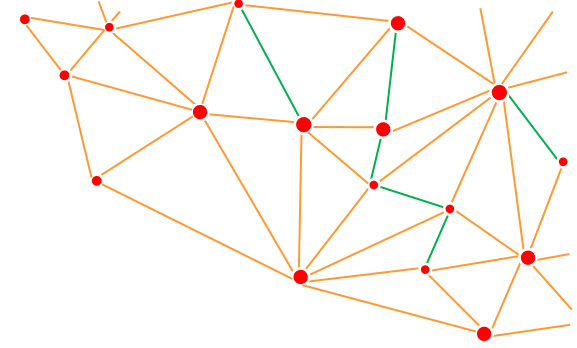# Design and Implementation of CLaDA-BG-Research System: Modeling Streets Through a Sofia Usecase
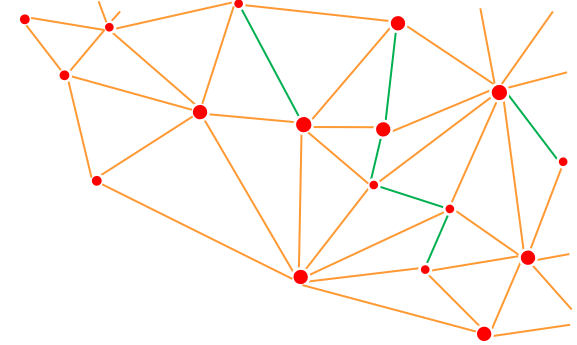
Zhivko Angelov*, Ivaylo Nachev^, Kiril Simov*

Institute of Information and Communication Technologies, BAS*
Institute of Balkan Studies and Centre of Thracology, BAS^

Plan of the Talk

- Introduction

- Architecture of CLaDA-BG-Research

- Functionalities

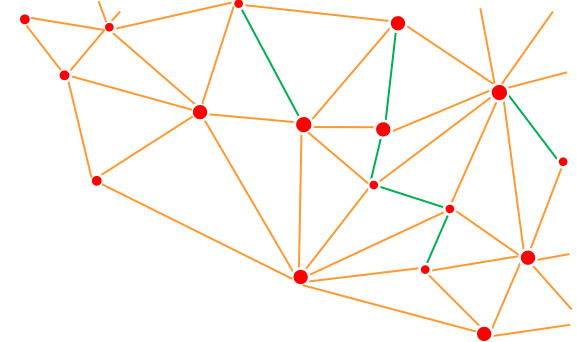- Usecase: Streets Through Sofia

Research Infrastructure

"Research infrastructures are complex agglomerations of
- knowledge,
- data,
- people, and
- services

that bring together diverse resources for a wide user base and make these resources (re)usable and available for an appropriately long term in order to support research (either individual or collaborative) and share the results of that research."
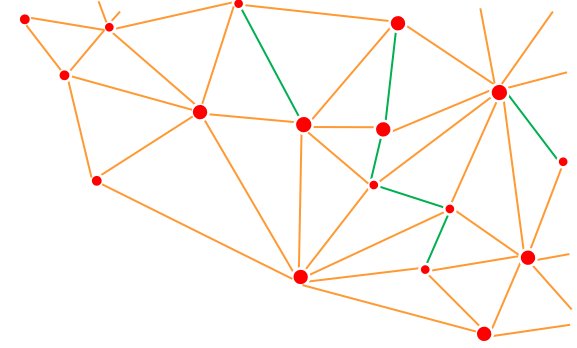
PARTHENOS Project (H2020 project of DARIAH and CLARIN)

Data and Knowledge

- Data – documents in different media and formats: text, hypertext, audio, video, images

- Knowledge – interpretation of the data with respect to their creation, content and interconnection

