AD06: Automatic Construction and Open Source Use of Trustworthy Knowledge Graph of Ancient Chinese and Italian Medical Texts

Emanuele Frontoni  
University of Macerata

Weimei Pan  
Tianjin Normal University,
Outline

01 Background
02 Analytical techniques versus ancient medical texts
03 Research goals
04 Expected products
Background

Ancient **Chinese** medical texts: Copied or printed on paper before 1911.

There are as many as 8,663 kinds of ancient Chinese medical texts dated before 1912 survived, accounting for 6% of all ancient Chinese books in China.

———Xue Qinglu, 《General Catalog of Chinese Ancient Medicine Books》
Chinese Medicine will help us conquer life-threatening disease worldwide, and that people across the globe will enjoy its benefits for health promotion.

Tu Youyou
If Tu Youyou had the assistance of AI tool, would she find the cure sooner or later?
Opinions on Promoting the Work of Ancient Texts in New Era

“...... greater efforts on the work of ancient texts, to properly protect, develop and utilize these precious treasures of our motherland.”

“...... actively carry out research and practice on text structuring, knowledge systemization, and intelligent use of ancient texts”
Challenges when utilizing ancient medical texts

- Unstructured text, complex characters, and classical Chinese
- Use of different expressions
- Fuzzy words
- Information fragmentation
- Tacit knowledge
Techniques have been used to analyze data in the ancient medical texts

- Database construction
- Cognitive linguistic analysis
- Fuzzy logic
- Data mining
Artificial Intelligence technology

- Machine learning algorithm
- Deep learning algorithm
- Knowledge graph
Knowledge Graph

Semantic network that describes objective entities, concepts, and the relationships between them in the real world. 

(Nickel et al., 2016)

- Map-based method of knowledge representation and organization

(Gao et al, 2020)
Opinions on Accelerating Scenario Innovation and High-level Application of AI to Promote High Quality Economy Development

“……insufficient design of major scenario system, insufficient openness of scenario opportunities, ecology for scenario innovation not well developed, and others……”
Italy’s concept and theory of cultural heritage protection and knowledge graph technology are among the top in the world.

- Venice Time Machine
- BabelNet
- ArCo: the Italian Cultural Heritage Knowledge Graph
Overall goal

This project aims to focus on terminology and study the underlying mechanism of trustworthy knowledge graph by examining the generation mechanism of trustworthy, explainable, and computable knowledge graphs, integrating lexical chain model and deep learning technology to develop usable, trustworthy domain knowledge graphs techniques for ancient medical texts, and using pandemic prevention and regimen as examples to develop demo applications.
Objectives

01
Trustworthy knowledge graph theories to improve the explainability and computability of knowledge graph

02
AI technologies for mining knowledge within traditional Chinese and Italian medical texts

03
Scenario-based open source demo applications
Interdisciplinary research that integrates knowledge from the fields of language, cognitive science, and computing science will be conducted to understand the mechanism of trust and provide bases for the development of trustworthy knowledge graph.
Trustworthy knowledge graph is essentially a sense-making process.

“Attention” research in cognitive science.

- Attenuation theory (Anne Treisman)

  “the attenuation of unattended stimuli would make it difficult, but not impossible to extract meaningful content from irrelevant inputs, so long as stimuli still possessed sufficient "strength" after attenuation to make it through a hierarchical analysis process.”
The process of knowledge graph generation is actually the process of human beings filtering information.

Words and sentences have different cognitive loads on people.

Graph generation often uses multiple tasks to execute in parallel,
02 Technologies for mining knowledge within traditional Chinese and Italian medical texts

(1) Restoring ancient texts using deep neural network technology

(2) Developing technology for the construction of knowledge graph based on lexical chain

(3) Methods for enhancing trustworthy knowledge graph based on entity linking and link prediction.
(1) Cross-language terminology service of Chinese and Italian traditional medical science.

(2) Designing “Digital Ricci” using metahuman technology.
Map terminologies in Chinese and Italian ancient medical texts and through which to aggregate medical knowledge in China and Italian ancient medical text.

Pandemic prevention (public health) and regimen (personal health care)
Matteo Ricci (Li Madou)  
(1552-1660, born in Macerata)

- An Italian Jesuit priest and one of the founding figures of the Jesuit China missions, also one of the first Western scholars to master Chinese script and Classical Chinese.
- Pioneer of Chinese-Western dialogue and cultural exchanges

“Digital Ricci” : Automatic questioning and answering service to the general public on Chinese and Italian medical information.
Thanks!

Emanuele Frontoni
emanuele.frontoni@unimc.it

Weimei Pan
weimeipan3@gmail.com