



# How to organize and reveal the knowledge in digital resources of cultural heritage? -- A case study of Dunhuang Grottoes

Xiaoguang Wang, Wanli Chang, Hongyu Wang

School of Information Management

Wuhan University, Hubei, China

wanghongyu@whu.edu.cn

JCDL 2019 Workshop 3, June 6

# CONTENT

01

Introduction

02

Graph Model Construction

03

Data Process & Result

04

Conclusion & Future Work

# 01

PART ONE

## Introduction

---



# 1. Introduction——Background



Cultural heritage carries abundant cultural connotation and knowledge.

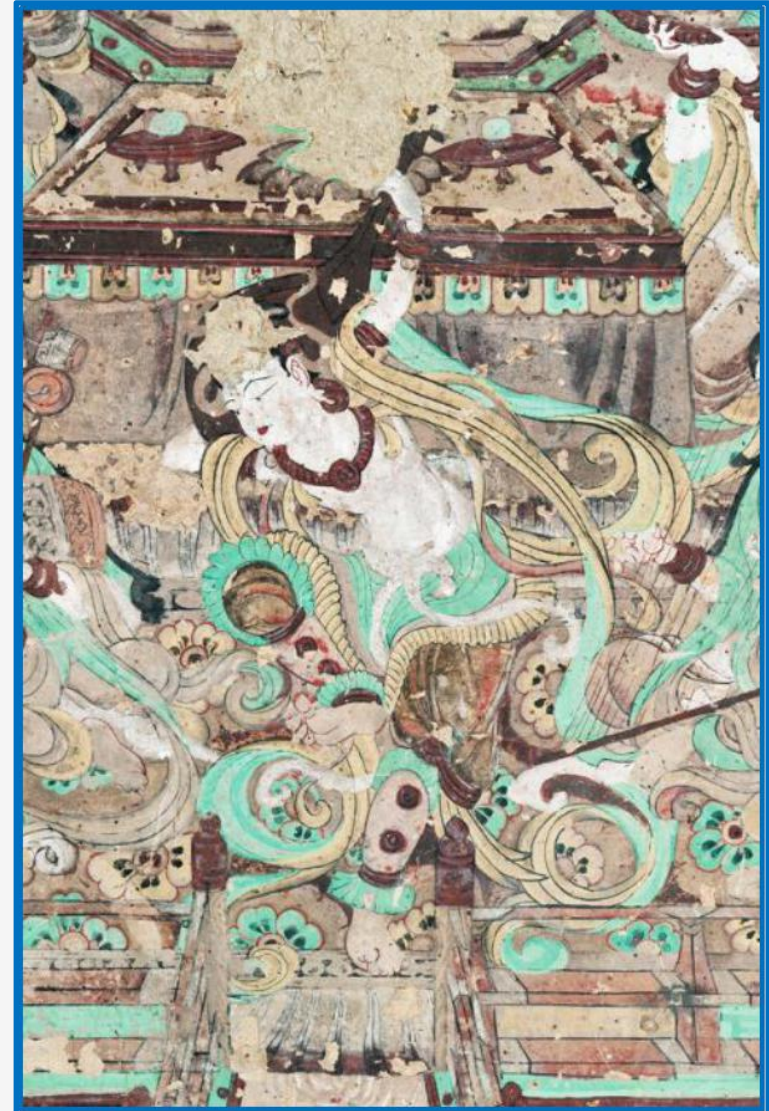


With the development of digital technologies, amounts of digital resources has been accumulated in the field of cultural heritage.



Digital resources of cultural heritage have brought a data-driven research pattern to humanists named DH.

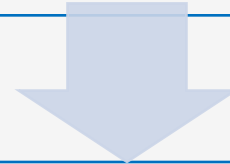
Advancing the research in digital humanities (DH) becomes a new mission of libraries, so as to provide high-quality digital resources and their relevant tools for C/O/A/U.





# 1. Introduction——Question

In the context of DH, the discovery and revelation of knowledge implied in cultural heritage, through the semantic and expandable organization of digital resources, will efficiently improve the inheritance, application and protection of cultural heritage.



As two commonly used KOS, metadata framework is inefficient in representing and reflecting relationships between knowledge units, and while, ontology model and/or RDF can reflect such relationships but are inability to qualify their instances.