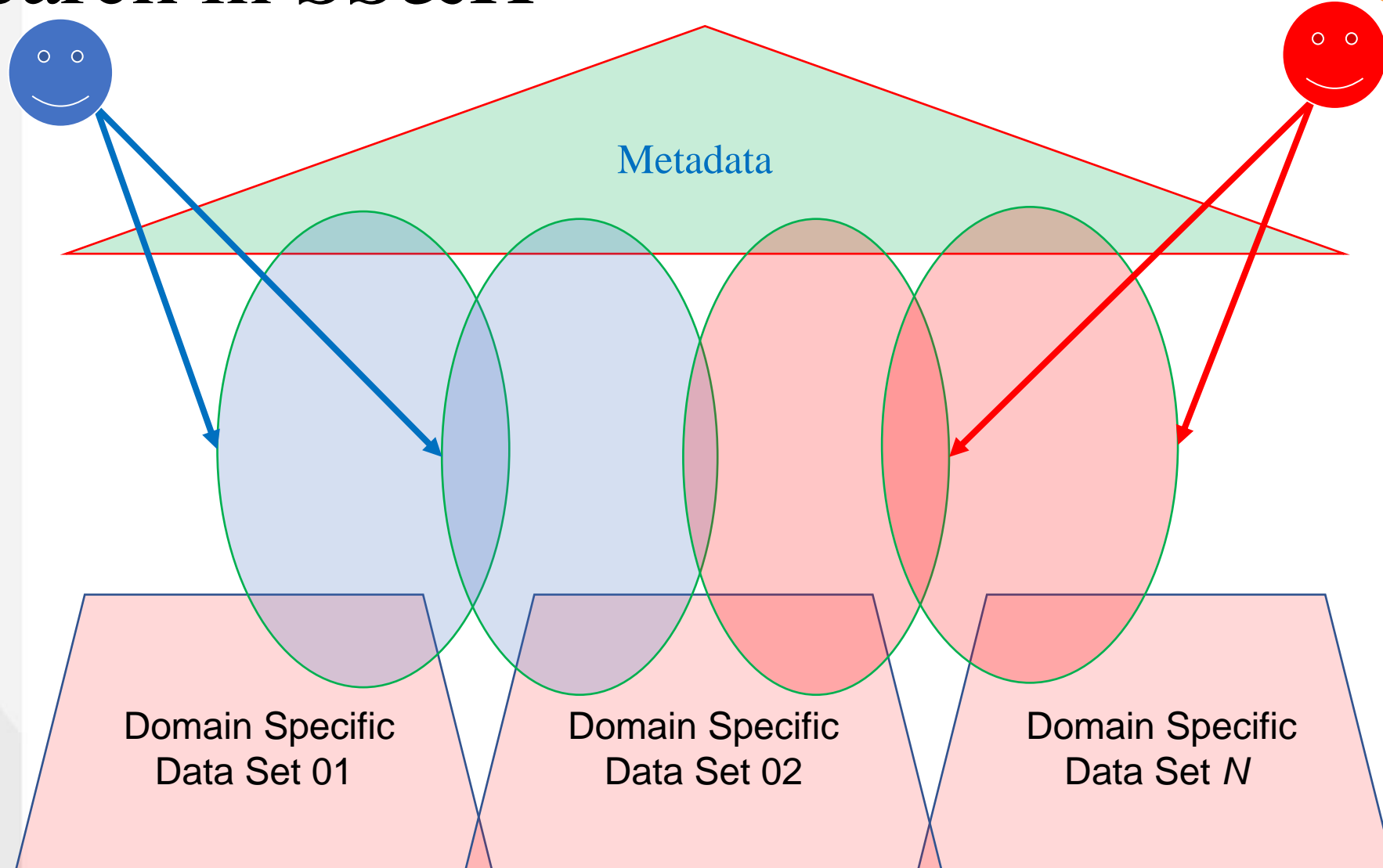- CLaDA-BG assumes that the main quantity of data is unstructural and non-homogeneous text data

