I Trust AI

Muhammad Abdul-Mageed
Canada Research Chair
in Natural Language Processing and Machine Learning
Co-Director, I(nterPARES) Trust AI
The University of British Columbia

ITrustAI

In this Part...

1. Central Question
2. Deep Learning
3. NLP
4. Computer Vision
5. Speech
6. Educational Platform

Our Central Objective

- Develop **AI systems** for **records and archives *competently* and *efficiently* while maintaining the *nature* and *trustworthiness* of the records**

Tasks

- Break down archival functions into **AI tasks**
- Good tasks are ones clearly defined
- Proceed from simple to complex

Methods & Tools

- Develop **AI methods** to tackle tasks
- Focus: SOTA, explainable, interpretable

Data

- Identify, acquire, and develop data
- **Several questions**, e.g., ethics, privacy, ownership, respect, community norms

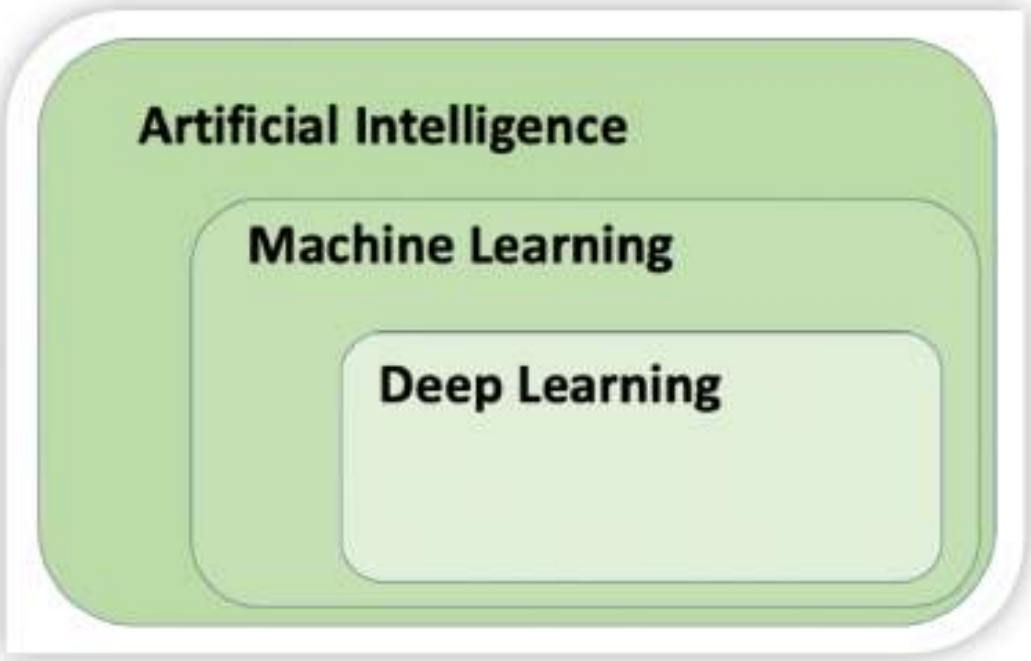
Evaluation

- What are **measures of success?**

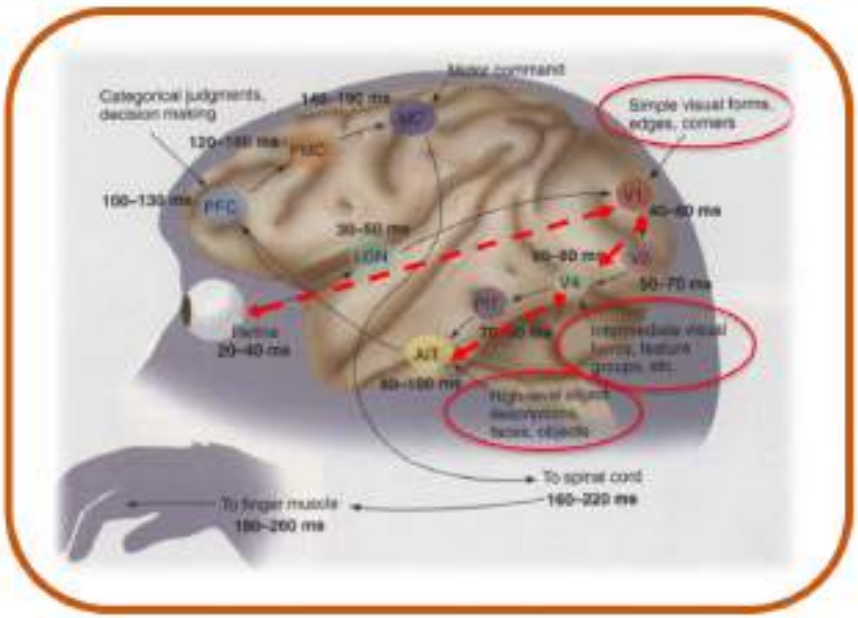
Human Resources

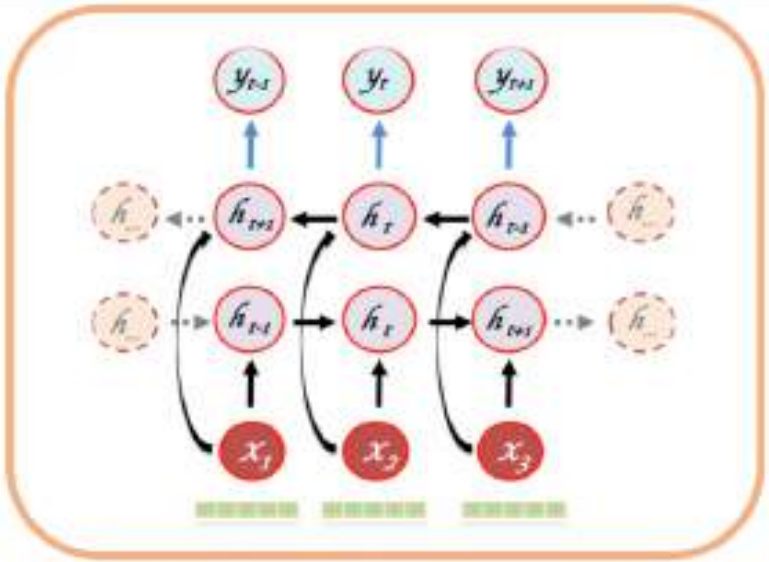
- A major training component
- Talks, tools, workshops, courses, tutorials, etc.

