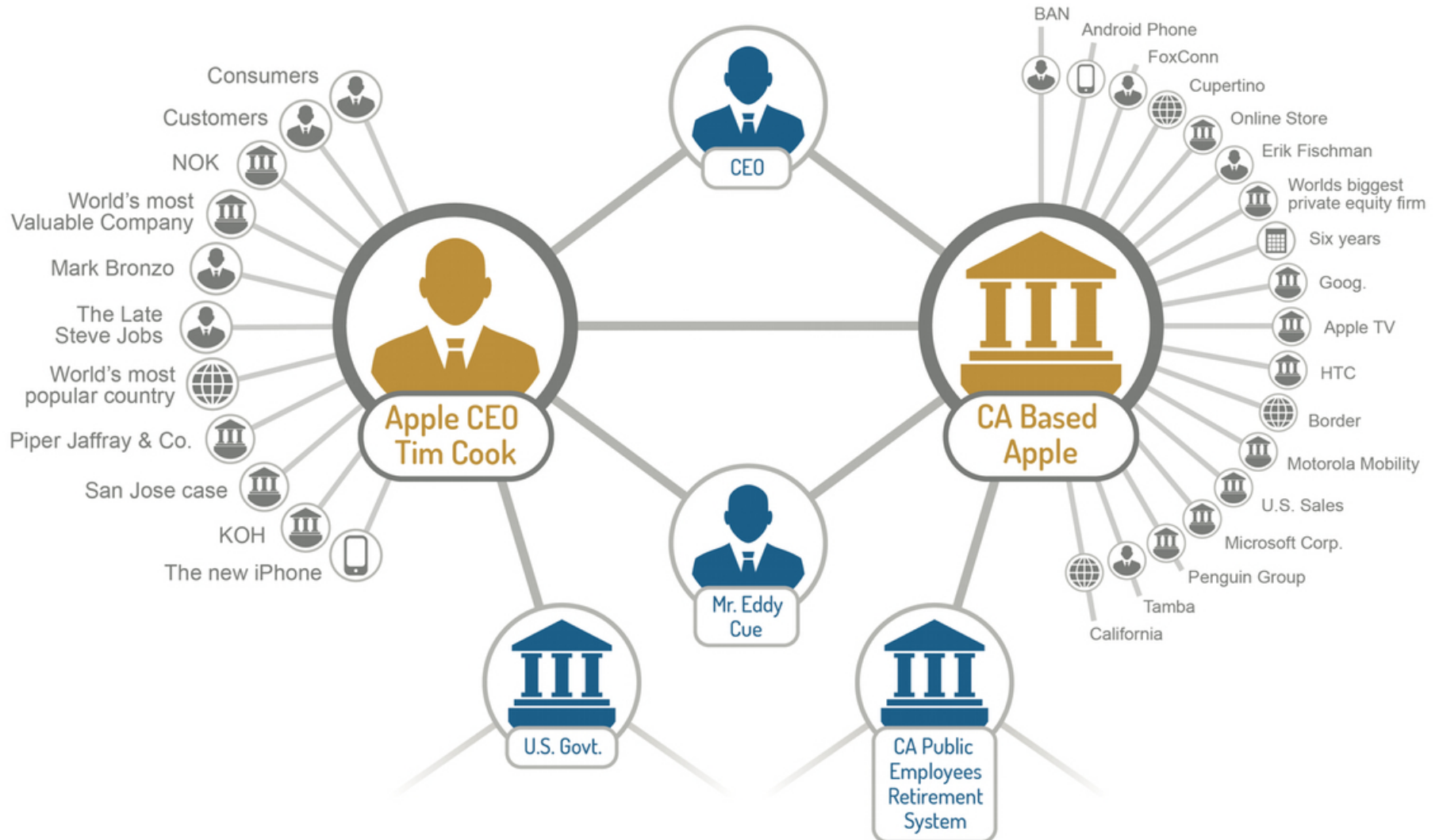# CULTURAL HERITAGE AND BIG DATA

By: Kun Hu, Marianne Marjanovic, Jianfeng Zhu

THE KNOWLEDGE GRAPH

LINKS TOGETHER BILLIONS OF ENTITIES, FACTS AND RELATIONSHIPS

Consumers

Customers

NOK

World's most Valuable Company

Mark Bronzo

The Late Steve Jobs

World's most popular country

Piper Jaffray & Co.

San Jose case

KOH

The new iPhone

CEO

Apple CEO Tim Cook

Mr. Eddy Cue

U.S. Govt.