„Language Technology is a core Big Data technology. Growth in the volume and variety of data is mostly due to the accumulation of unstructured text data; in fact, up to 80% of all data is unstructured text data." http://roadmap2018.esfri.eu/media/1066/esfri-roadmap-2018.pdf (p. 108)
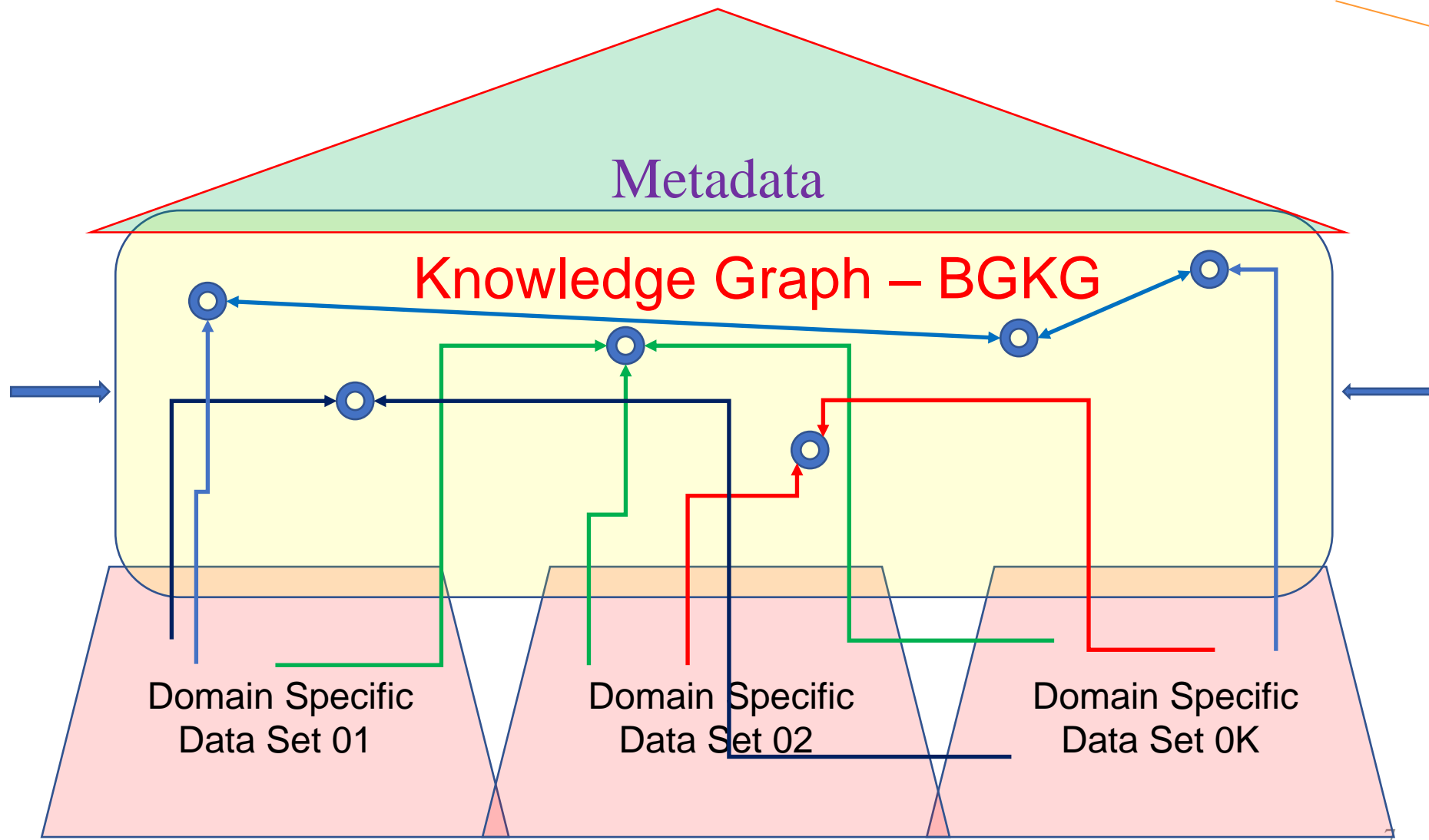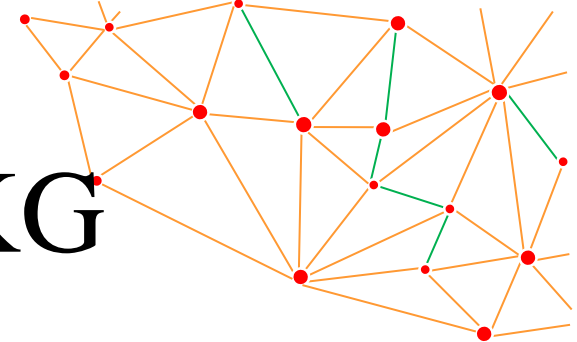
Research in SS&H

- A huge variety of research objects

- The research objects frequently share the same information

- A wide range of specific technologies for creation, generalization, storage, search, etc.

