** if training data not carefully curated

Explainability is important

nature International weekly journal of science

https://www.nature.com/articles/483531a

47 of 53 landmark publications in cancer research could not be reproduced


COMMENT

AVIAN INFLUENZA Shift expertise to track mutations where they emerge p.534

EARTH SYSTEMS Past climates give valuable clues to future warming p.537

HISTORY OF SCIENCE Descartes' lost letter tracked using Google p.540

OBITUARY Wylie Vale and an elusive stress hormone p.542

Many landmark findings in preclinical oncology research are not reproducible, in part because of inadequate cell lines and animal models.

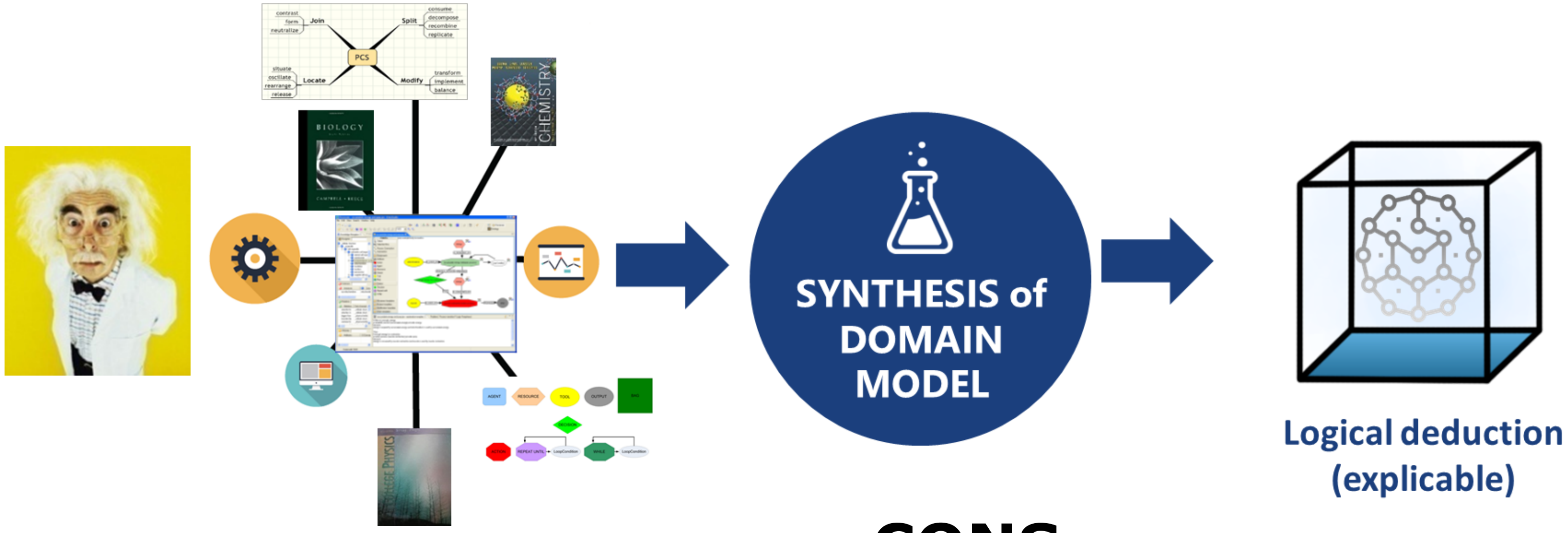
Raise standards for preclinical cancer research

C. Glenn Begley and Lee M. Ellis propose how methods, publications and incentives must change if patients are to benefit.

Efforts over the past decade to characterize the genetic alterations in human cancers have led to a better understanding of molecular drivers of this complex set of diseases. Although we in the cancer field hoped that this would lead to more effective drugs, historically, our ability to translate cancer research to clinical success has been remarkably low[1]. Sadly, clinical trials in oncology have the highest failure rate compared with other therapeutic areas. Given the high unmet need in oncology, it is understandable that barriers to clinical development may be lower than for other disease areas, and a larger number of drugs with suboptimal preclinical validation will enter oncology trials. However, this low success rate is not sustainable or acceptable, and investigators must reassess their approach to translating discovery research into greater clinical success and impact.

Many factors are responsible for the high failure rate, notwithstanding the inherently difficult nature of this disease. Certainly, the limitations of preclinical tools such as inadequate cancer-cell-line and mouse models[2] make it difficult for even

The knowledge-based approach



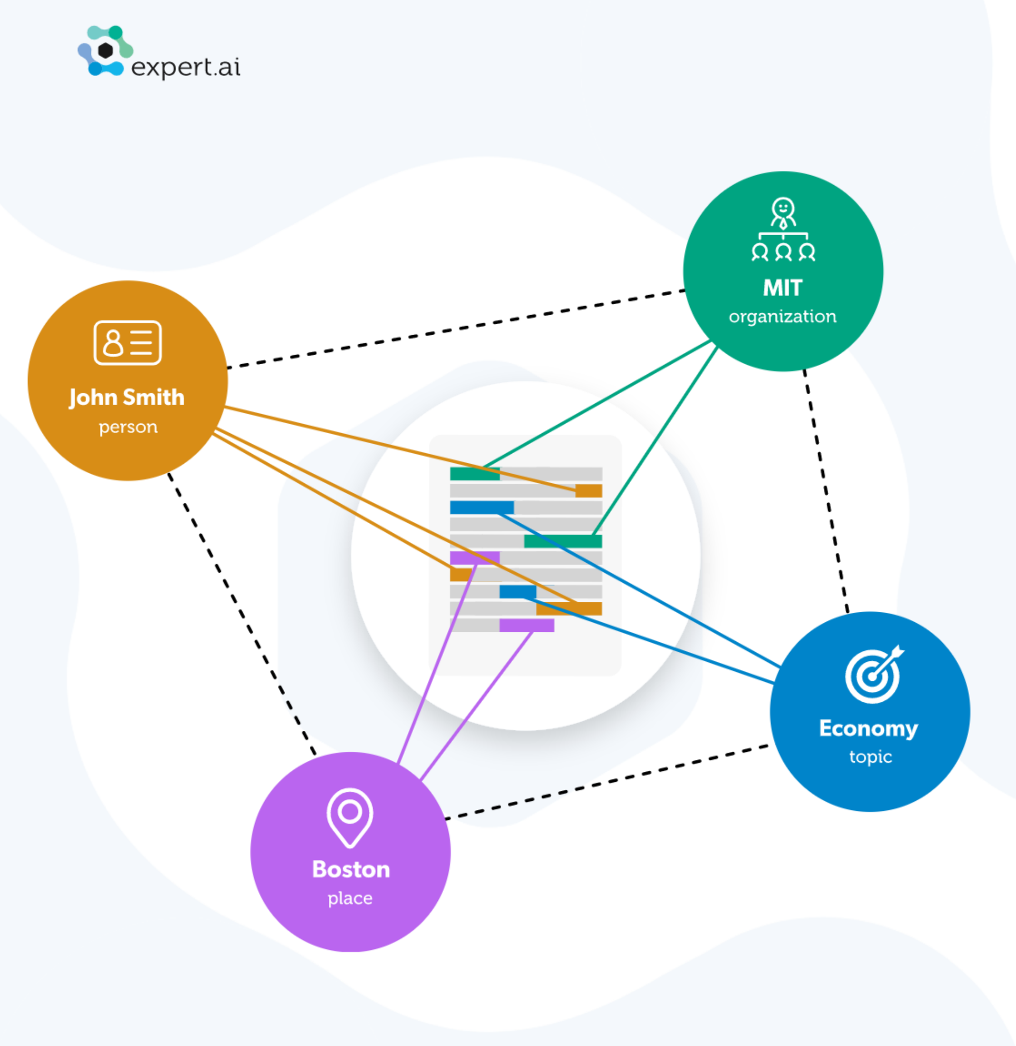
PROS

- Based on highly curated resources
- No need for training, just a few examples (in theory)
- Logically interpretable, explainable
- Structured representations are great for tasks like WSD
- Good modeling tools available

CONS

- Representations can be rich and deep but also rigid and brittle
- Automation can be challenging
- Well trained labor needed to manually model a domain can be expensive
- May be hard to scale

Expert.ai



Developer Portal
Start using the API

Documentation
How to use the API

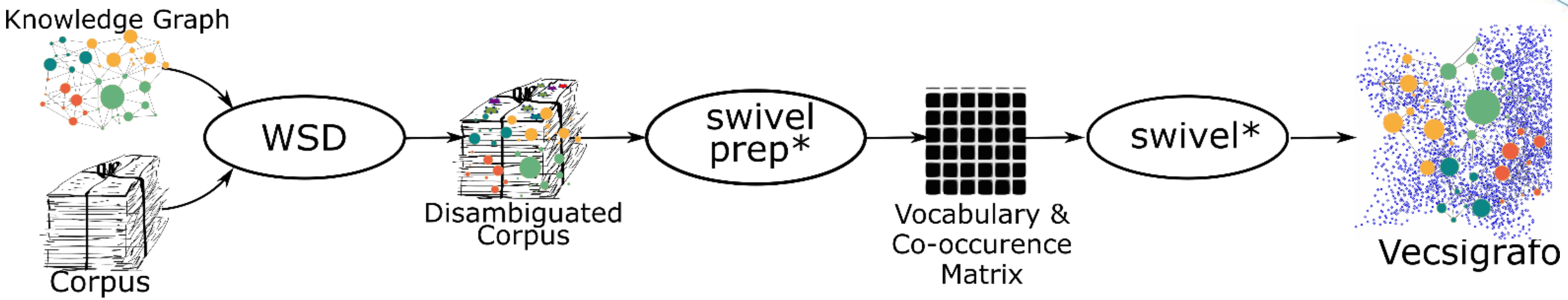
expert.ai Natural Language API

Add language intelligence to your application right now! expert.ai Natural Language API provides deep language understanding without any IT infrastructure or installation, and scales with your needs so you can start developing intelligent applications today!

Live Demo

Sign Up or Login

Vecsigrafo



- Unlike Knowledge Graph Embeddings, Vecsigrafo
  - **Combines corpus-based and graph-based** approaches
  - **Jointly learns word and concept embeddings**
- Considers **both lexical and semantic entries** as part of the vocabulary
  - The corpus is lemmatized following different tokenization strategies, disambiguated and expanded with grammatical (PoS) and semantic information
  - Word and semantic (lemma and concept from the KG) embeddings are then jointly learnt
- For our experiments, we use **Sensigrafo, expert.ai's KG**
  - 300K concepts, 400K lemmas and 80+ relation types 2.8 million links) per language (14)
  - Other lexical KGs like WordNet can be equally used

UMBC WebBase Corpus

SN SciGraph

Semantic Scholar

R Denaux, JM Gomez-Perez. 2019. *Vecsigrafo: Corpus-based word-concept embeddings.* Semantic Web Journal, *10*, 881-908.