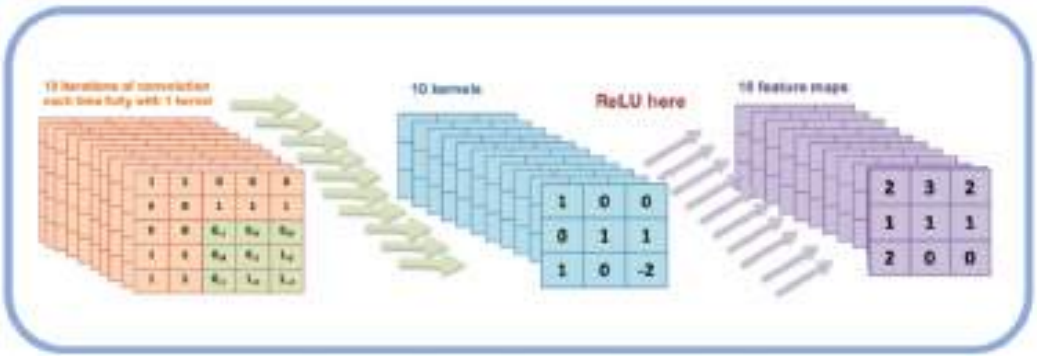
Artificial Intelligence

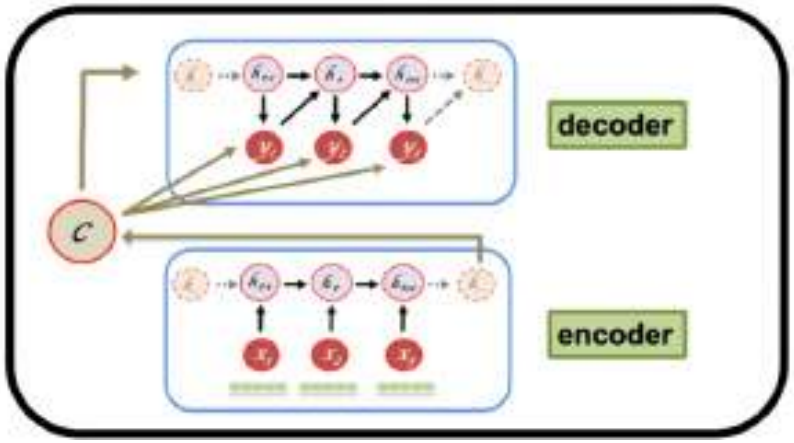


Deep Learning









DL for **Scientific Discovery**

Biology

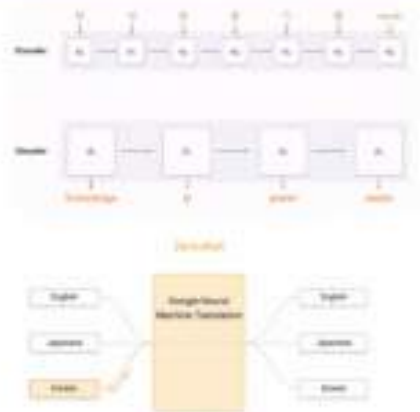
Astronomy

Agriculture

Chemistry

Image, Speech, and Language Processing

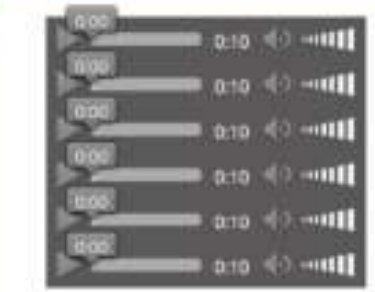
Machine Translation

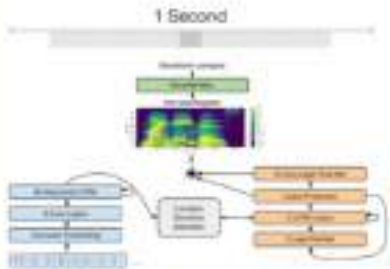


https://research.googleblog.com/