- The research in SS&H requires we to put the various types of data in the context of each other

- We name this process of integration of the data *contextualization*

Research in SS&H

Metadata
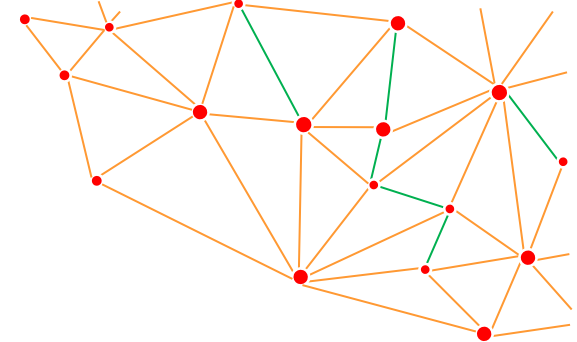
Domain Specific
Data Set 01

Domain Specific
Data Set 02

Domain Specific
Data Set *N*

Data Hierarchy in CLaDA-BG – BGKG

Metadata

Knowledge Graph – BGKG

Domain Specific Data Set 01

Domain Specific Data Set 02
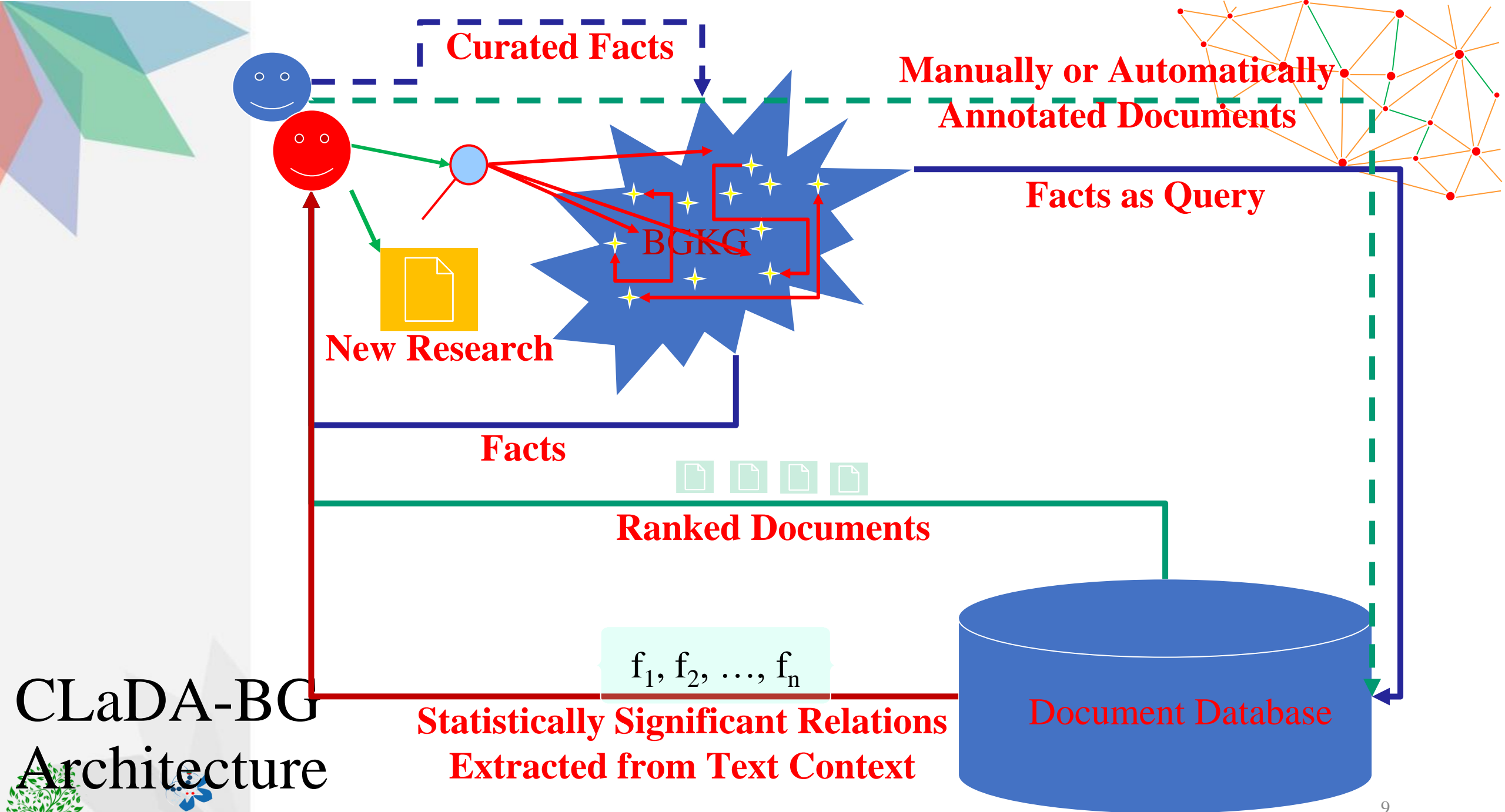
Domain Specific Data Set 0K

CLaDA-BG Contextualization

We consider contextualization as a network of interlinked descriptions of people, events, geographical entities, objects, documents, authors, opinions, etc.