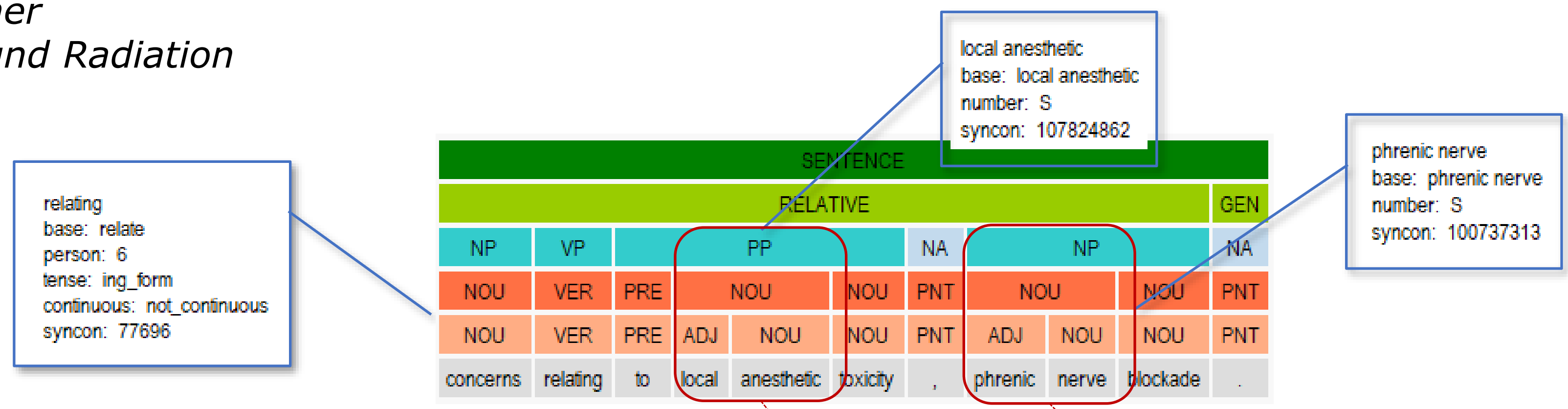
Scientific terminology & tokenization

- **Domain-specific terminology requires specialized lexical resources**
- **Homonymy is frequent, also with named entities (Galileo the space probe vs. the scientist)**
- **Multiple-word expressions requires specific tokenization**
  - *Molecularly imprinted polymer*
  - *Cosmic Microwave Background Radiation*
  - …

PoS	example	#swe	#mwe
ADJ	single+dose	104.7	0.514
ADV	in+situ	21	1.98
CON	even+though	34.6	3.46
ENT	august+2007	1.5	1.24
NOU	humid+climate	216.9	15.24
NPH	Ivor+Lewis	0.917	0.389
NPR	dorsal+ganglia	22.8	5.22
PRO	other+than	13.89	0.005
VER	take+place	69.39	0.755

Millions of single vs. multiple expressions

SciGraph



[…] concerns relating to local anesthetic toxicity, phrenic nerve blockade, and pneumothorax.

Concept	en%2326973	en%2377696	-	en%23107824862	en%23100274160	-	en%23100737313	en%23101569578
Grammar	N	V	P	NP	N	PNT	NP	N
Lemma	concern	relate	to	local anesthetic	toxicity	,	phrenic nerve	blockade
Surface Form	concerns	relating	to	local anesthetic	toxicity	,	phrenic nerve	blockade
Token	concerns	relating	to	local anesthetic	toxicity	,	phrenic nerve	blockade

A Garcia-Silva, R Denaux, JM Gomez-Perez. 2019. Learning Embeddings from Scientific Corpora using Lexical, Grammatical and Semantic Information. 3rd Intl. Workshop on Capturing Scientific Knowledge (SciKnow'19)

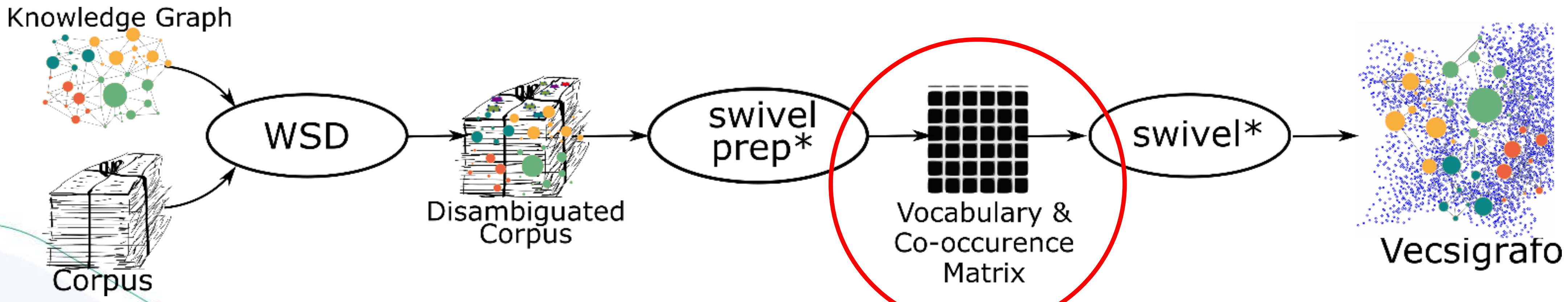
Vecsigrafo - Encoding

A Novel WASP Gene Mutation in a Chinese Boy with Wiskott-Aldrich Syndrome
Hui Wu[1], Cheng Hu[2], Dan Dang[1], Ying-Jie Guo[3]

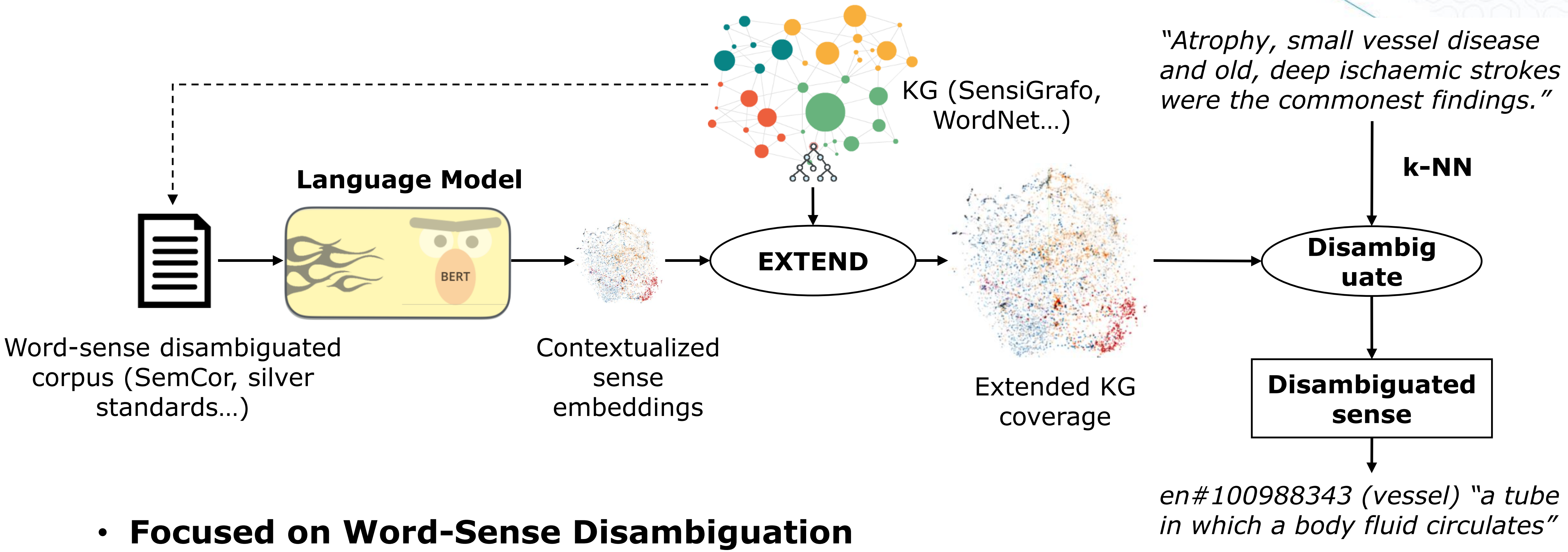
PubMed.gov
https://pubmed.ncbi.nlm.nih.gov/25332617/

```
t       a novel wasp gene mutation in a chinese boy with wiskott-aldrich syndrome

sf_l_c  a novel|lem_novel|en%23100008261 wasp|lem_Wiskott-Aldrich+syndrome+protein
        gene+mutation|lem_gene+mutation|en%23101415380 in a chinese|lem_Chinese|en%2398003
        boy|lem_boy|en%2346011 with wiskott-aldrich+syndrome|lem_Wiskott-Aldrich+syndrome
```



Transigrafo: Transformers + KGs

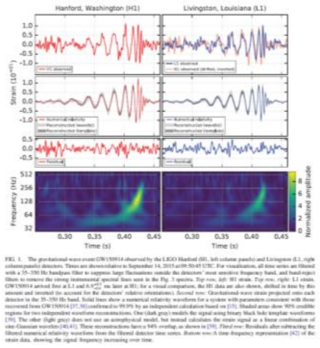


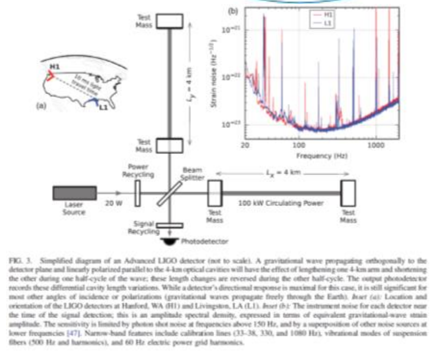
- **Focused on Word-Sense Disambiguation**
- **Decouples language and knowledge representations,** allowing for parallel development by independent, possibly unrelated teams
  - **LM** models human language and how sentences are built
  - **KG** human-engineered, interpretable conceptualization of a domain

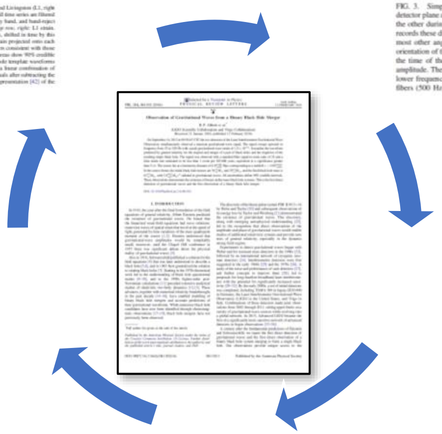
Transigrafo based *on work by Loureiro and Jorge:* ***Language modelling makes sense: Propagating representations through WordNet for full-coverage word sense disambiguation.*** *In: Proceedings of the 57th Annual Meeting of the ACL, pp. 5682–5691.*

