TrueWeb
A Proposal for Scalable Semantically-Guided Data Management and Truth Finding in Heterogeneous Web Sources

Amgad Madkour (Purdue)
Walid G. Aref (Purdue)
Sunil Prabhakar (Purdue)
Mohamed Ali (U Washington, Tacoma)
Siarhei Bykau (Bloomberg)
Motivation

• We envision a **responsible web environment**, where a user should be able to find out whether any *sentence* on the web is *true* or *false*

• The user should be able to **track the provenance** of any sentence or paragraph on the web

• **Compose factual knowledge** from web resources about any subject of interest and assign some *belief factor* for each fact

**Figure**: Guidelines to spot fake news provided by the International Federation of Library Associations and Institutions
Problem Definition

- Users *manipulate* (read, share and comment on) posts on social media affecting other users' mindsets, attitudes and responses

- Extracting *credible* pieces of *information* out of a *mixture of news* coming from sources with variable degrees of *trustworthiness*

- The *credibility* of various posts on *social media* is hard to evaluate
TrueWeb
Overview

• Create a **semantically-guided system** for knowledge graphs

• Create reliable **truth finding techniques**

• Our use cases are mostly drawn from **validating** sentences

• This can still be **extended** for other cases such as prediction of crimes, climate change etc.
• **Social media** provides a vast amount of information which contains some important facts or observations.

• **News agencies** aim at monitoring the mainstream social media and extracts valuable messages, posts or tweets which can be used as a source for its news articles.

• The main challenge is to **distinguish** real facts and gossips or intentionally false evidences (e.g. vandalism).

• This requires **data processing** at a large scale of millions of social media messages, hundreds of thousands of news articles and billions of web pages.
Some news agencies may ask their readers to report news and then uses those reports to provide up-to-date coverage of events and accidents.

The main problem is that reports may contain inaccurate or even false information.

Example: A user reports a traffic accident at Grant St.
- Verify if there is an accident at all or the user reports some false information.
- Assess the quality of his/her report accuracy, e.g., whether there is an accident at Grant St. or maybe it is at Salisbury St.
We propose to develop and prototype for a semantically guided system for the management of knowledge graphs.

We adopt the notion of Knowledge cubes (KC) for the prototype [Madkour 2013].

Each KC is responsible for a certain semantic topic, e.g., sports, US presidents, or certain geographical regions.
Knowledge Cube (KC)

Description

- Data extracted from the Internet will be directed towards the relevant KCs for further investigation and scrutiny.
- A KC is an unsupervised and adaptive database instance of knowledge.
- A KC is capable of storing, analyzing, and searching linked-data components in the form of RDF triplets.
Knowledge Cubes (KC)
Architecture (1/2)

• The catalog maintains all the information related to the data sources it fetches.

• The information extraction component employs text analysis techniques in order to extract and learn from structured and unstructured sources.

• The search and query component provides a rich set of constructs that semantically parses queries.
Knowledge Cubes (KC)
Architecture (2/2)

- The **data store and indexing** component provides the scalable storage and indexing mechanisms.
- The **discovery of data sources** component identifies data sources relevant to the KC.
- The **data sources update and extension** component create a time-oriented snapshot of the current knowledge store data.
- The **semantic query processing** component infers if a certain statement is true or false.
Knowledge Cubes (KC)  
Research Challenges — KC Construction

- How to extract RDF data (i.e. subject-predicate-objects) from textual resources that will be used to construct the knowledge cubes.

- How to identify co-occurrence of entities within textual resources to exposes implicit relations.

- How to use the spatial dimension to answer non-spatial topics that in turn can help answer investigative queries that are not possible to answer otherwise.
Knowledge Cubes (KC)
Research Challenges — Semantic Query Processor

• The semantic query process will make heavy use of the credibility and trustworthiness modules

• **Query processing** will operate on the **KC attributes** such as its *topical, spatial, temporal, and contextual* aspects to **validate** a given statement under investigation or respond to a user's query

• How to decide on the **order of execution** with respect to the KC attributes
Truth Finding Techniques

- **Every single entity**, e.g., user, news reporter, and organization, is tagged with a dynamically changing **trustworthiness score**

- **Every post** is tagged with a dynamically changing **credibility score** to reflect how far this post is believed to be true

- TrueWeb **adjusts** scores based on **how entities respond to a post**

- These adjustments are a **continuous** process as the posts hop from one entity to another in the social media graph
Credibility and Trustworthiness
Architecture (1/2)

• Build and **continuously maintain credibility** and trustworthiness scores in a social network.

• The **connectivity and relationship analyzer** takes as input the social network graph and analyzes the connectivity and relationships among entities.

• The **data integration manager** considers all posts in the social network, correlates these posts together and decides the credibility score.
Credibility and Trustworthiness Architecture (2/2)

• The **data lineage and provenance monitor** tracks posts as they hop from one entity to another and how they are handled by entities.

• The **semantic and sentiment analyzer** helps decide on the reaction an entity showed in response to a post.

• The **reputation builder** elevates or de-elevates the trustworthiness score of entities across all dimensions.
Credibility and Trustworthiness
Research Challenges - Semantic Interpretation and Conflict Detection

• If some of the parts of the semantic RDF graph contains false information, then this affects the accuracy of its neighbors.

• How to represent possible worlds and find the most-likely state of conflicting data that maximizes the observations seen so far.

• How to identify the correlations among concept attributes and use these correlations detect semantics-based conflicts among the underlying data.
Credibility and Trustworthiness
Research Challenges — Detection of Source Independence and Conflict of Interest

• How to discover conflict of interest through the network of relations of an entity or through the content of the post

• How to detect conflict in the content using a user profile for each entity

• Detecting conflict through the network of relations requires monitoring and assessing the similarity in behavior among the entire network of an entity
Credibility and Trustworthiness
Research Challenges — Assessment of Proactiveness, Reactiveness, and False Proactiveness

This research task addresses the ability to classify users into three categories:

1. **Proactive Sources**: Represents the original sources of the post

2. **Reactive Sources**: Represents the entities that compile their posts from other sources

3. **False Proactive Sources**: Represents the entities that initiates a post that is based on posts from other sources without proper citation
Credibility and Trustworthiness
Research Challenges - Influence-based Entity Ranking

• How to find the correct items that the user can provide feedback on given a limited budget?

• How to solicit user feedback to improve the accuracy

• How to take advantage of voting relationships and dependencies among facts and sources
Credibility and Trustworthiness
Research Challenges - Semantic-based Analysis and Classification Techniques

- How to model a post (e.g., in RDF) based on it being a fact or an opinion?

- How to classify a user post as a past fact vs. future speculation?

- However, a degree of incorrectness in posts that refer to tentative events in the future may be acceptable without imposing a large penalty on the trustworthiness of the posting entity.
Conclusion

• We envision TrueWeb as an oracle for validating the truthfulness of sentences

• We plan to study predictive queries based on the structured knowledge available in TrueWeb

• We also plan to investigate extending the provenance of TrueWeb where we can indicate whether a sentence was true given a specific possible world scenario

• We also plan to investigate utilizing the current TrueWeb prototype in order to discover entities that can be masquerading as different individuals over the web
Questions ?