https://translator.microsoft.com/help/articles/neural/

Speech





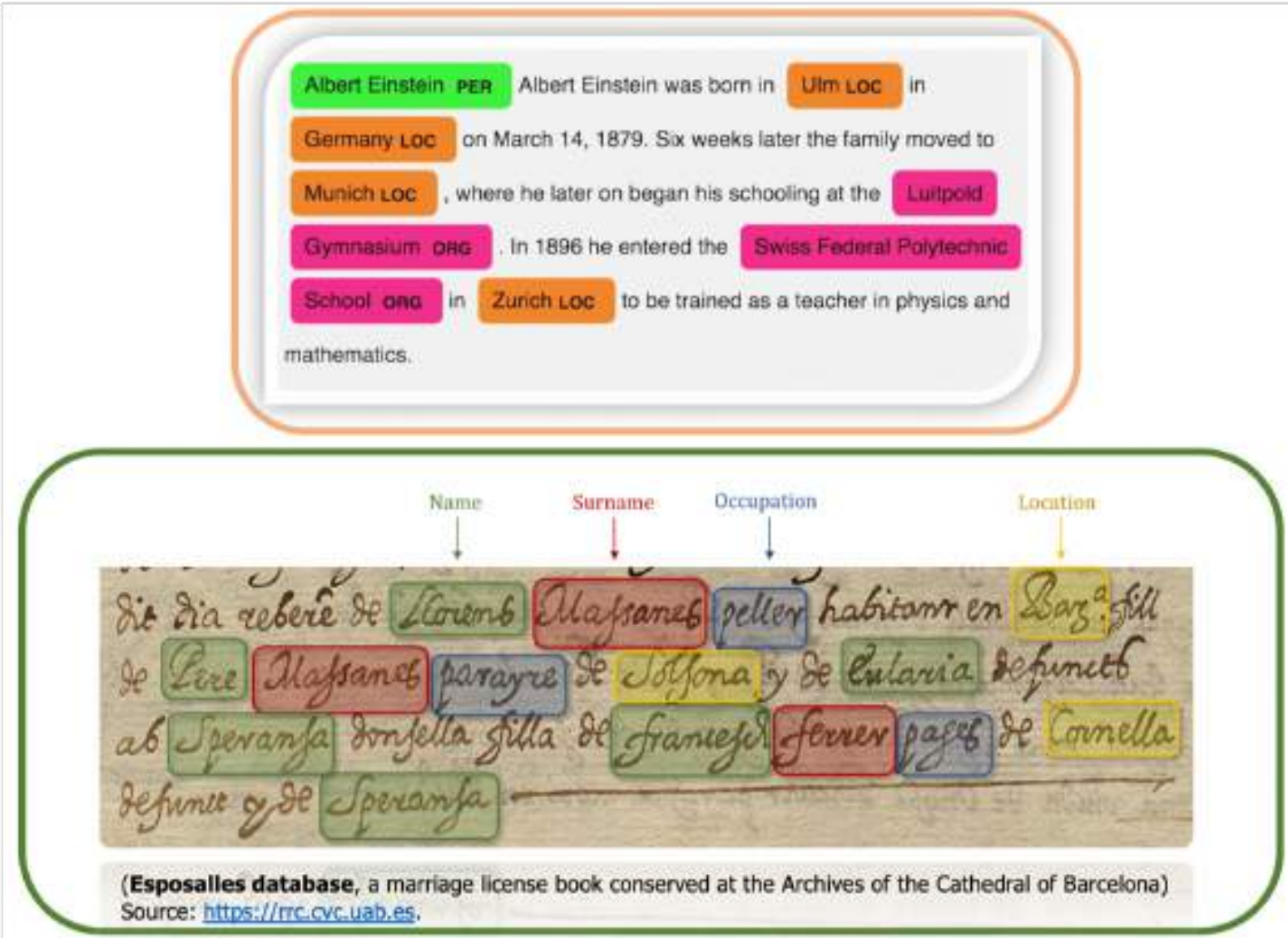
https://research.googleblog.com/

Image Captioning

[Microsoft Research: He et al., 2015]

https://research.googleblog.com/

Named Entity Recognition (NER)



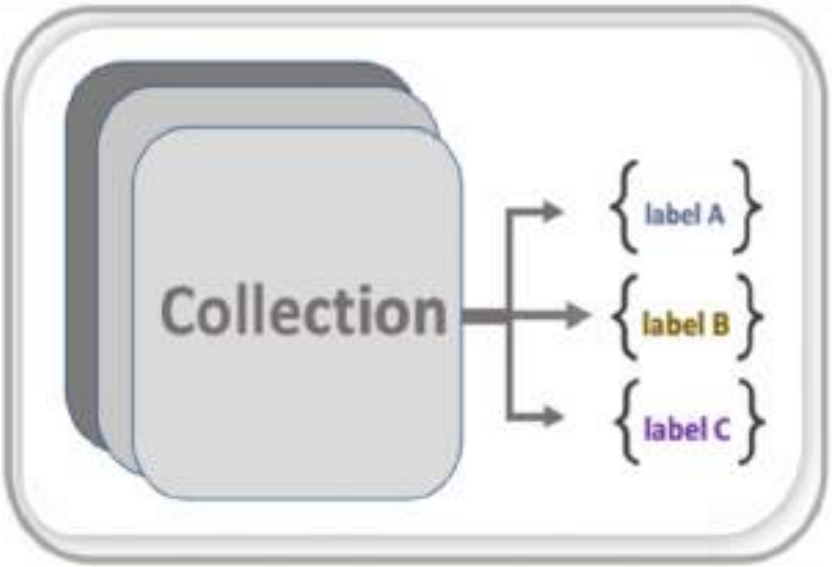
Topic Modeling

Topic modeling provides methods for automatically organizing, understanding, searching, and summarizing large electronic archives.