Reading a scientific document

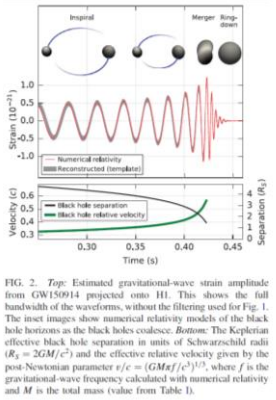
- **The scientific discourse usually adopts the form of a narrative**
- **However, we rarely read a whole paper sequentially**
  - We may start with the abstract, then check figures and tables to get an idea of the methods and experimental results, and iterate until we acquire an overall understanding
- **To facilitate human understanding,** scientific information is represented in mutually supportive ways across modalities
- **This entails understanding text, but also figures, diagrams and tables**

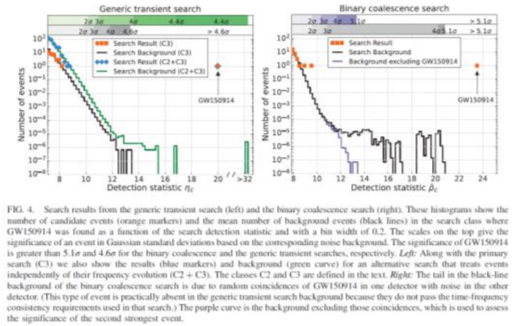






[Abbott et al., 2017] Observation of Gravitational Waves from a Binary Black Hole Merger





Relating scientific language and visual information

Captions are a source of FREE supervision!

Definition
What is this?

Description
What is the experimental setting?

Supporting information
What are the details?

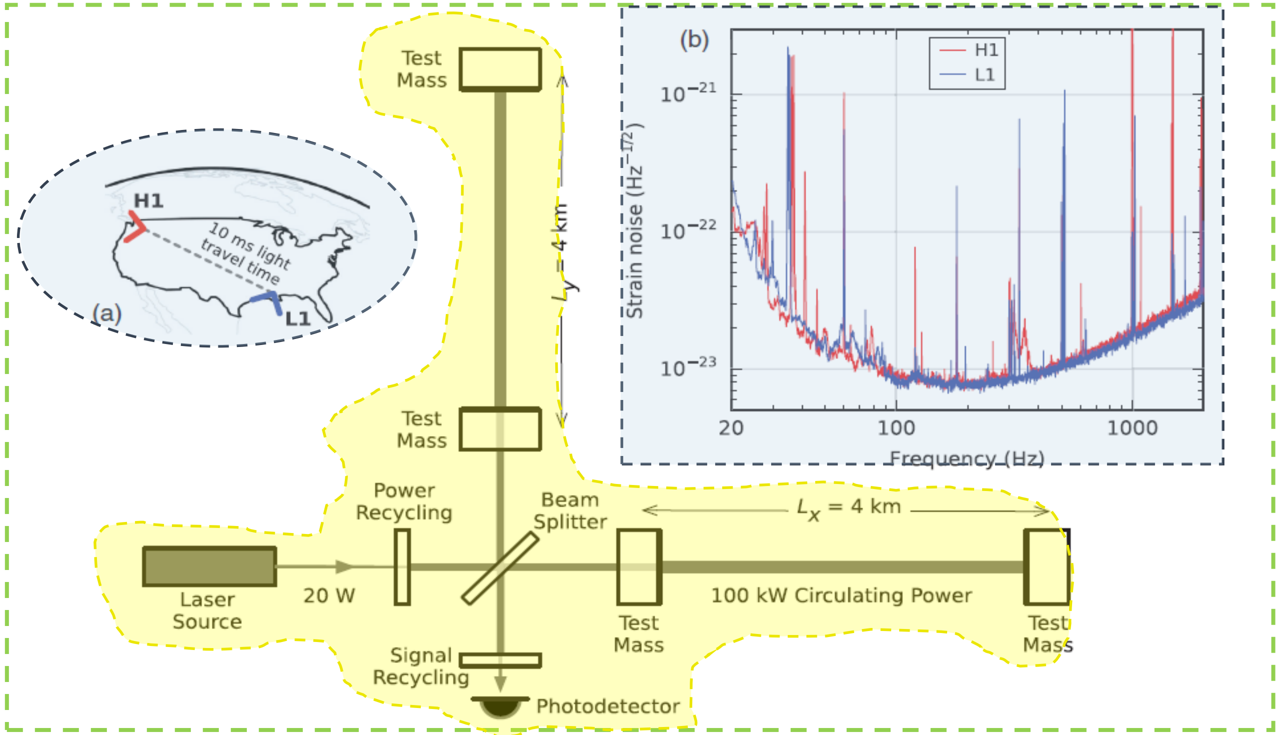


FIG. 3. Simplified diagram of an Advanced LIGO detector (not to scale). A gravitational wave propagating orthogonally to the detector plane and linearly polarized parallel to the 4-km optical cavities will have the effect of lengthening one 4-km arm and shortening the other during one half-cycle of the wave; these length changes are reversed during the other half-cycle. The output photodetector records these differential cavity length variations. While a detector's directional response is maximal for this case, it is still significant for most other angles of incidence or polarizations (gravitational waves propagate freely through the Earth). *Inset (a):* Location and orientation of the LIGO detectors at Hanford, WA (H1) and Livingston, LA (L1). *Inset (b):* The instrument noise for each detector near the time of the signal detection; this is an amplitude spectral density, expressed in terms of equivalent gravitational-wave strain amplitude. The sensitivity is limited by photon shot noise at frequencies above 150 Hz, and by a superposition of other noise sources at lower frequencies [47]. Narrow-band features include calibration lines (33–38, 330, and 1080 Hz), vibrational modes of suspension fibers (500 Hz and harmonics), and 60 Hz electric power grid harmonics.

Figure-Caption Correspondence

How can we tap on this source of supervision?

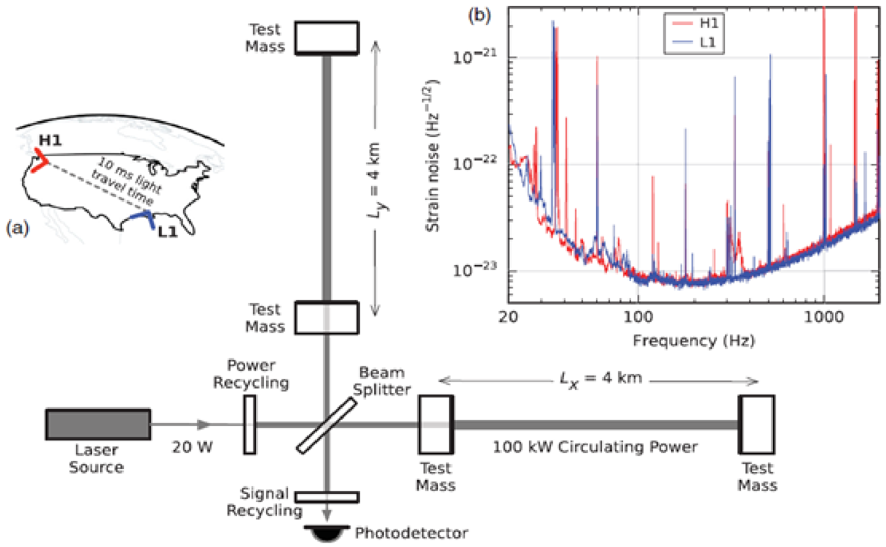
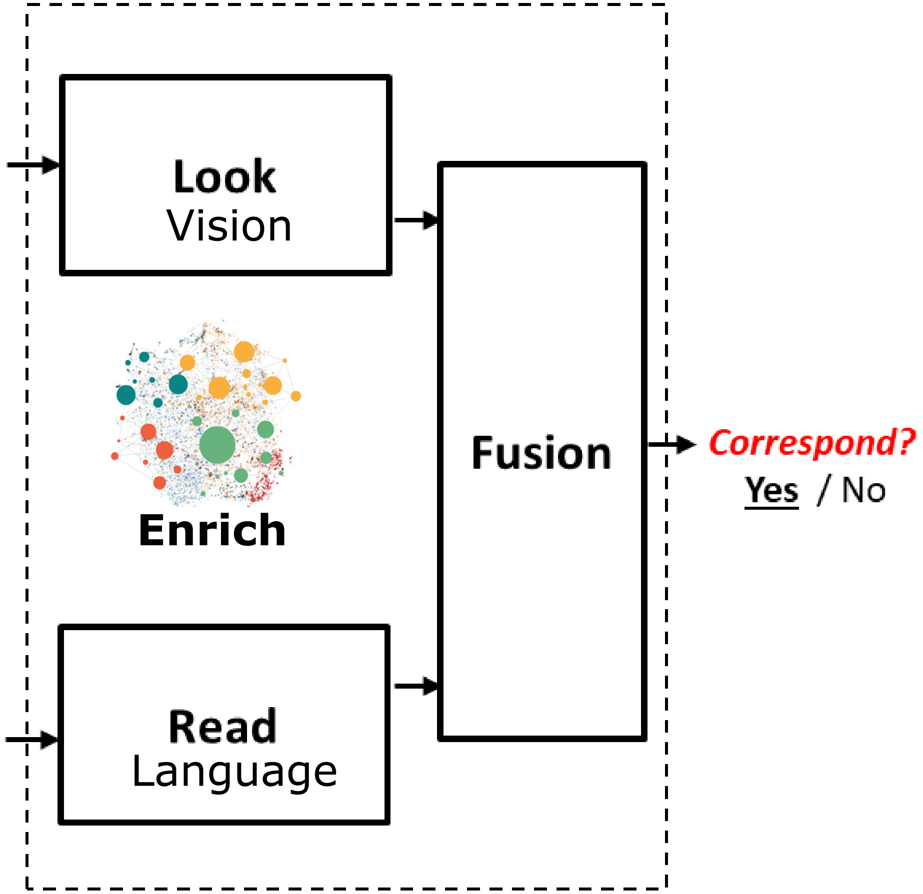


FIG. 3. Simplified diagram of an Advanced LIGO detector (not to scale). A gravitational wave propagating orthogonally to the detector plane and linearly polarized parallel to the 4-km optical cavities will have the effect of lengthening one 4-km arm and shortening the other during one half-cycle of the wave; these length changes are reversed during the other half-cycle. The output photodetector records these differential cavity length variations. While a detector's directional response is maximal for this case, it is still significant for most other angles of incidence or polarizations (gravitational waves propagate freely through the Earth). *Inset (a):* Location and orientation of the LIGO detectors at Hanford, WA (H1) and Livingston, LA (L1). *Inset (b):* The instrument noise for each detector near the time of the signal detection; this is an amplitude spectral density, expressed in terms of equivalent gravitational-wave strain amplitude. The sensitivity is limited by photon shot noise at frequencies above 150 Hz, and by a superposition of other noise sources at lower frequencies [47]. Narrow-band features include calibration lines (33–38, 330, and 1080 Hz), vibrational modes of suspension fibers (500 Hz and harmonics), and 60 Hz electric power grid harmonics.