In addition, when adding and associating new kinds of resources, the ontology model needs to be redesigned, so that the knowledge cannot be extended conveniently and efficiently.



# 1. Introduction——Our Work



This paper proposes an **attribute graph model** according to the characteristics of digital resources of Dunhuang, and constructs **a platform of knowledge discovery and revelation** by knowledge graph, which **organizes, links and stores** the domain knowledge contained in these digital resources semantically.

We provide **a novel method and example** for the knowledge organization and analysis **in digital resources of cultural heritage**.

# 02

PART TWO

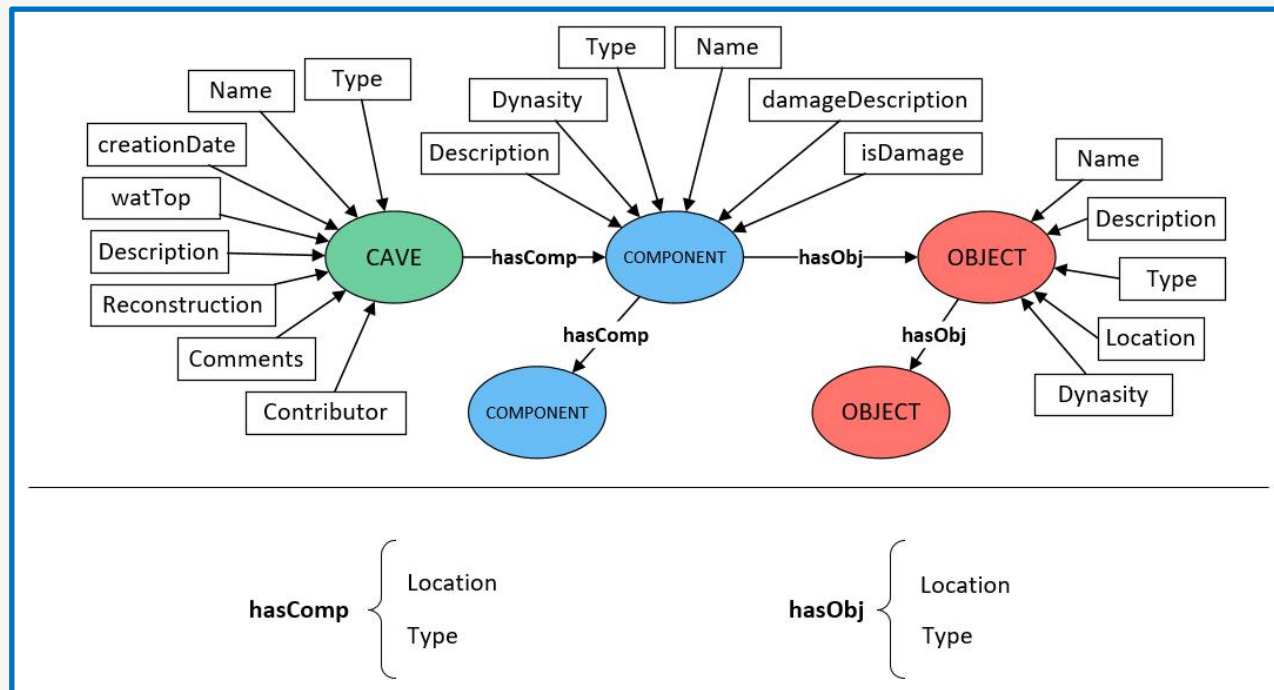
## • Graph Model Construction •



## 2. Graph Model Construction

The **KO model** of Dunhuang grottoes based on attribute graph is **the basis of knowledge integration** and **the premise of semantic revelation** of digital resources.

Dunhuang Grottoes and cultural elements therein are complicated in structure. For example, Grotto No.61 consists of an **arch** and a **main room** with **four walls**, which each wall contains **various objects** of cultural elements. Therefore, the relevant resources and elements are divided into three types according to the website of **e-dunhuang** and related **metadata standards** in this paper. Finally, we built the following KO model.







## 2. Graph Model Construction Cont.

The attribute graph model enables us to extract structured knowledge from digital resources of Dunhuang Grottoes for the purpose of constructing a knowledge graph.

