

Linked Open Data and Knowledge Graph

Application Topic: Never Ending Language
Learning (NELL)

Agenda

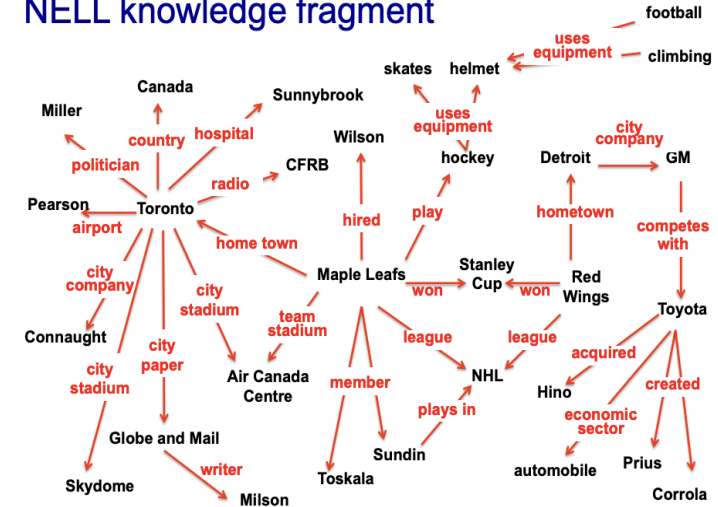
1. Introduction
2. Motivation
3. Tasks of Learner
4. Candidate facts
5. Beliefs
6. NELL architecture
7. NELL learning overtime
8. Typical ML Vs NELL
9. Applications
10. Challenges and limitations
11. Conclusion



Introduction to NELL

1. NELL stands for Never Ending Language Learning.
2. It is an AI system developed in 2010 by a research team at Carnegie Mellon University.
3. It is a system that continuously learns from Web.
4. It extracts, Categorize and refines knowledge much like a human does.

NELL knowledge fragment



https://www.cs.cmu.edu/~tom/pubs/NELL_aaai15.pdf



Why was NELL created?

The goal is to build an AI that can learn autonomously, evolving its understanding over time without human intervention.

- To mimic human like learning.
- To develop a machine that continuously read, learn and improve.
- To handle vast unstructured information available online.
- Domain Adaptability
- Iterative Improvement
- Create a reliable ever-growing knowledge base.



<https://bernardmarr.com/the-evolution-of-ai-transforming-the-world-one-algorithm-at-a-time/>



What Never Ending Language Learner does?

It is an intelligent computer agent that runs 24 hours per day, 7 days per week, forever, performing the following two tasks

Reading Task

Extract information from web text to further populate a growing knowledge base of structured facts and knowledge.

Learning/Refining Task

Learn to read better each day than the day before, NELL reviews its existing knowledge to validate or correct previously learned facts and categories.

What are Candidate facts?

- Facts are pieces of information extracted by NELL from web data.
- These are Candidate Facts, which means that they have been identified as potentially correct but are not yet confirmed.

EXAMPLE: “Paris is the capital of France.”

Candidate Facts: Paris -> CapitalOf -> France

Paris -> City

France -> Country

At this stage NELL is not entirely confident about the correctness and stores these as Candidate Facts in the Knowledge Base.

What are Beliefs?


- Beliefs are facts that NELL has validated and considers trustworthy.
- These are the confirmed pieces of knowledge stored in NELL's permanent knowledge base.

EXAMPLE: "Paris is the capital of France."