JM Gomez-Perez and R Ortega. 2019. *Look, Read and Enrich - Learning from Scientific Figures and their Captions.* 10th International Conference on Knowledge Capture (K-CAP '19). Association for Computing Machinery, New York, NY, USA, 101–108

Look, Read... and Enrich

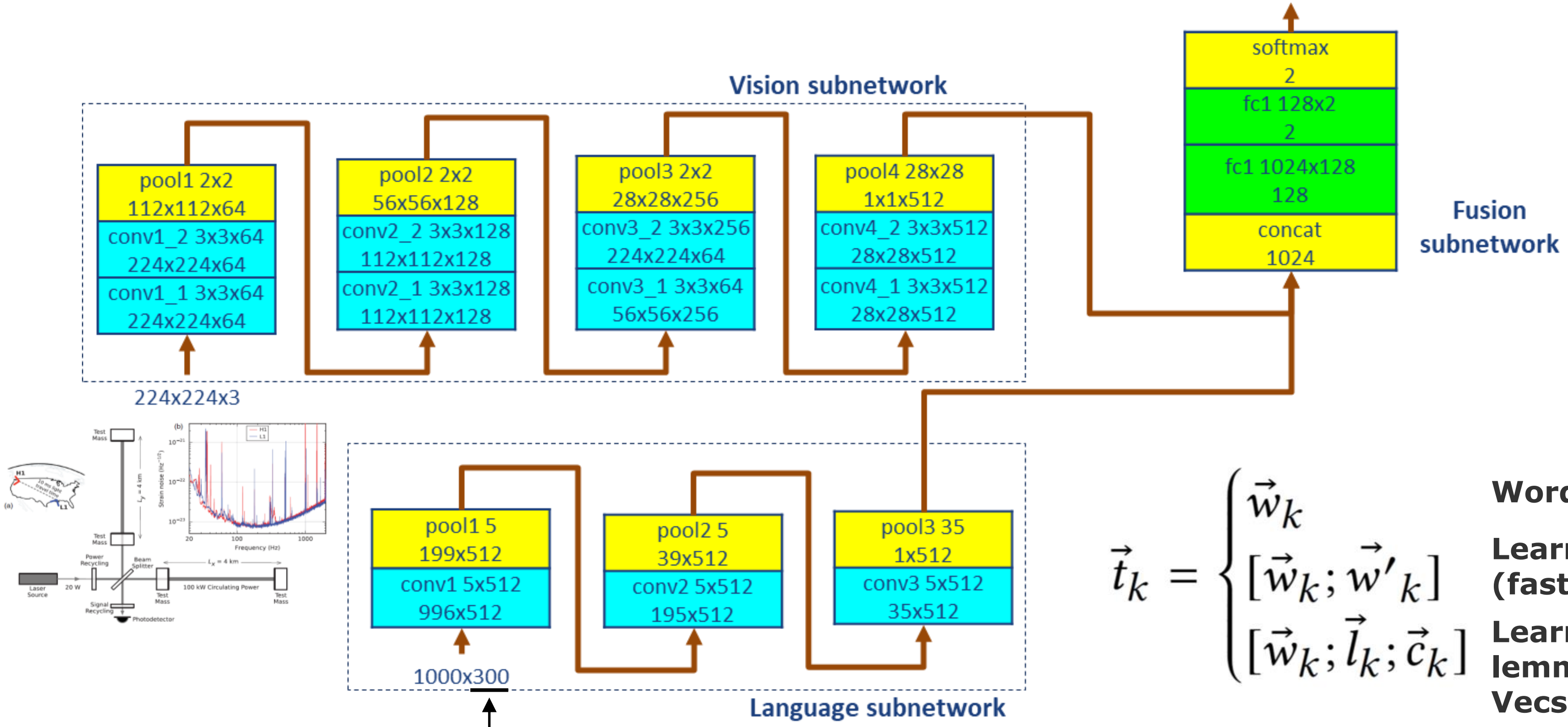


FIG. 3. Simplified diagram of an Advanced LIGO detector (not to scale). A grav detector plane and linearly polarized parallel to the 4-km optical cavities will have the e the other during one half-cycle of the wave; these length changes are reversed durin records these differential cavity length variations. While a detector's directional respons most other angles of incidence or polarizations (gravitational waves propagate fre orientation of the LIGO detectors at Hanford, WA (H1) and Livingston, LA (L1). *Inse* the time of the signal detection; this is an amplitude spectral density, expressed i

$\vec{t}_k$ = Input word d-D embedding, D=300

$$\vec{t}_k = \begin{cases} \vec{w}_k \\ [\vec{w}_k ; \vec{w'}_k] \\ [\vec{w}_k ; \vec{l}_k ; \vec{c}_k] \end{cases}$$

Word embeddings learnt by the network

Learnt and pre-trained word embeddings (fastText)

Learnt word embeddings, pre-trained lemma and concept embeddings (HolE, Vecsigrafo)

SN SciGraph

7M publications

Semantic Scholar

39M papers

Some Experimental Results

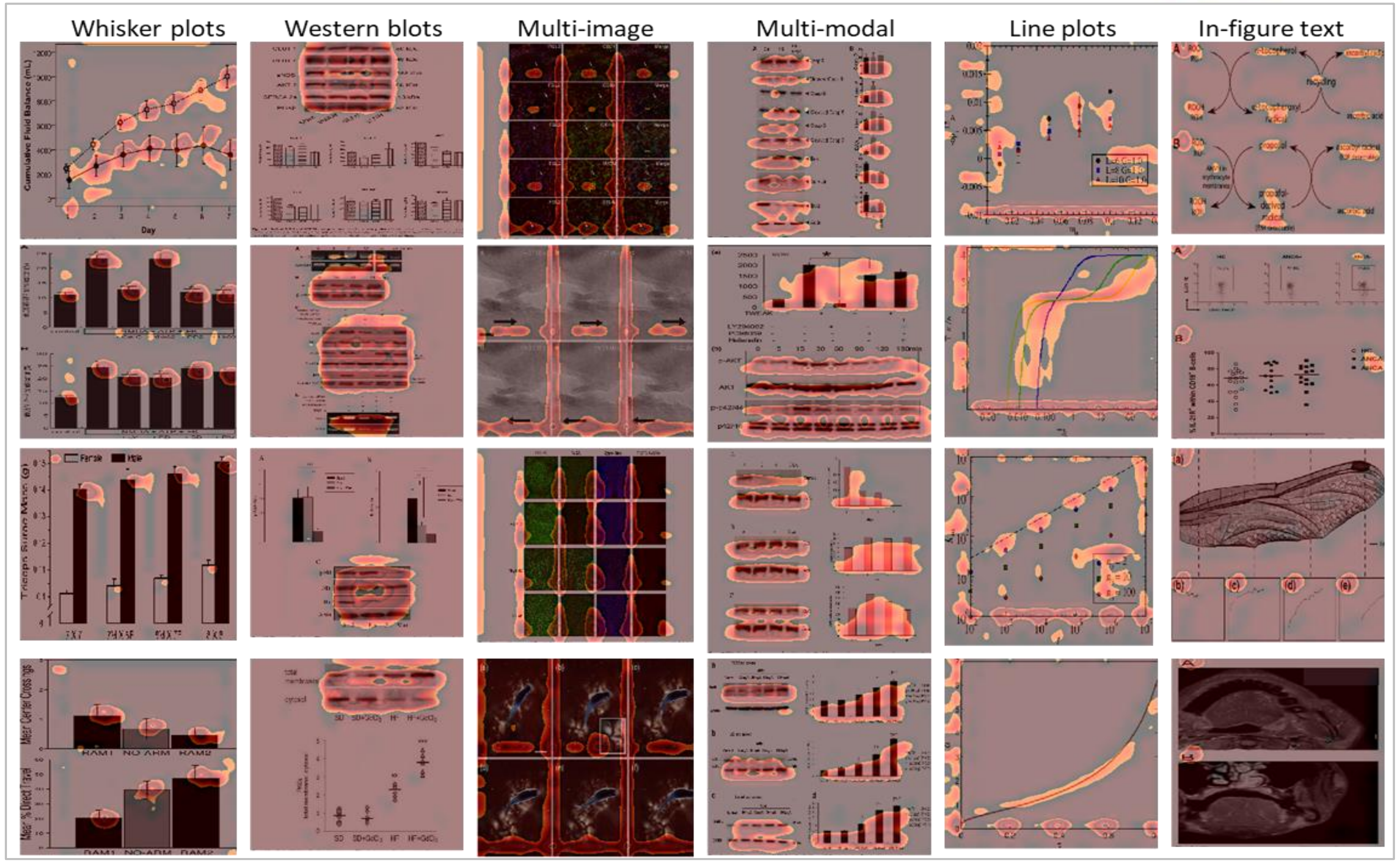
Figure-caption correspondence

	Corpus	Word rep.	Acc_{vgg}.	Acc.
Direct	SciGraph	$\vec{w}_k$	60.30	
Pre-train		$\vec{w}_k$	68.40	
FCC_1	SciGraph	$\vec{w}_k$	78.09	78.48
FCC_2		$[\vec{w}_k;\vec{w'}_{k\_sem}]$	79.75	80.35
FCC_3		$[\vec{w}_k;\vec{l}_{k\_holE};\vec{c}_{k\_holE}]$	78.64	78.08
FCC_4		$[\vec{w}_k;\vec{l}_{k\_wiki};\vec{c}_{k\_wiki}]$	79.71	80.50
FCC_5		$[\vec{w}_k;\vec{l}_{k\_sem};\vec{c}_{k\_sem}]$	**80.50**	**81.97**
FCC_6	SemScholar	$\vec{w}_k$	80.42	81.44
FCC_7		$[\vec{w}_k;\vec{l}_{k\_sem};\vec{c}_{k\_sem}]$	**82.21**	**84.34**