CA Based Apple

BAN

Android Phone

FoxConn

Cupertino

Online Store

Erik Fischman

Worlds biggest private equity firm

Six years

Goog.

Apple TV

HTC

Border

Motorola Mobility

U.S. Sales

Microsoft Corp.

Penguin Group

Tamba

California

CA Public Employees Retirement System

Introduction

- ☐ Problem Statement
- ☐ Our Goal
- ☐ Our Process
- ☐ Future Work

Problem Statement

- Lack of organized digitized information linking knowledge of cultural heritage and relics.

- Today's implementation of digitization of cultural heritage and relics is inadequate for a friendly user experience.

Our Goal

- Create a Semantic Web Application.

- To show the ancient coins during the Roman Empire and provide additional information regarding to the coin and the ruling people of the time.

Visual Studio Code

Visualization

Request Setting, User Interface Design, Web Components

Server Service   **SuperAgent**

Data Query, User Request Transformations   *Fetch Tool*

Mongo DB

Selected Attributes, Data Structure

Data Source   **JSON-LD & Nomisma Ontology**

OCRE, MANTIS Data and DBpedia   *Schema*

Basic : JavaScript , Mongo DB , SuperAgent

Fetch Data

- Data Sources:
    - OCRE Data: http://numismatics.org/ocre/apis
    - MANTIS Data: http://numismatics.org/search/apis
    - Dbpedia: http://wiki.dbpedia.org/

Region

Data Access

Individual Records

Numishare supports delivery of individual records in a variety of models and serializations through both REST and content negotiation. Content negotiation (with the accept header) requests should be sent to the URI space http://numismatics.org/ocre/id/. Requesting an unsupported content type will result in an HTTP 406: Not Acceptable error.

Model	REST	Content Type
HTML	http://numismatics.org/ocre/id/{$id}	`text/html`
NUDS/XML	http://numismatics.org/ocre/id/{$id}.xml	`application/xml`
KML	http://numismatics.org/ocre/id/{$id}.kml	`application/vnd.google-earth.kml+xml`
RDF/XML	http://numismatics.org/ocre/id/{$id}.rdf	`application/rdf+xml`
Turtle	http://numismatics.org/ocre/id/{$id}.ttl	`text/turtle`
JSON-LD	http://numismatics.org/ocre/id/{$id}.jsonld	`application/ld+json`
geoJSON	http://numismatics.org/ocre/id/{$id}.geojson	`application/vnd.geo+json`

Search Results

Search results (the browse page) are returned in HTML5, but Numishare supports Atom and RSS via REST, as well as Atom and raw Solr XML via content negotiation of the browse page URL, http://numismatics.org/ocre/results. The REST-based Atom feed sorts by the Lucene syntax 'timestamp desc' by default, but the sort parameter may be provided manually to alter the default field and order.

Model	REST	Content Type
HTML	http://numismatics.org/ocre/results	`text/html`
Atom	http://numismatics.org/ocre/feed/	`application/atom+xml`

8

1 Querying DBpedia

The DBpedia data set enables quite astonishing query answering possibilities against Wikipedia data.

1.1 Public SPARQL Endpoint

There is a public SPARQL endpoint over the DBpedia data set at http://dbpedia.org/sparql. The endpoint is provided using OpenLink Virtuoso as both the back-end database engine and the HTTP/SPARQL server.

The public endpoint does NOT include all available DBpedia data sets. See the list of all DBpedia data sets that are currently loaded into the public SPARQL endpoint.

You can ask queries against DBpedia using:

- the OpenLink Interactive SPARQL Query Builder (iSPARQL) at http://dbpedia.org/isparql;
- the SNORQL query explorer at http://dbpedia.org/snorql (does not work with Internet Explorer); or
- any other SPARQL-aware client(s).

NOTE: Please read the documentation and usage notes about the public SPARQL endpoint carefully, before querying the service. In terms of a Fair Use Policy, restrictions and limitations may apply for complex queries.

1.2 Triple Pattern Fragments

JSON_LD

- Flexible and loose structure to organize data.

- Convenient and light weight to transfer online.

- Can be applied with schemas to describe content and restrain the form.

MongoDB

- ☐ Compatible with JSON format.

- ☐ Suitable for big data processing and dynamic data storage.

- ☐ Easy to learn and use.

Visualization

- 44,000+ records of coins used.
- Web component technique in the UI.

Future Work

- Extend our project into a Knowledge Graph.

- Perform in depth analysis of data.
  - Classify semantic terms.
  - Calculate document distance.

- Present more attributes into our project.

- Create more Visualization according to user interest and data characteristics.

Thank you!