- People – biographical data – events in the their life, their roles
- Geographical Entities – history of countries, cities, etc., the names have their own life
- Objects – creation, materials, form, discovery
- Events – time, place, participants, connection to other events
- Documents – authors, contents, opinion about peoples, events, …
- Text as main source of information
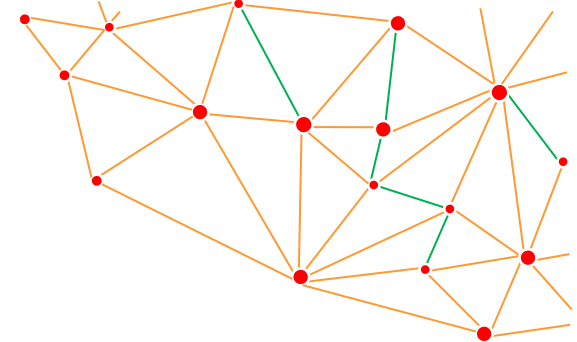- Technology – Linked Open Data and Knowledge Graph

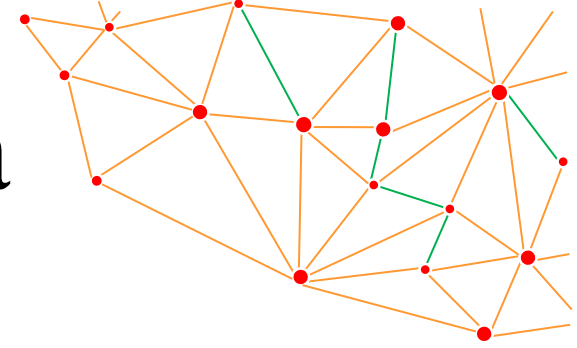CLaDA-BG Architecture

Implemented Elements of the Architecture

- BGKG + BTB-WN and corresponding software infrastructure

- Document database – indexing on NE IDs and terms from BTB-WN

- CLaDA-BG-Research Document Editor
  - Creation of internal representation of external documents (Document Referent)
  - Create of metadata to the internal document
  - Annotation of the document – manual and automatic
  - Creation of new documents:
    - BGKG Entity Document
    - Research papers
    - Research notes
    - etc

CLaDA-BG Research Projects

- Goal is to use the current implemented infrastructure to support different SS&H project providing support for all the stages of the project
  - Processing the necessary archive and other type documents
  - Annotation of relevant named entities and key words, terms
  - Linking to BGKG and BTB-WN and terminology
  - Linking to other resources like bibliographies
  - Extensions of existing resources
- Example of such project: Streets Through Sofia

BulTreeBank   CLaDA BG

Street Names/Name Changes in BG, a Complex and Multifaceted Story

- Street names as an essential component of urban memory and identity (yet sometimes underestimated); symbolic value, symbolically loaded, more than just instrument for practical spatial orientation within the city

- Streetscapes used in many places and various epochs to inscribe ideologies and dominating visions of the past, showing an official version of the past; two large renaming waves in the 1940s and the 1990s and many other smale-scale name changes