Caption and figure classification

Model	Caption		Figure	
	Non-trainable	Trainable	Non-trainable	Trainable
Random	39.92	78.20	44.19	61.21
VGG16	n/a	n/a	58.43	n/a
Ours FCC6	61.31	**79.24**	58.57	**63.60**
Ours FCC7	**67.40**	79.11	**60.19**	63.49

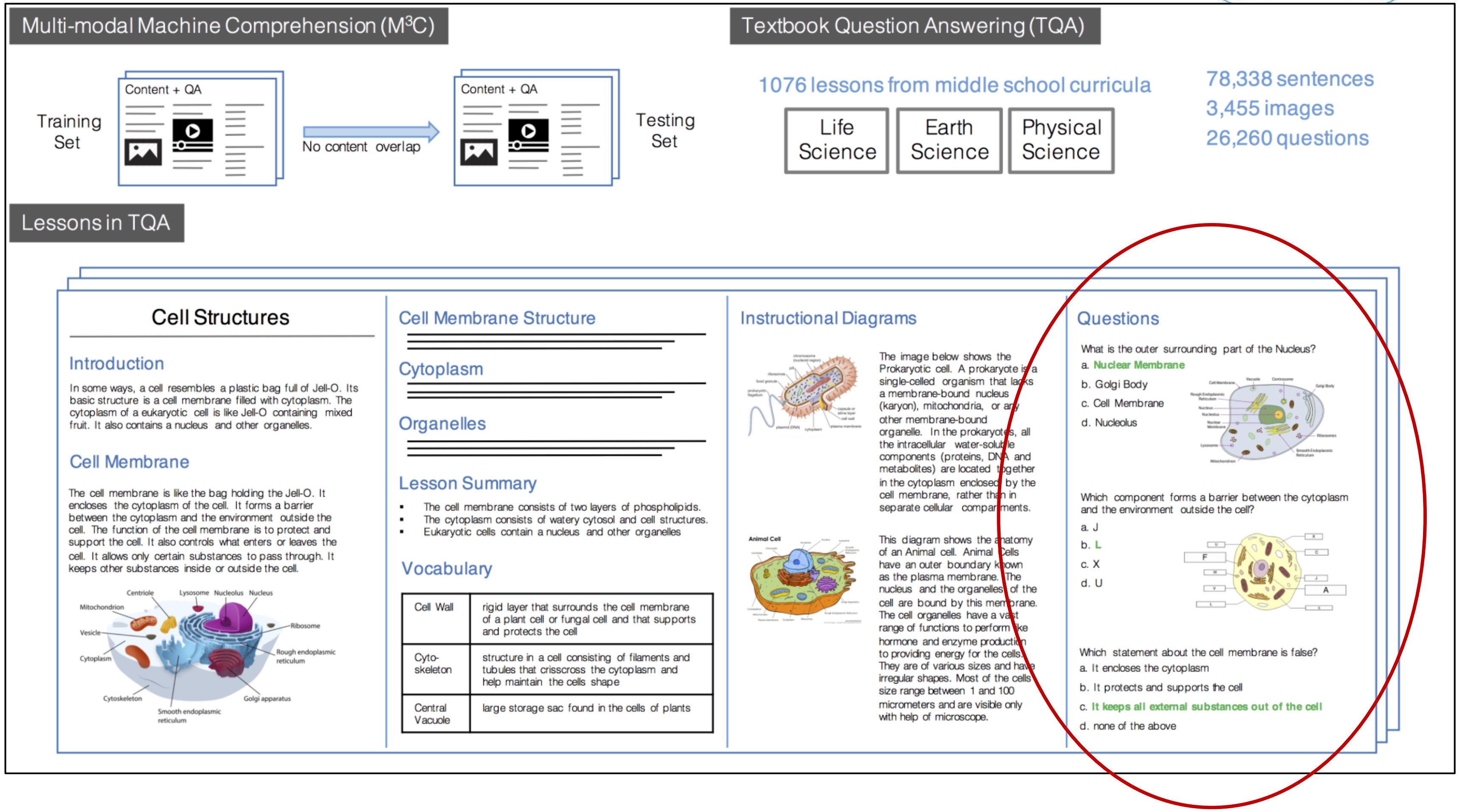
Emerging visual patterns



- **Vecsigrafo boosts performance** in FCC and downstream tasks
- FCC leads to detailed **textual and visual discrimination** through cross-modal learning

Textbook Question Answering

- TQA describes complex scientific phenomena through a combination of text and diagrams
- Answering TQA questions may involve language, visual information or both
- Often, TQA questions cannot be answered by correlation or lookup
- Negation, conjunction, polarity, qualities (high frequency), quantities (20.000 Hz)
- Diagrams describe concepts hard to represent in a single natural image, like mitosis
- Also, they comprise constituents and relationships whose semantics needs to be captured



TQA facts: 10% questions formulated assertively, 20% multi-paragraphs or multi-lesson, 40% diagram questions require complex diagram parsing, only 2% can be answered with an OCR.

TQA SotA

Model	Text T/F	Text MC	Text All	Diagram MC	All
Random	50,10	22,88	33,62	24,96	29,08
MemN+VQA	50,50	31,05	38,73	31,82	35,11
MemN+DPG	50,50	30,98	38,69	32,83	35,62
BiDAF+DPG	50,40	30,46	38,33	32,72	35,39
FCC+Vecsigrafo	-	40,21	-	35,30	-
IGMN	57,41	40,00	46,88	36,35	41,36
f-GCN1+SSOC	**62,73**	**49,54**	**54,75**	**37,61**	**45,77**

Nice progress, but still poor performance!

Mastering TQA with ISAAQ

ISAAQ: Intelligent System for Automatically Answering Textbook Questions

Background retrieval

Solve

Ensemble

ISAAQ leverages transformers and cross-modal attention

Pre-training on related datasets is key

- Text: RACE, ARC-Easy, ARC-Challenge, OpenBookQA
- Diagrams: VQA abstract scenes, AI2D

JM Gomez-Perez, R Ortega. 2020. ISAAQ - mastering textbook questions with pretrained transformers and bottom-up and top-down attention. 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP), pp 5469–5479. Association for Computational Linguistics

Background retrieval

IR

- Scope: whole textbook
- *What is the most related sentence to the question?*

"Wave erosion threatens many homes and beaches on the ocean. Deposits by waves include beaches. (…) Wave-cut cliffs form when waves erode a rocky shoreline. (…)"

Information retrieval

NSP

- Scope: question lesson
- *What is the most likely sentence following the question?*

"Erosion by waves can create unique landforms (figure 10.12) such as wave-cut cliffs, sea arches, and sea stacks. Other wave deposits are spits, sand bars, and barrier islands. (…)"

Next sentence prediction

NN

- Scope: question lesson
- *What is the most similar sentence to the question?*

"Deposits by waves include beaches. (…) wave-cut cliffs form when waves erode a rocky shoreline. wave erosion threatens many homes and beaches on the ocean. (…)"

Nearest neighbors

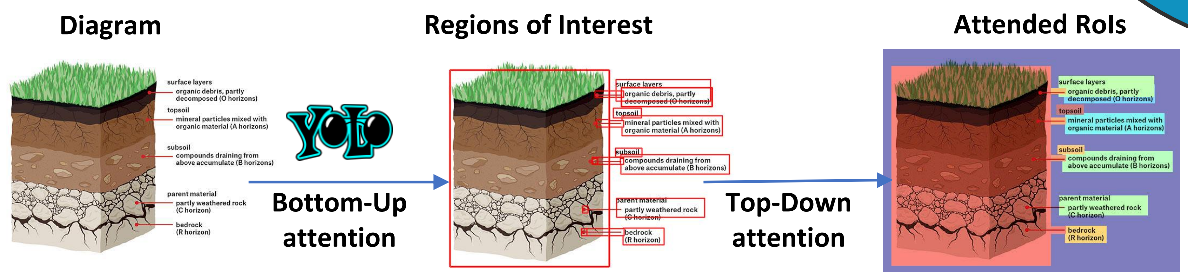
Solvers

TQA divided in three individual sub-tasks, one for each question type

TQA

Diagram MC

Multimodal Multiple-Choice Classification task with transformers and BUTD (Bottom-Up and Top-Down) attention



True/False

Entailment task

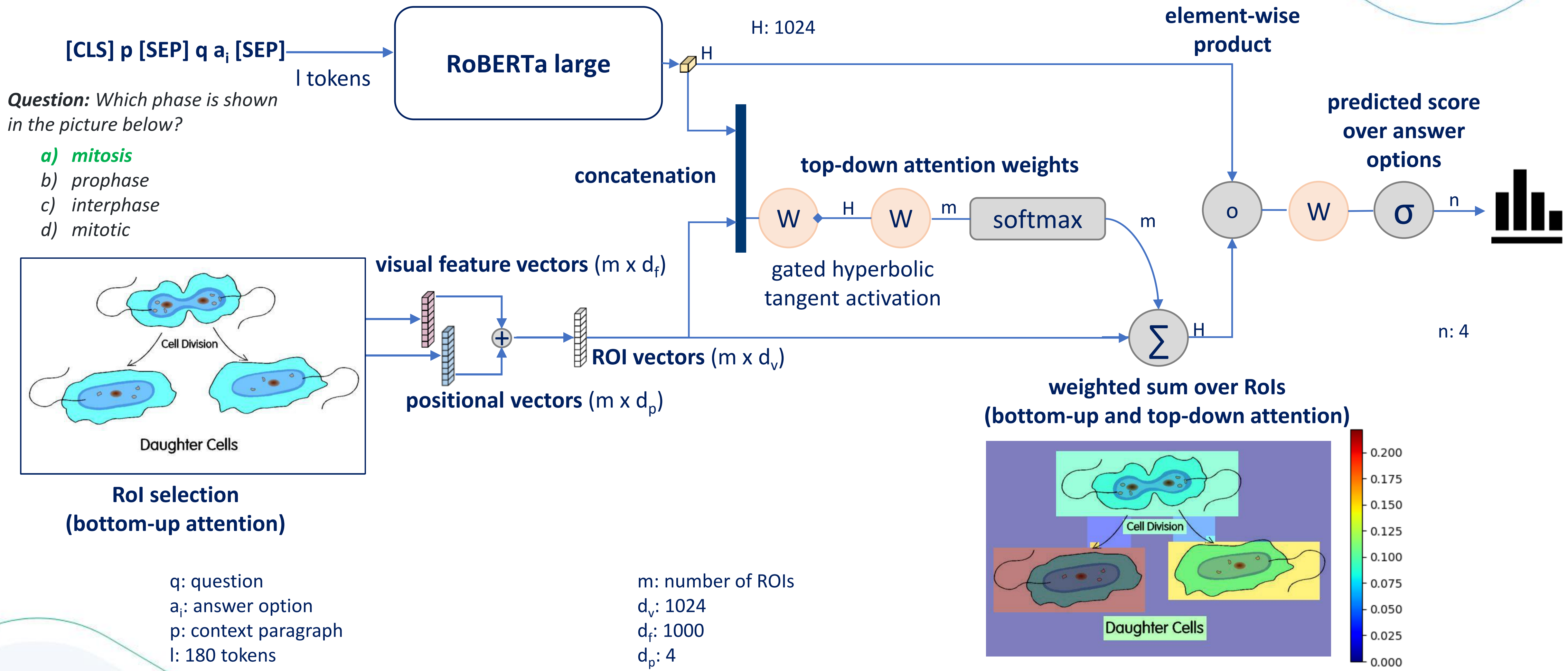
Background (Premise)
Question (Hypothesis)
Transformer
True or False

