

# Automatic Georeferencing and Search on Structured Georeferenced Bibliographic Data

Dragan Ivanović<sup>1</sup> Marjan Čeh<sup>2</sup>

<sup>1</sup>University of Novi Sad, Serbia

<sup>2</sup>University of Ljubljana, Slovenia

1st International KEYSTONE Conference, 2015

# Outline

1 Introduction

2 Challenges

# Georeferencing

- Georeferencing means associating something with locations in physical space
- GEOPOLO is a software system for georeferencing scientific-research publications, i.e. it is a system which enables joining geospatial data to scientific-research publications
- Visualization of publications by geographic regions and search by geographic data
- Currently, librarians manually add geospatial data to publications which is a weak point of the system
- Automatic extraction of toponyms could significantly improve the system and search of the stored data.

# Georeferencing

- Georeferencing means associating something with locations in physical space
- GEOPOLO is a software system for georeferencing scientific-research publications, i.e. it is a system which enables joining geospatial data to scientific-research publications
- Visualization of publications by geographic regions and search by geographic data
- Currently, librarians manually add geospatial data to publications which is a weak point of the system
- Automatic extraction of toponyms could significantly improve the system and search of the stored data.

# Georeferencing

- Georeferencing means associating something with locations in physical space
- GEOPOLO is a software system for georeferencing scientific-research publications, i.e. it is a system which enables joining geospatial data to scientific-research publications
- Visualization of publications by geographic regions and search by geographic data
- Currently, librarians manually add geospatial data to publications which is a weak point of the system
- Automatic extraction of toponyms could significantly improve the system and search of the stored data.

# Georeferencing

- Georeferencing means associating something with locations in physical space
- GEOPOLO is a software system for georeferencing scientific-research publications, i.e. it is a system which enables joining geospatial data to scientific-research publications
- Visualization of publications by geographic regions and search by geographic data
- Currently, librarians manually add geospatial data to publications which is a weak point of the system
- Automatic extraction of toponyms could significantly improve the system and search of the stored data.

# Georeferencing

- Georeferencing means associating something with locations in physical space
- GEOPOLO is a software system for georeferencing scientific-research publications, i.e. it is a system which enables joining geospatial data to scientific-research publications
- Visualization of publications by geographic regions and search by geographic data
- Currently, librarians manually add geospatial data to publications which is a weak point of the system
- Automatic extraction of toponyms could significantly improve the system and search of the stored data.

# Specific challenges

- How NLP and data mining techniques can be used in order to find out toponyms (geospatial keywords) from a structured data source in order to enable automatic georeferencing of those data?
- Dealing with toponyms which represent homonyms - Moscow is the capital city of Russia or a city in Kansas?
- Can Google Maps JavaScript API or some similar API be used for automatic georeferencing keywords from a query in order to enable keywords-based search on georeferenced structured data source even though the certain keyword from the query is not mentioned in some data which geospatial data are close to or are in relation superset-subset with geospatial data of the certain keyword?



# Specific challenges

- How NLP and data mining techniques can be used in order to find out toponyms (geospatial keywords) from a structured data source in order to enable automatic georeferencing of those data?
- Dealing with toponyms which represent homonyms - Moscow is the capital city of Russia or a city in Kansas?
- Can Google Maps JavaScript API or some similar API be used for automatic georeferencing keywords from a query in order to enable keywords-based search on georeferenced structured data source even though the certain keyword from the query is not mentioned in some data which geospatial data are close to or are in relation superset-subset with geospatial data of the certain keyword?

## Specific challenges

- How NLP and data mining techniques can be used in order to find out toponyms (geospatial keywords) from a structured data source in order to enable automatic georeferencing of those data?
- Dealing with toponyms which represent homonyms - Moscow is the capital city of Russia or a city in Kansas?
- Can Google Maps JavaScript API or some similar API be used for automatic georeferencing keywords from a query in order to enable keywords-based search on georeferenced structured data source even though the certain keyword from the query is not mentioned in some data which geospatial data are close to or are in relation superset-subset with geospatial data of the certain keyword?

# Scientific fields

- **Interdisciplinary open issue**
- Information retrieval, NLP, data mining, manual and semi-automated georeferencing, GIS, collection management, software engineering, protocol standards, data standards, etc.

# Scientific fields

- Interdisciplinary open issue
- Information retrieval, NLP, data mining, manual and semi-automated georeferencing, GIS, collection management, software engineering, protocol standards, data standards, etc.

# Questions / Discussion

- Thank you for your attention!!!