Contested Names, Contested Central Avenues
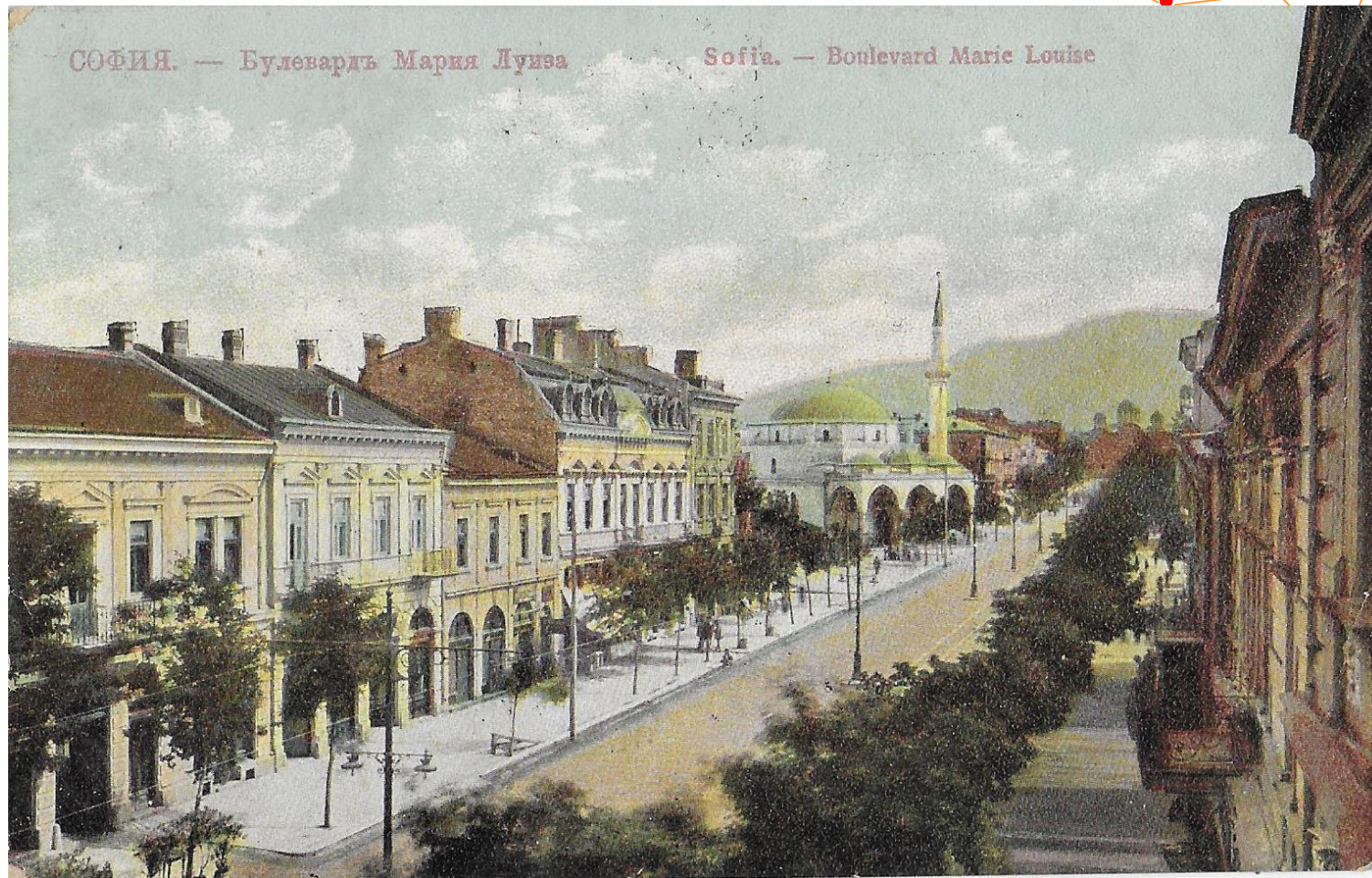## Maria Luisa blvd.

Balkana

Lomska

Vitosha (1883-1893)

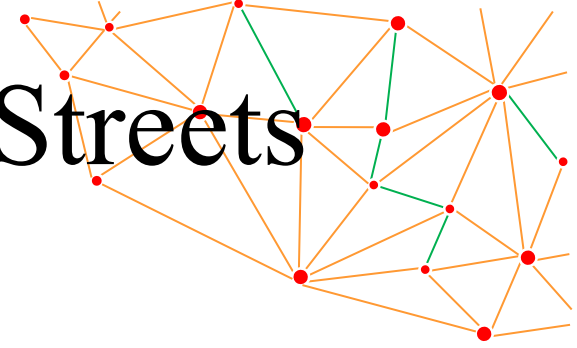Dimitar Petkov (1893-1895)

Vitosha (1895-1899)

Maria Luisa (1899-1946)

Georgi Dimitrov (1946-1992)

Maria Luisa (since 1992)