Text MC

Text Multiple-Choice Classification task

Background + Question + Answer a
Background + Question + Answer b
Background + Question + Answer c
Background + Question + Answer d
Transformer
Correct Answer

Merging Language and Vision

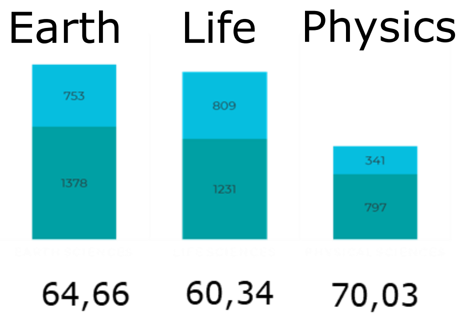


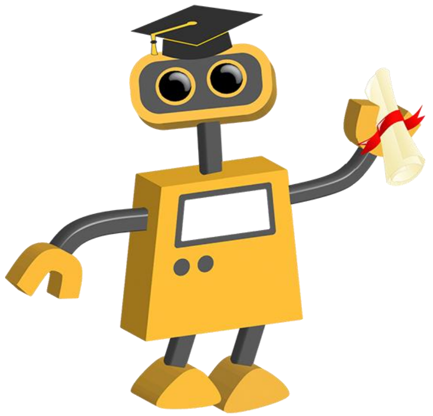
q: question
a_i: answer option
p: context paragraph
l: 180 tokens

m: number of ROIs
d_v: 1024
d_f: 1000
d_p: 4

Results

Model	Text T/F	Text MC	Text All	Diagram MC	All
Random	50,10	22,88	33,62	24,96	29,08
MemN+VQA	50,50	31,05	38,73	31,82	35,11
MemN+DPG	50,50	30,98	38,69	32,83	35,62
BiDAF+DPG	50,40	30,46	38,33	32,72	35,39
FCC+Vecsigrafo	-	36,56	-	35,30	-
IGMN	57,41	40,00	46,88	36,35	41,36
f-GCN1+SSOC	62,73	49,54	54,75	37,61	45,77
RoBERTa + VQA	76,85	62,81	68,38,	41,14	54,09
ISAAQ	**81,36**	**71,11**	**75,16**	**55,12**	**64,66**





Average 19% accuracy points over the previous SotA
10% points over a RoBERTa + VQA baseline (14% in Diagram MC)

Some examples

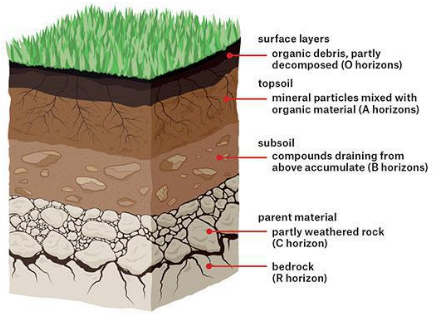
ISAAQ discriminates the key visual information necessary to answer the question

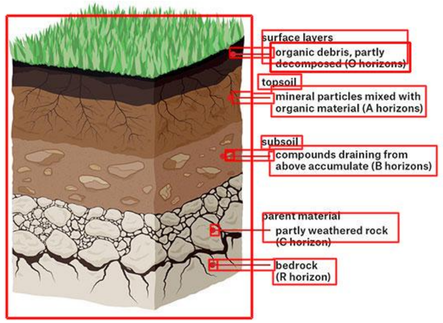
expert.ai

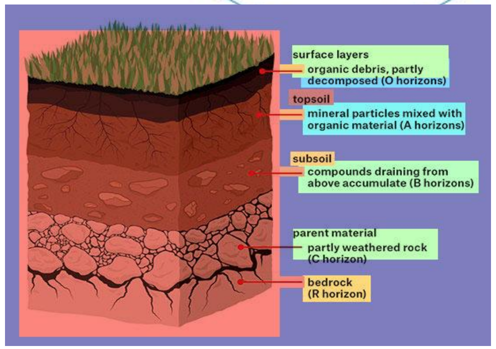
Question	Diagram	Regions	Attention

Which of the following layers comprise mineral particles?

a) bedrock
b) subsoil
c) surface layers
d) topsoil

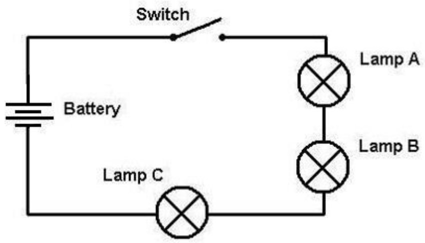


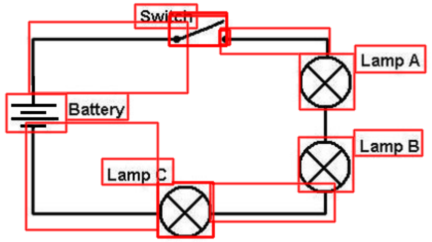


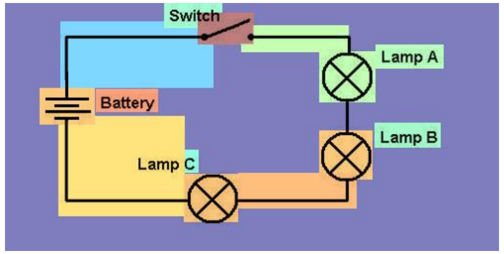


Which lamps would turn on if switch is connected?

a) b
b) a
c) a, b, c
d) c

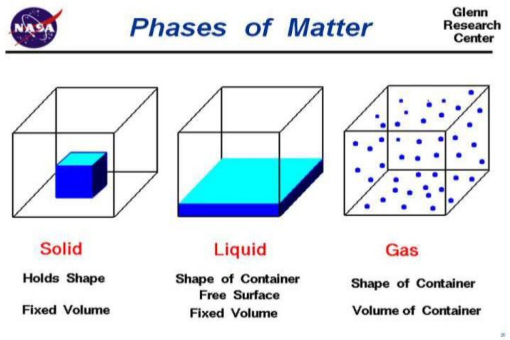


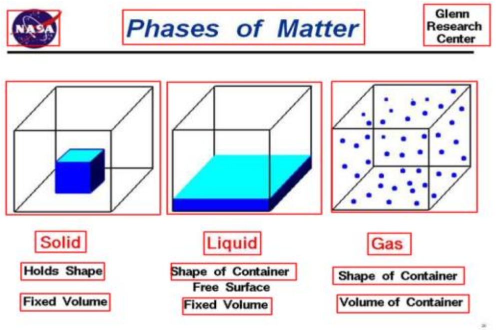


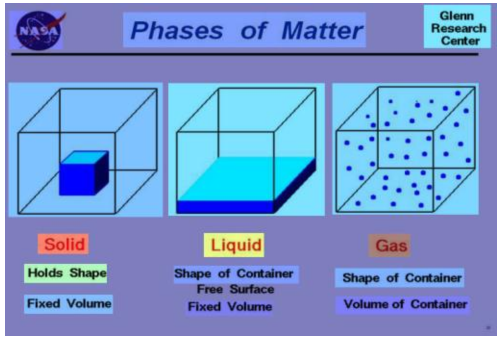


In which state does the substance hold shape?

a) solid
b) liquid
c) gas
d) none

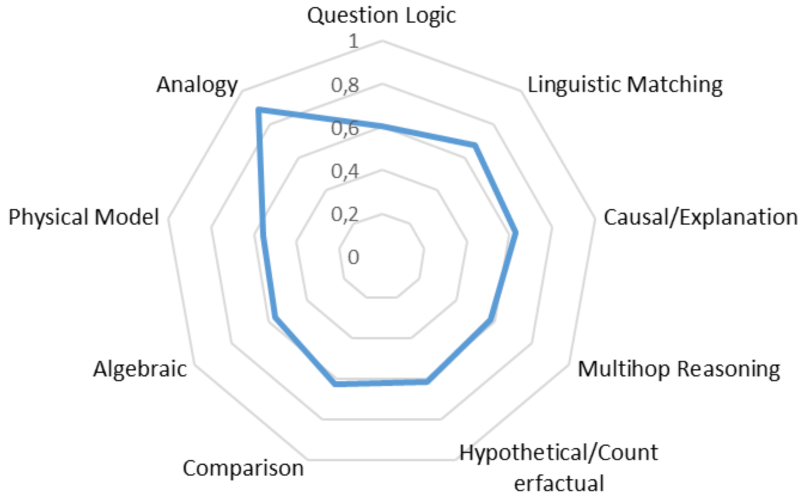




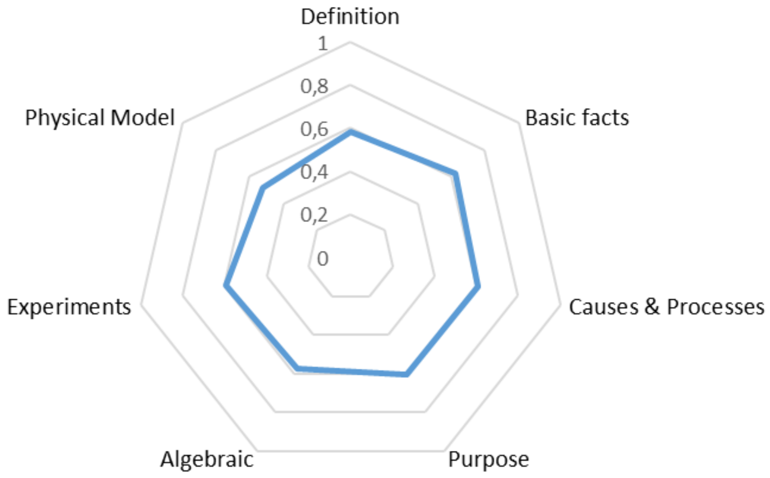


What do we know about our transformer-based models?