1. Discover the hidden themes that pervade the collection.
2. Annotate the documents according to those themes.
3. Use annotations to organize, summarize, and search the texts.

(Slide credit: David Blei)

Text Classification



(1) Just got chased through my house with a bowl of tuna fish. 😩 ing. **[Disgust]**

(2) I love waiting 2 hours to see 2 min. Of a loved family members part in a dance show 😒 #sarcasm **[Sarcastic]**

(3) USER Awww 😔 😔 CUPCAKES SUCK IT UP. SHE LOST 😡 😡 GET OVER IT 😡 😡 **[Offensive]**

Descriptions of Visual Archives

(Google image search)

Machine Translation

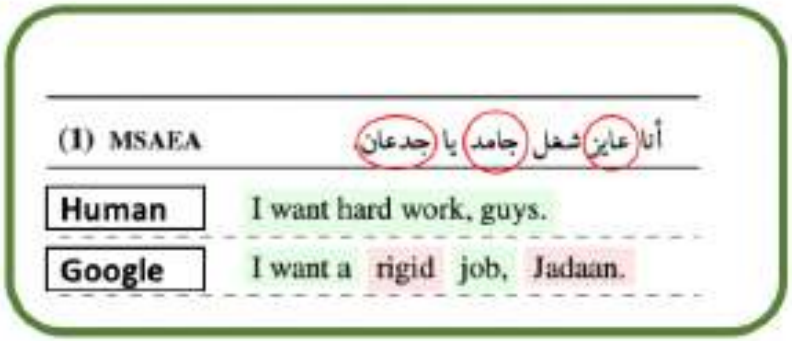
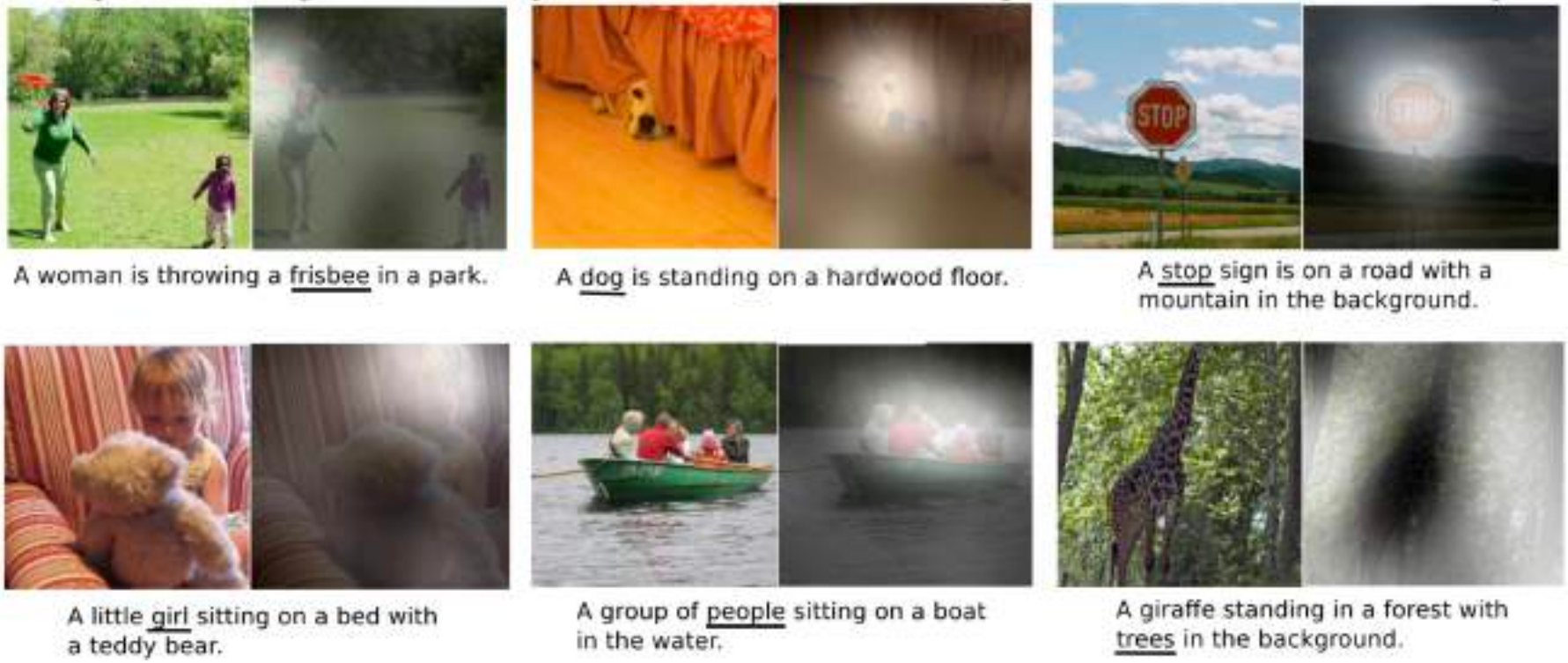