СОФИЯ. — Булевардъ Мария Луиза   Sofia. — Boulevard Marie Louise
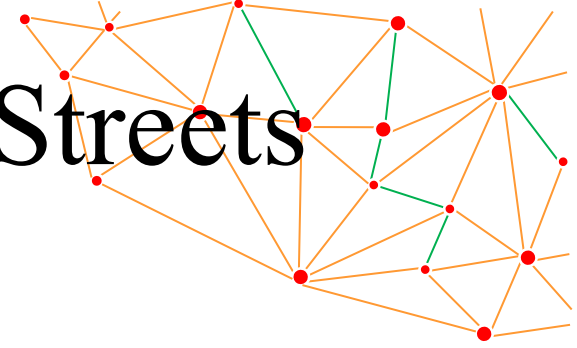
BulTreeBank CLaDA BG

Model for describing History of Sofia's Streets

1. Name (current name)

2. Name changes (historical names)

3. Documents related to historical names/biography of the street

4. Physical parameters: coordinates of different points, width, numbers (identifying plots and buildings), other characteristics, intersections with other streets, boundaries through objects

5. Important buildings and sites (existing/non-existing)

6. People who were related with the street

7. Important events that happened on the street

8. Links to other objects, including links to info for historical persons or geographic objects

9. Connections with neighborhoods, parts of the city

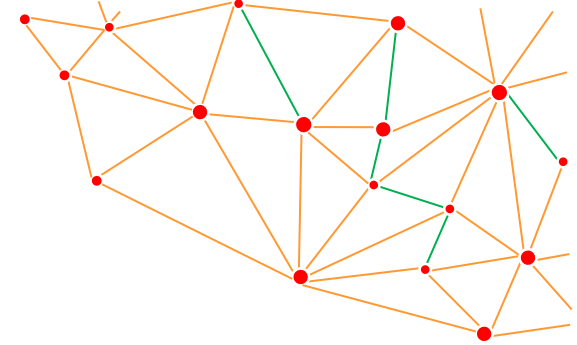# Model for describing History of Sofia's Streets

1. Name (current name)

2. Name changes (historical names)

3. Documents related to historical names/biography of the street

4. Physical parameters: coordinates of different points, width, numbers (identifying plots and buildings), other characteristics, intersections with other streets, boundaries through objects