- **We executed ISAAQ on a sample of 203 text MC questions from ARC-Challenge** annotated (*) against 7 knowledge and 9 reasoning types
- Results in-line with the overall iSAAQ accuracy on ARC-Challenge (60.34%)
- Notice the **spike in analogical reasoning** (90% accuracy), a key reasoning type for textbook question answering
- Consistent with recent findings (**) on the **reasoning ability** of transformer language models
- **Probing: Are models actually learning knowledge and reasoning skills when trained on benchmark tasks?** (***)



Reasoning types



Knowledge types

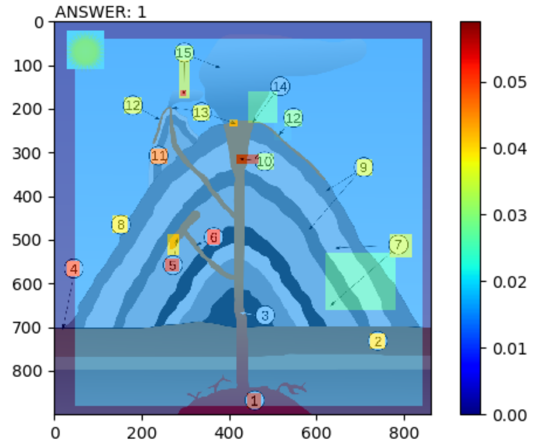
(*) M Boratko et al. 2018. *A systematic classification of knowledge, reasoning, and context within the ARC dataset. Workshop on Machine Reading for Question Answering*, pp. 60–70, Melbourne, Australia. Association for Computational Linguistics.

(**) P Clark, O Tafjord, and Kyle Richardson. 2020. *Transformers as soft reasoners over language.* ArXiv, abs/2002.05867.

(***) K Richardson, A Sabharwal, A. 2019. What Does My QA Model Know? Devising Controlled Probes Using Expert Knowledge. Transactions of the Association for Computational Linguistics, 8, 572-588.

What do we know about our (cross-modal) transformer-based models?

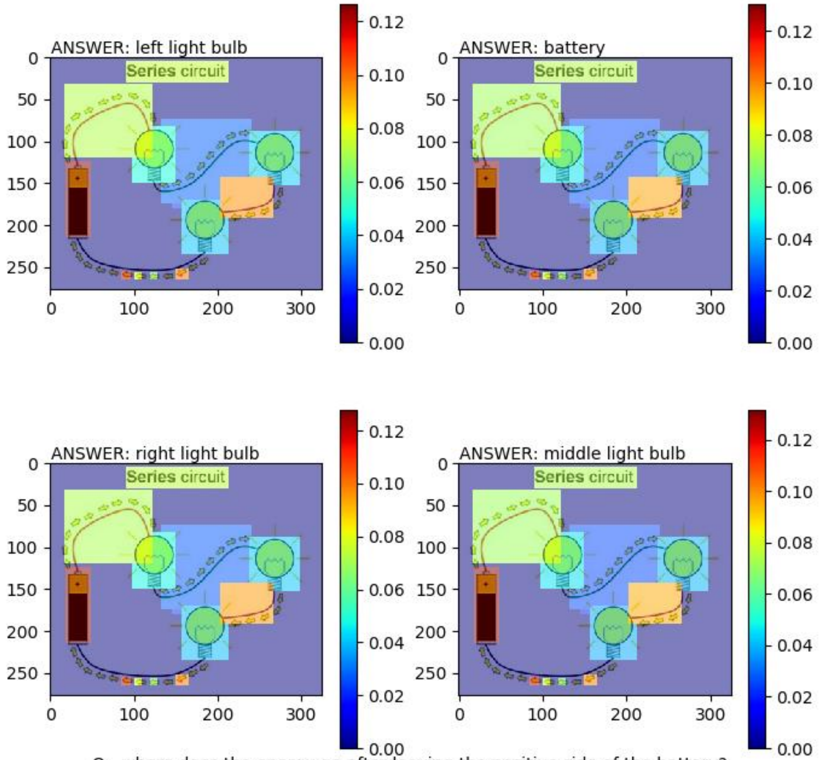
The intensity of visual attention is generally low…



Q: on the diagram, which number is pointing to the magma chamber? 1
CORRECT ANSWER:
CONFIDENCE:0.56827074

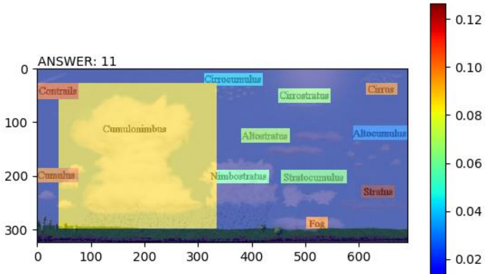
…even when confidence on the predicted answer is high

The intensity of the visual attention across the candidate answers tends to be very similar



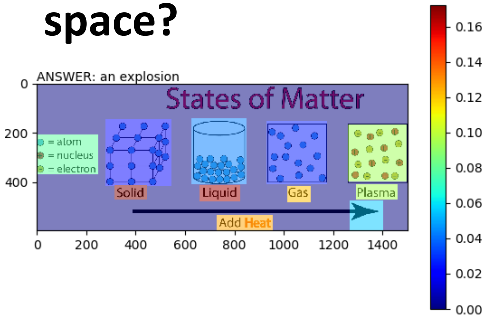
Q: where does the energy go after leaving the positive side of the battery?
CORRECT ANSWER: left light bulb
CONFIDENCE:0.31029627

Can ISAAQ count?



Q: how many types of condensation can occur?
CORRECT ANSWER: 12
CONFIDENCE:0.34962934

…and reason with space?



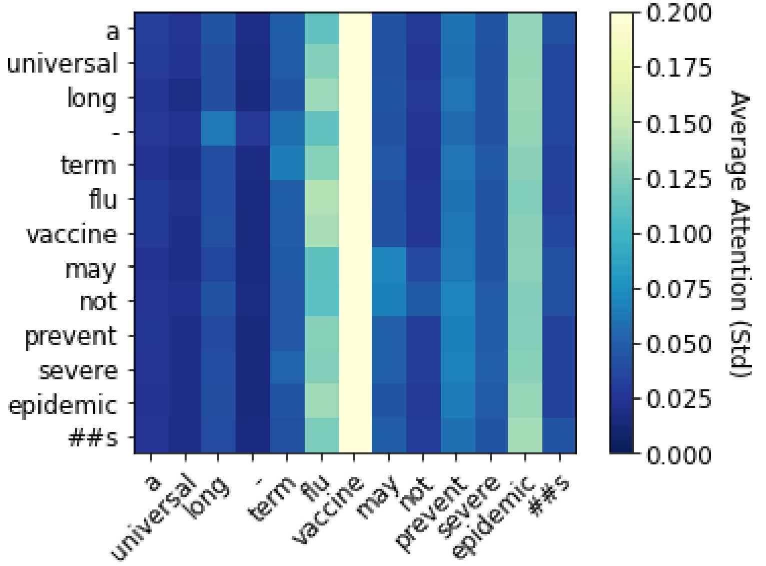
Q: if you cool gas, what do you get?
CORRECT ANSWER: liquid
CONFIDENCE:0.6637937

Self-attention as feature selection

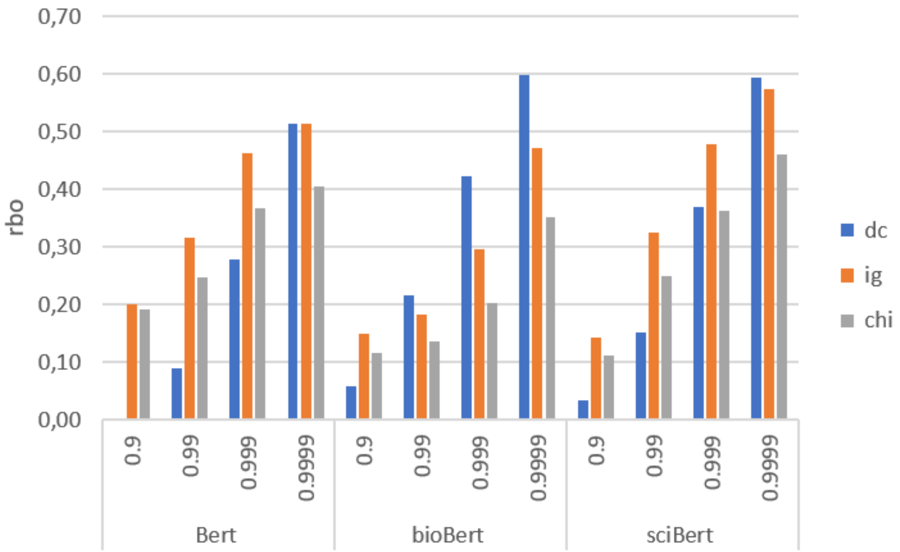
SN SciGraph

ConceptNet
An open, multilingual knowledge graph

Vardavas et al. 2010. ***A universal long-term flu vaccine may not prevent severe epidemics.*** BMC research notes, 3, 92.



BERT average weights in the self-attention heads of the last hidden state, <u>fine-tuned</u> on a classification task over SciGraph.