Due to its rich content, multitude of objects in Dunhuang Grottoes, and the varied description in different literature and documents (e.g. "Sakyamuni" and "Tathagata" as refer to the same entity) , it's necessary to disambiguate the entities and fuse the knowledge in Dunhuang digital resources.

With reference to the category structure of AAT (Art & Architecture Thesaurus) vocabulary, a vocabulary list in the field of Dunhang Grottoes has already been produced by our team.

An efficient and accurate knowledge organization of Dunhuang digital resources is achieved in this paper with the aid of the vocabulary.

# 03

PART THREE

## Data Process & Result

---

### 3. Data Process & Result——Dataset

#### **Dunhuang Grottoes Content Catalogue**

, including detailed information for all 492 caves. All have been structured and stored in knowledge graph.

#### **Dictionary of Dunhuangnology**

, including 6,925 entries of Dunhuang manuscripts and academic studies under 60+ categories.

#### **Academic literatures of**

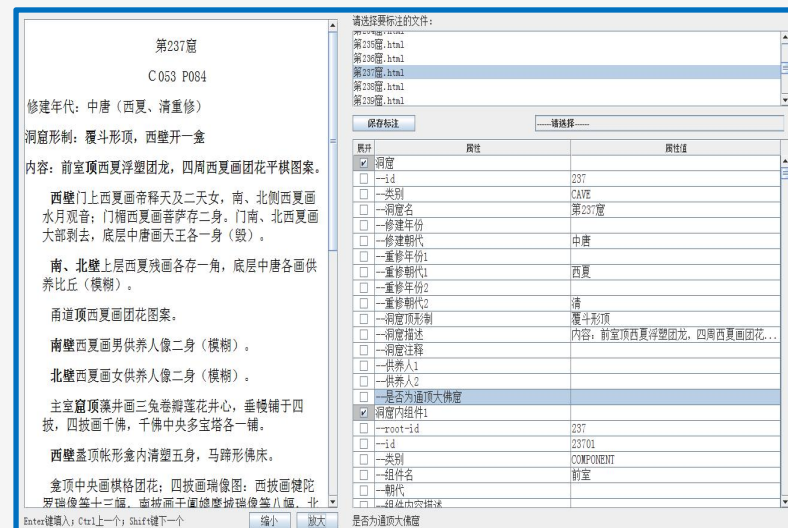
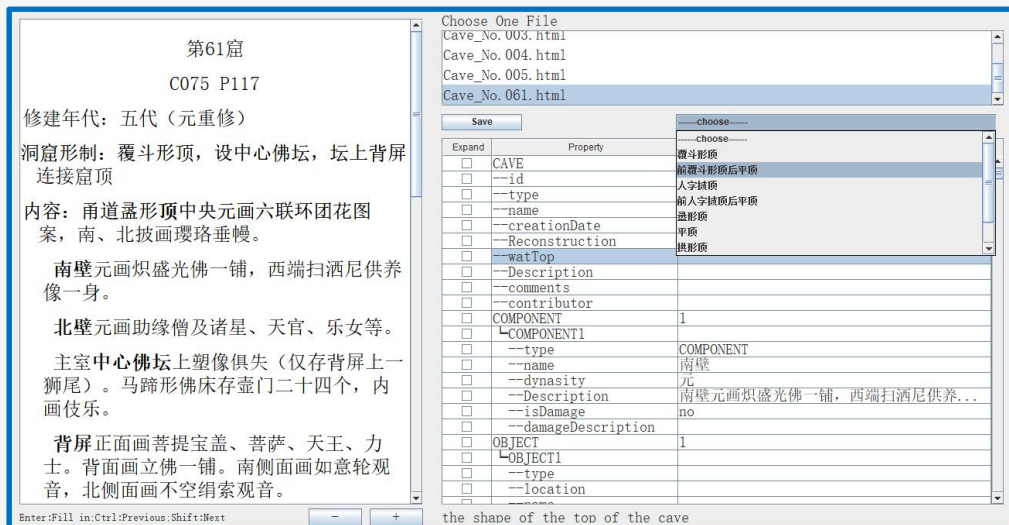
**Dunhuang research.** 7,929 e-books, 557GB. 5,083 papers, 784 of them have been textualized.