5. Important buildings and sites (existing/non-existing)

6. People who were related with the street

7. Important events that happened on the street

8. Links to other objects, including links to info for historical persons or geographic objects

9. Connections with neighborhoods, parts of the city
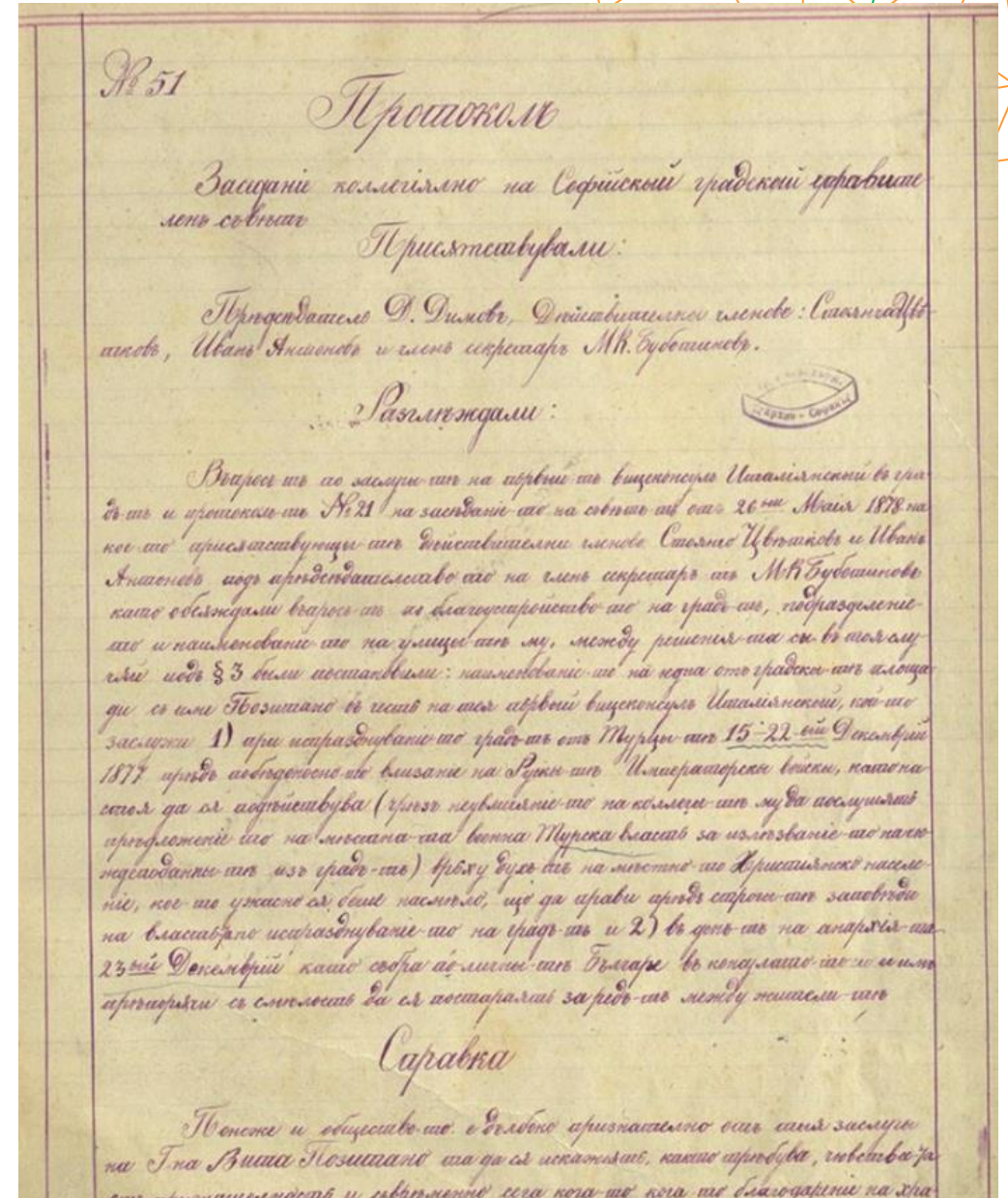
Available Materials

- Archival documents, incl. city-council proceedings; city-council proceedings officially sanctioned and discussed name changes

- Old city maps: convenient overview of the situation in various epochs

- Pictures

- Stories, memoirs
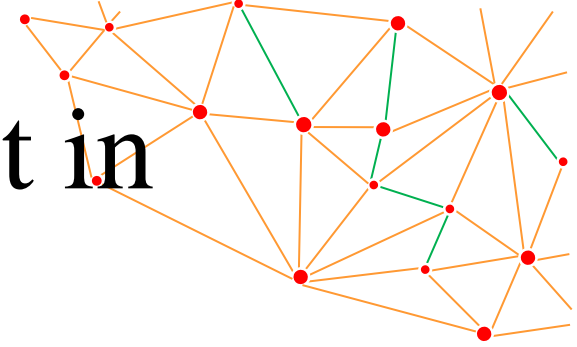
Available Materials: some challenges

Naming a square after Italian diplomat Vittorio Positano, December 1878

- street names among the first activities of new city council, as early as 1878
- well described the contribution of Positano to the city
- the document mentions popular sentiments – "people are deeply grateful"

Informs about the creation of special committee for naming streets, established in the same year
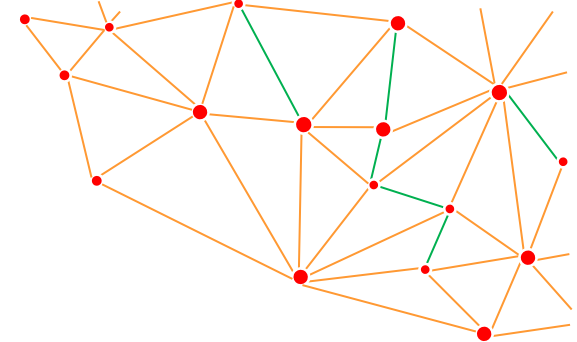
Example of One Street: Pirotska Street in Downtown Sofia

An important commercial artery starting from a main square in front of the Central Public Baths, a place with intensive traffic since the Ottoman period

- Intersections with several other central arteries, one of the first streets with a tram route since the beginning of the 20th century.
- A number of important buildings, including three schools, one of the newly-built churches, a bank.
- The street the central artery of the Yuchbunar neighborhood, historical part with many lower-class inhabitants and migrants. Several colorful descriptions of the life of the street during the first half of the 20th century

BulTreeBank   CLaDA BG

Conclusions and Future Work

- We defined CLaDA-BG Architecture including BGKG as a contextualization device and annotated document database as a research repository
- CLaDA-BG Architecture is already partially implemented with respect the combination of BGKG+BTB-WN
- Defining of new conceptualization and functionalities via implementation of small project within current implementation
- Usecase: Streets Through Sofia