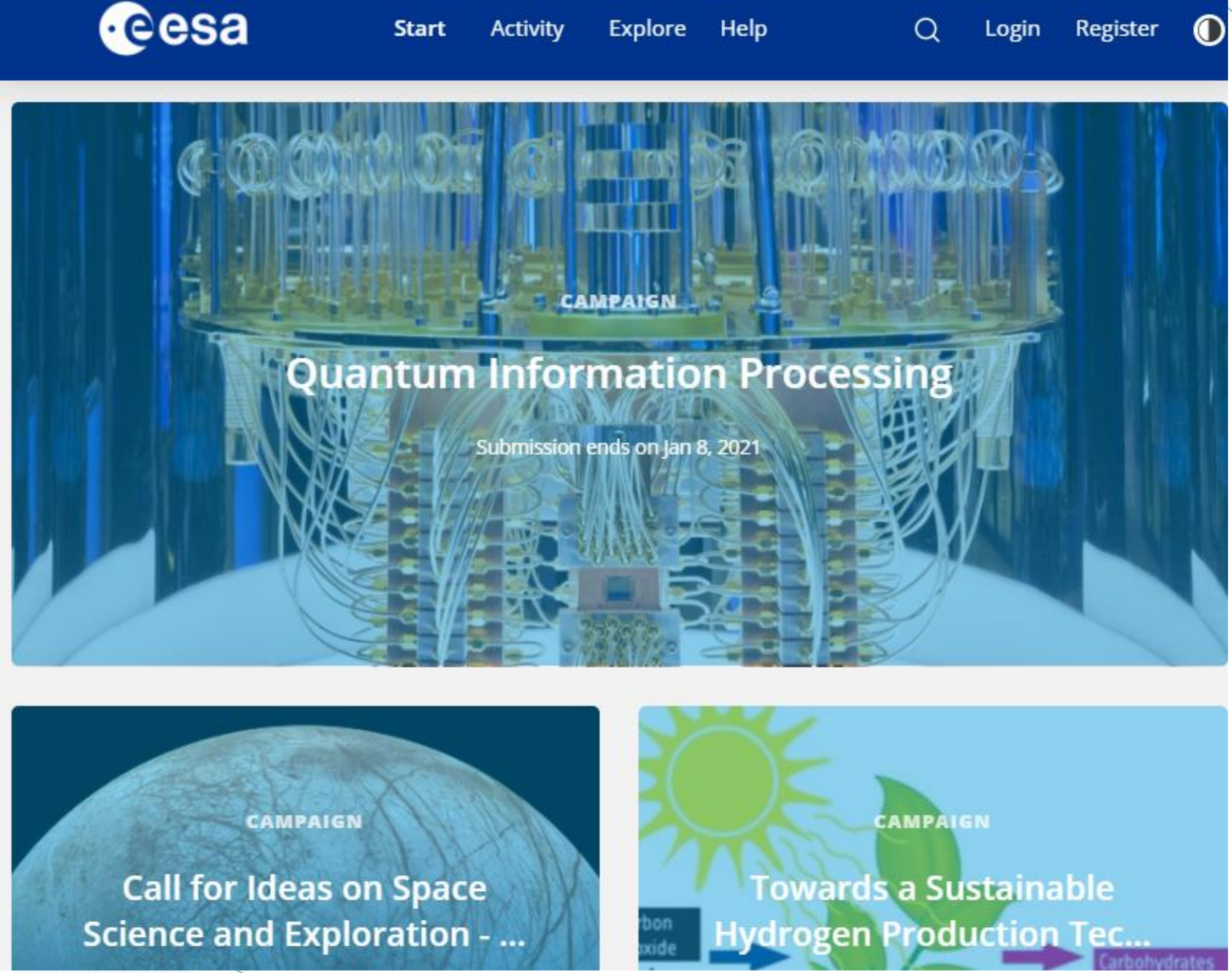
Rank biased overlap between most attended words and feature selection algorithms

- Key terms (*flu, vaccine, epidemics*) for the relevant class (*Medical and Health Sciences*) are attended more intensely
- Most attended words cover only 16% of the vocabulary
- Compared against words selected through conventional feature selection methods: chi-square (chi), information gain (ig), document frequency(df), and categorical proportional diference (pd)
- Relevant overlap, especially with ig and dc
- Mapped against ConceptNet (HasContext), showed more domain relevance than conventional methods

A Garcia, JM Gomez-Perez. 2021. *Classifying Scientific Publications with BERT - Is Self-Attention a Feature Selection Method?* European Conference on Information Retrieval (ECIR 2021), to appear.

What else?

- **Question generation**
  - to evaluate scientific knowledge comprehension
- **Hypothesis generation**
  - to propose new experimental work
- **Novelty evaluation**
  - to estimate the potential impact of new work
  - to verify its coherence with the SotA



The Open Space Innovation Platform

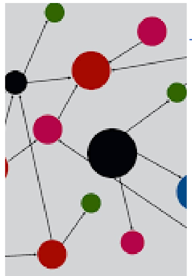
- **Campaigns:** temporary strategic calls
- **Channels:** permanent calls
- **Ideas**

Semantic similarity as novelty indicator

EXISTING SIMILAR IDEAS

EXISTING SIMILAR PROJECTS

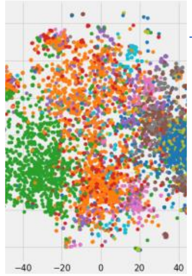
Building the model



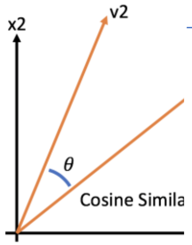
Extract domain-specific terminology and extend the KG

Automatically spot problems and proposed solutions in idea descriptions

Extract explicit semantic metadata



Learn latent representations with transformer language models



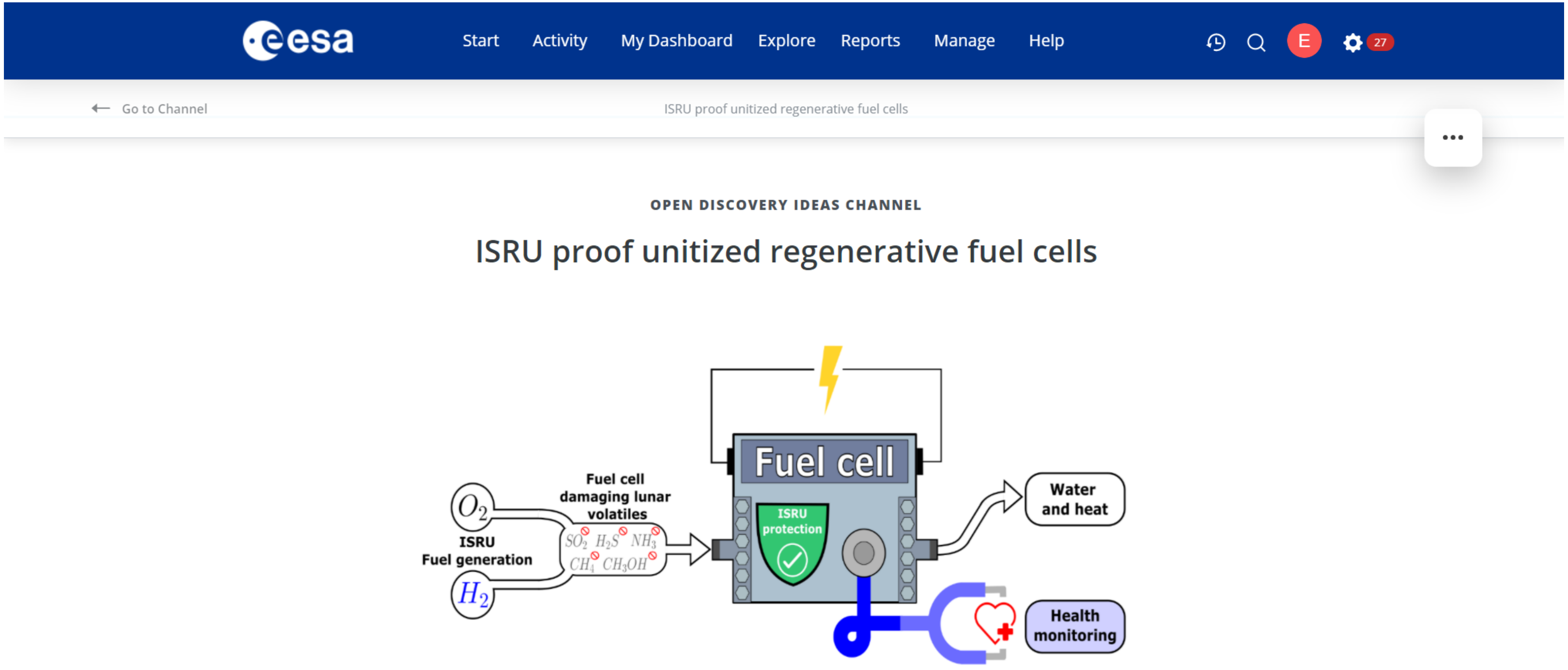
Model semantic similarity and novelty evaluation based on joint idea representations

EXISTING SIMILAR IDEAS

EXISTING SIMILAR PROJECTS

Evaluating ideas

Not only evaluate novelty but also explain why



Scientific claim analysis

Scientific claim

Space debris pose a huge risk for space exploration

Scientific literature

Claim verification

Satellites have to move out of the way of all this incoming space junk to make sure they don't get hit and potentially damaged or destroyed

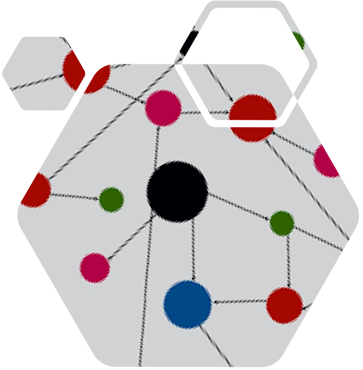
SUPPORTS

Collisions are rare: the last satellite to collide and be destroyed by space junk was in 2009.

REFUTES

Linking

The OSIP KG



Claims
Publications
Data
Authors

Interesting related work by: D Wadden et al. 2020. *Fact or Fiction: Verifying Scientific Claims.* 2020 Conference on Empirical Methods in Natural Language Processing (EMNLP). Association for Computational Linguistics

Wrapping up

Understanding

AI that assists scientists

Elements of language, logic, and cognition

Findable Accessible Interoperable Reusable

doi DC[1] Data Citation Principles SN SciGraph Semantic Scholar researchobject.org

Learning to read

AI that understands science

Machine reading comprehension, novelty evaluation, claim verification

ISAAQ

ARISTO

esa

Threshold understanding
(ability to develop domain expertise)

Reading to learn

AI that does scientific work

Generate scientific questions and hypotheses, design experiments, evaluate grants…

Collaboration, problem solving and thought partnership
(unlocking our creative potential)

Will AI Write the Scientific Papers of the Future?

Time

expert.ai

EUROPEAN LANGUAGE GRID

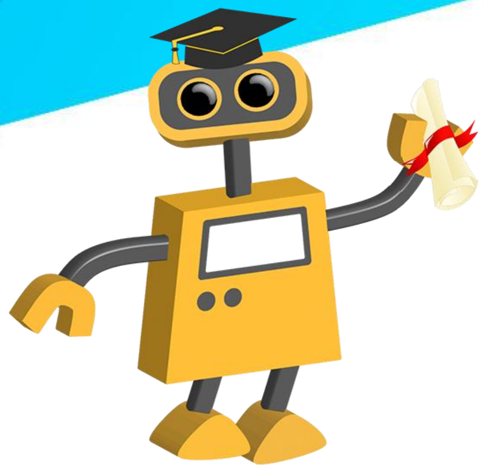
Reliance

Thank You

Linkedin.com/company/expert-ai

Twitter.com/expertdotai

jmgomez@expert.ai



Scientific publications today

Several initiatives to gain better visibility, reuse capabilities and to foster experimental **reproducibility and data/software accessibility**

- **Digital Object Identifiers**
- **Data Citation principles**
- **Software Citation principles**
- **Fair Principles**
- **Research Objects**
- **Academic Search Engines**
- **Scientific Knowledge Graphs**