### 3. Data Process & Result——TDSAT

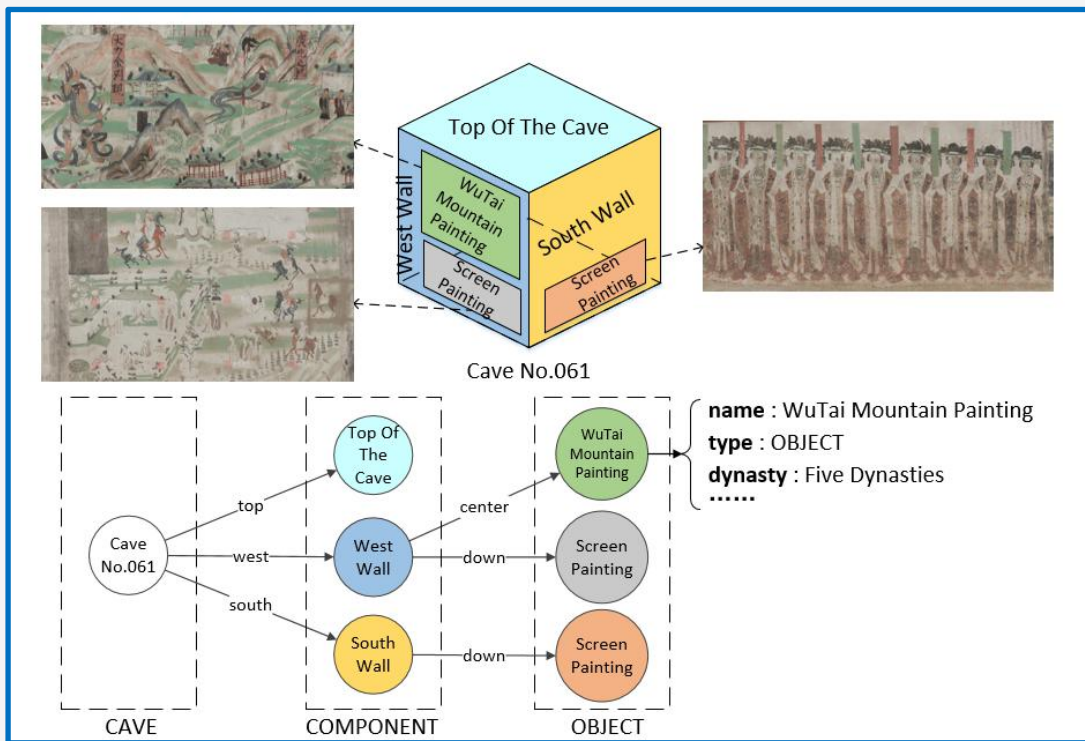
We developed the **Text Deep Semantic Annotation Tool** based on the proposed KO model to annotate the text of Dunhuang Grottoes and construct the Knowledge Graph for all 492 grottoes.

Then, we stored these data into the **Neo4j** after data-structuring.



With the aid of this tool, three types **entities** and their **relationships** are extracted, and their **attributes** are annotated, which realized the **knowledge organization** and **semantic association** of Dunhuang digital resources in the form of attribute graph.

### 3. Data Process & Result——Example



As a step further, the relevant multi-media resource are stored together with these knowledge in Neo4j.

Taking Grotto No.61 for example, the contents of its west and south wall are extracted according to the Dunhuang Grottoes KO Model, as shown in Figure.