Image Captioning



(Xu et al., 2016)

Museum Image Captioning

Captions: (1) Lion-Shaped Furniture Leg. (2) In both Egypt and Nubia the lion was associated with the sun god and symbolized royalty.

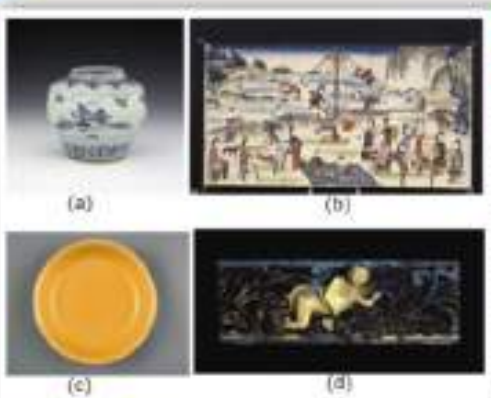


Figure 3: Four ancient Chinese artwork images.

Images	(a)	(b)	(c)	(d)
Artwork type	Jar	Print	Dish	Wall-tile
Model NIC	Porcelain bowl with underglaze blue decoration	Woodcut	There is a man in underglaze blue on the base	There is an inscription on the base
Model SA	Porcelain jar with underglaze blue decoration	Woodcut	Porcelain dish with rounded sides	Made of blue glazed porcelain
Model LSTM-A_5-MC	This ovoid jar has a short neck with a thickened rim and a recessed base	Woodcut	Porcelain dish with rounded sides	Made of blue glazed porcelain
Model LSTM-MC-$OUT_{dynamic}$	Porcelain jar with underglaze blue decoration	Ink and colours on paper	Porcelain dish with rounded sides	Earthenware wall tile with moulded decoration and green glaze

Figure 4: Generated captions for the images in Figure 4 by model NIC, SA, LSTM-A_5-MC and LSTM-MC-$OUT_{dynamic}$.

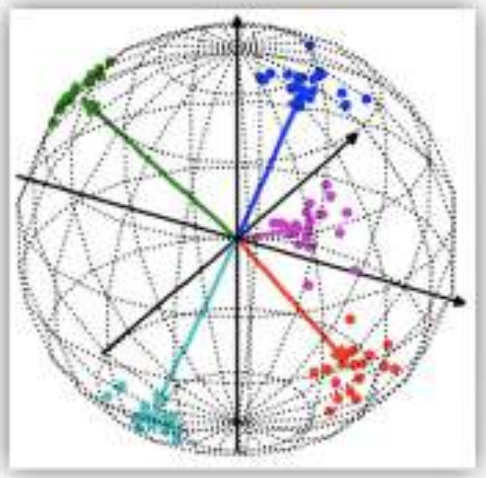
(a)	Porcelain jar with underglaze blue decoration. This jar has a narrow raised neck, rounded body and tapering foot. On the unglazed base the spiral marks resulting from turning the jar on the potter's wheel are clearly visible. On the outside an old man is shown in a landscape on one side and another, walking with a servant carrying a 'qin' wrapped in a textile cloth, on the other side, they are separated by plants and rocks. Around the neck is a double ring with banana leaves forming a collar; stylized lappets, like vertical long and short stripes, surround the foot. The cobalt blue appears black where it has not been covered by the blue-green glaze. Originally the jar would have had a domed cover with lotus-bud finial
(b)	Woodcut. Recreation. Acrobats and jugglers and women playing musical instruments. Printed in ink, colour on paper
(c)	Porcelain dish has yellow enamel over thin plain terapathic glaze. There is an inscription on the base, which is glazed
(d)	Wall tile with relief moulded decoration and fahua'-palette glazes. This 'fahua'-palette tile is moulded and carved in relief with a celestial boy flying in a contorted pose through scrolling foliage. He is naked except for a torque necklace with three pierced beads. He holds the stem of lotus foliage in his right hand and raises the other hand. The tiny curl in the centre of his forehead indicates his youth. The turquoise glaze is fugitive