Candidate Facts: Paris -> CapitalOf -> France

Paris -> City

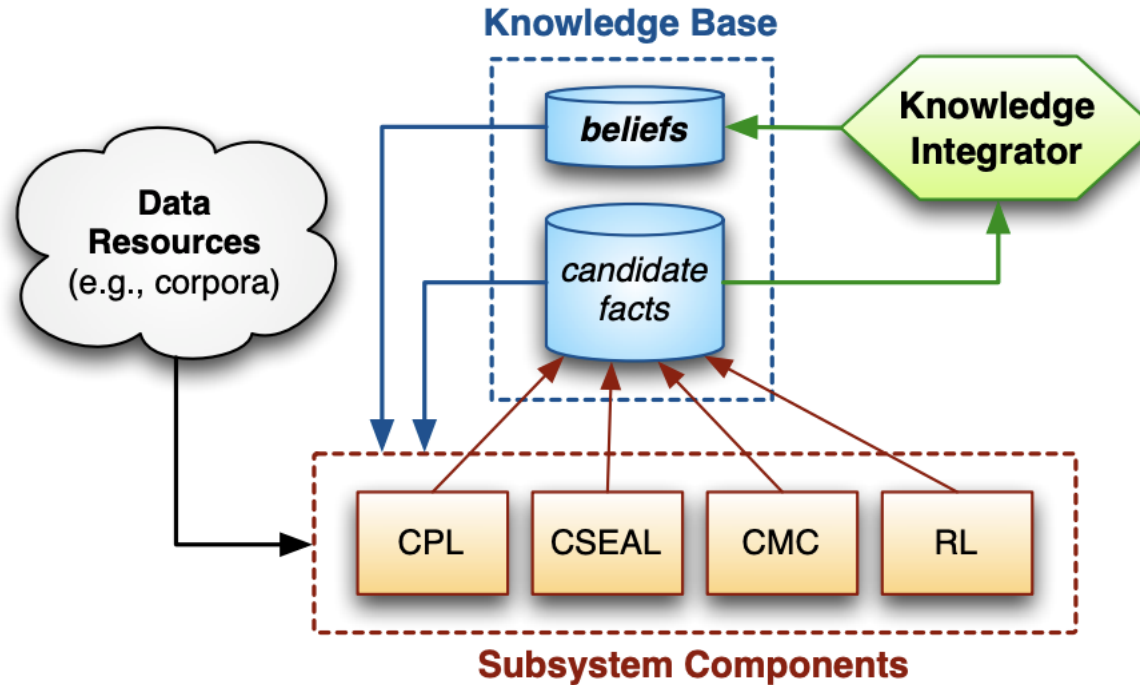
France -> Country

- 
1. Cross-referencing sources
 2. Internal consistency
 3. Confidence levels

After NELL encounters multiple sources supporting the candidate facts listed above, it elevates these to a belief in its knowledge base.



Architecture of NELL





Architecture of NELL

1. **CPL (Category pattern Learner):** Learn patterns to identify categories. E.g: X is a city
2. **CSEAL (Coupled Pattern Learner):** Focuses on extracting relations between entities. E.g: X is located in Y.
3. **CMC (Coupled Morphological Classifier):** Analyzes words or terms to classify them based on morphology. E.g: Apple is a fruit vs Apple is a company.
4. **RL (Relation Learner):** Learns and Identifies relationships between entities. E.g: CapitalOf, part of.

These systems produces candidate facts from the input data.



What NELL has learned over time?

- So far, NELL has accumulated over 50 million candidate beliefs by reading the web.
- NELL has high confidence in 2,810,379 of these beliefs.
- NELL can be tracked online at <http://rtw.ml.cmu.edu>
- URL: <http://rtw.ml.cmu.edu/rtw/kbbrowser/>

NELL Knowledge Base Browser
CMU Read the Web Project

log in | preferences | help/instructions | feedback

categories relations

everypromotedthing
location
building
airport
bridge
hotel
placeofworship
retailstore
museum
monument
restaurant
stadiumeventvenue
shoppingmall
skyscraper
hospital
trainstation
geopoliticallocation
county
continent
stateorprovince
country

al_shifa_hospital (hospital)
literal strings: [Al-Shifa Hospital](#), [Al-Shifa hospital](#), [Al Shifa Hospital](#)

Help NELL Learn!

NELL wants to know if this belief is correct.
If it is or ever was, click thumbs-up. Otherwise, click thumbs-down.

• [al_shifa_hospital](#) is a [hospital](#) 👍👎

categories

• [hospital](#) (100.0%)

- CPL @1095 (98.2%) on 17-jan-2018 ["Medical sources in _" "intensive care unit at _" "ICU" at _ "care unit at _" "trauma center at _" "Medical sources at _" "medical source at _" "medical sources at _"] using al_shifa_hospital
- CMC @1103 (100.0%) on 08-mar-2018 [PREFIX=hos 5.16625 PREFIX=hosp 3.57541 PREFIX=hospi 3.57319 SUFFIX=pital 3.40002 SUFFIX=ital 3.18492 PREFIX=ho 3.13044 LASTSUFFIX=al 3.04441 CHARS -0.25793 WORDSHAPE=Aaaaa -0.68963 WORDS -3.55178] using al_shifa_hospital



Typical ML vs NELL

Features	Typical Machine Learning	NELL
Learning Style	Static Learning	Continuous Learning
Adaptability	Limited, cannot handle new data unless retrained	High, processes new data daily
Human Intervention	Significant manual intervention needed	Facts extraction, validation and correction is automated
Error correction	Errors requires retraining and debugging by humans	Errors are self corrected via feedback loops and revalidation
Knowledge Base	Doesn't maintain a persistent knowledge base	Maintain a structured knowledge base (Beliefs and Candidate Facts)



Applications of NELL

1. Knowledge Base Construction
2. NLP Enhancement
3. Semantic Search & Q&A



Challenges and Limitations

1. Knowledge Accuracy
2. Language Ambiguity
3. Lack of Reasoning
4. Noisy/Biased Data
5. Limited Supervision
6. Domain Limitations
7. Implicit Knowledge
8. Error Propagation



Conclusion

- NELL aims to demonstrate that an AI can learn autonomously, adapt to new information, and continuously improve without needing human supervision.
- It's an innovative step toward creating intelligent systems capable of understanding and interacting with the complexity of the real world.
- In essence, NELL is like a "curious child" that never stops reading and learning, with the capacity to understand the web's vast amount of information in a structured and organized way.
- Its learning spans everything from simple facts to nuanced, multi-meaning concepts, making it a powerful example of autonomous learning in AI.



References

1. https://en.wikipedia.org/wiki/Never-Ending_Language_Learning
2. https://www.cs.cmu.edu/~tom/pubs/NELL_aaai15.pdf
3. <http://rtw.ml.cmu.edu/rtw/kbbrowser/>
4. <http://rtw.ml.cmu.edu/papers/carlson-aaai10.pdf>
5. <https://techcrunch.com/2010/10/09/nell-computer-language-carnegie-tctv/>
6. https://videolectures.net/videos/akbcwekex2012_mitchell_language_learning
7. <https://www.datacamp.com/blog/what-is-continuous-learning>