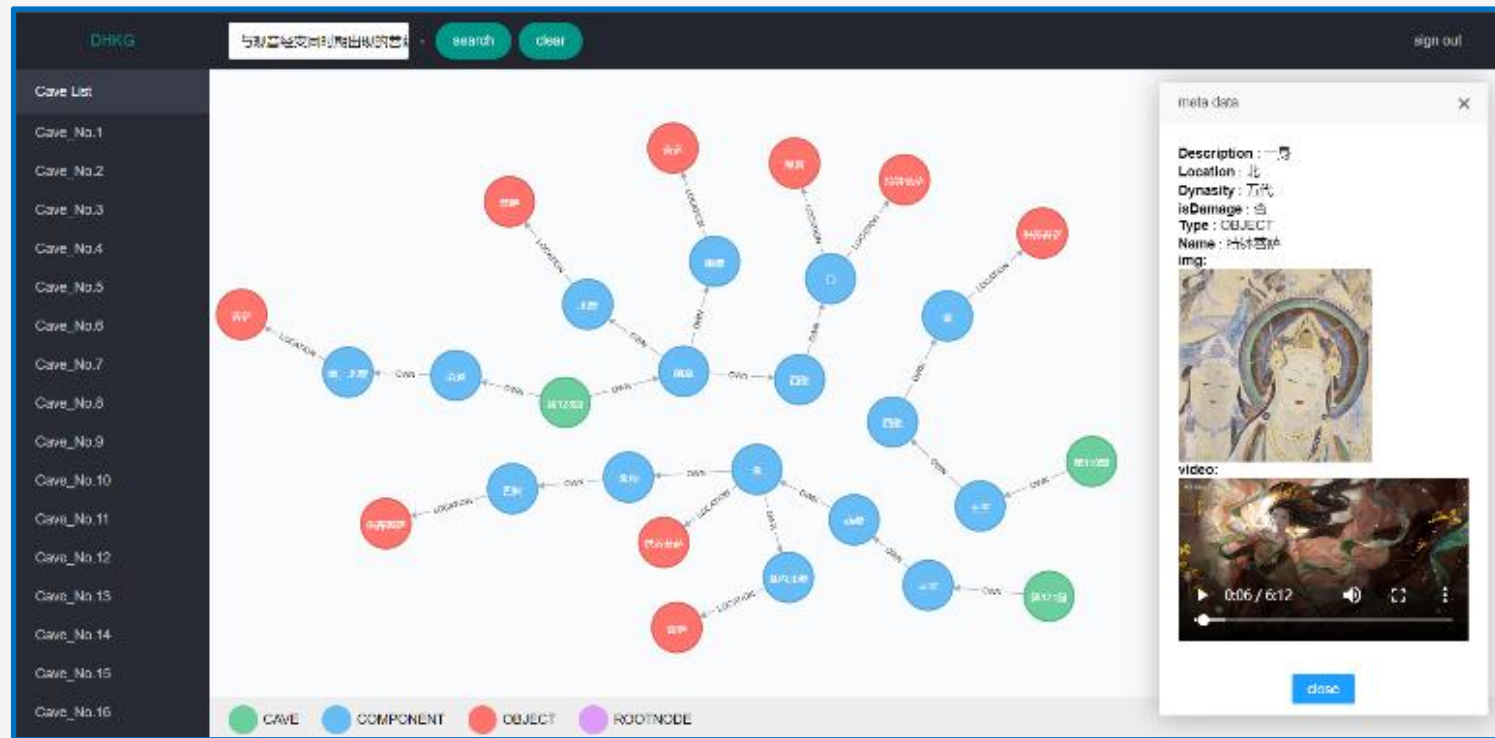




### 3. Data Process & Result——Result

The **platform** built in this paper with its function of knowledge navigation and semantic retrieval in Dunhuang Grottoes, can be applied to assist in further Studies.

For example, if the query of "**Bodhisattvas appearing at the same time as the Avalokitesvara Sutra**" is input, the answer will show all the suitable bodhisattvas with their **knowledge networks' structure, metadata, and associated multimedia resources** such as photos and videos. This result has shown in figure.



# 04

PART FOUR

## Conclusion & Future Work

---



## 4. Conclusion & Future Work

### Conclusion

Using Dunhuang Grottoes as example, this paper illustrates the process of attribute graph model-based knowledge organization and revelation of digital resources of cultural heritage.

Attribute graph model enables an handy and efficient semantic organization of entities in cultural heritage knowledge, their relationships and attributes, which further enables a fine-grained analysis and revelation of cultural heritage.

This paper provides another perspective for the organization of digital resources and the revelation of their cultural connotations. In addition, it provides a flexible way to consider the infrastructure development in digital humanities.

### Future Work

In the following, the research will attempt to further standardize the metadata of Dunhuang Grottoes on the basis of international standards such as CDWA, CIDOC-CRM, etc., and work towards the associated release and sharing of datasets.





THANK YOU

For Your Listening !

W u h a n  
University

Hongyu Wang