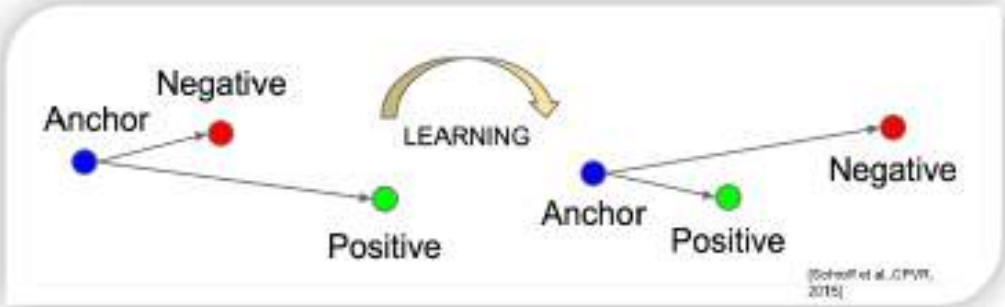
Figure 5: Ground-truth captions for the images in Figure 3.

(Sheng & Moens, 2019)

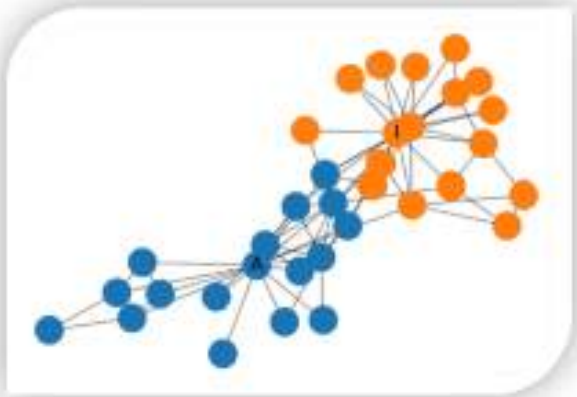
Active Research

Contrastive Learning





Graph Representation Learning



Educational Platform

(https://github.com/UBC-NLP/itrustai-tutorials)

AI and ML Tutorial Repository

This repository houses tutorials created for the iTrustAI SSHRC Partnership Grant. These tutorials will grow over time and will be used in hands-on workshops for training purposes.

Natural Language Processing

Natural Language Processing (NLP) is the field focusing on developing methods and tools for understanding and generating human language. As such NLP actually covers a wide range of technologies, with many different applications, that are often used in concert with one other. For instance, a type of text classification such as sentiment analysis might try to utilize part of speech (POS) tagging to disambiguiate word meaning through the POS tag, and therefor improve performance on the text classification task. In this section we will introduce code to train models for many of the core NLP tasks.

Part of Speech (POS) Tagging

	Category	Descriptions	Link
1	POS Tagging	POS with spaCy	notebook
2	POS Tagging	Train BiLSTM with PyTorch from Scratch	notebook
3	POS Tagging	Finetune with BERT from Scratch	notebook

Find Us

www.interparestrustai.org

@itrustai

www.facebook.com/interparestrust