Acta Geologica Sinica (English Edition), 2019, 93(supp.1): 34–36

The Sedimentological Knowledge Tree: Significance, Method and Progress

LI Chao[1], HU Xiumian[1, *], HOU Mingcai[2], WANG Chengshan[3], YANG Jianghai[4] and the DDE Sedimentary Data Group

[1] *Nanjing University, Nanjing, Jiangsu 210023, China*
[2] *Chengdu University of Technology, Chengdu, Sichuan 610059, China*
[3] *China University of Geosciences (Beijing), Beijing 100083, China*
[4] *China University of Geosciences (Wuhan), Wuhan, Hubei 430074, China*

Citation: Li et al., 2019. The Sedimentary Knowledge Tree: Significance, Method and Progress. *Acta Geologica Sinica* (English Edition). 93 (supp. 1): 34-36

On the platform of the Deep-time Digital Earth Program (DDE), sedimentary data are essential for achieving its scientific objectives. These data will take stratigraphic units as their core data carrier, for quantitative or qualitative data analysis. The DDE Sedimentary Data Group is responsible for the management of the sedimentary data on the DDE platform and has now developed into a group of nearly 40 disciplinary experts. This group's main tasks are constructing sedimentary knowledge tree, guiding sedimentary data collection and storage, and mining the sedimentary data. The sedimentary knowledge tree is the most critical task in the early stage of DDE program. Several important agreements have been reached (as follows).

1 The significance of the sedimentary knowledge tree

A knowledge tree lies at the core of the DDE sedimentology platform and has the following aspects.

(1) Establishment of a multidimensional and comprehensive sedimentology database

Based on the sedimentary knowledge tree, data-platform engineers can build a well-structured and comprehensive sedimentary database to accommodate the massive data accumulated by sedimentologists over the past centuries including quantitative measurement data and qualitative descriptive data.

(2) Establishment of a unified subject language under the DDE framework

Sedimentary data contains a large number of concepts and terminology. Before collecting and mining vast amounts of data, scientists, developers, engineers and users must reach a unified understanding of these concepts. The knowledge tree and the definitions contained therein can serve as a unified subject language, reducing the risk of confused data interpretation and utilization.

(3) Establishment of system logic for big data analytics

Intelligent analysis of multidimensional sedimentary data is the scientific mission of the DDE sedimentology platform. A consistently structured sedimentary knowledge tree, and the logical relationships contained in it, can help developers build powerful data-analysis algorithms.

2 Building the sedimentary knowledge tree

After continuous discussions the Sedimentary Data Group reached agreement on the construction ideas and display forms of the sedimentary knowledge tree. It is presented in the form of a tree diagram. Elements such as hierarchies, branches, indicators, and links in the tree diagram represent the various concepts in sedimentology and the relationships among them. A complete knowledge tree must meet the following conditions.

(1) Maximum segmentation of research subjects.

Sedimentary research objects can be divided into multiple levels, such as sections, stratigraphic units, samples, heavy minerals, and single detrital minerals. The knowledge tree needs to cover research indicators for all levels of research objects.

(2) Maximum inclusion of research indicators and filling methods.

The knowledge tree needs to improve various concepts and indicators that are usually applied in sedimentary research continuously. Then the database developers will be able to reserve space to fill data with confidence and authority.

(3) Maximum coverage of data sources.

By including as many research objects and research methods as possible, the knowledge tree will cover as many previous topics and studies as possible to obtain sedimentary data in a comprehensive and cost-effective manner.

The common languages of the knowledge tree are English and Chinese, currently. Each node of the tree (a sedimentary concept or indicator) will be defined in both Chinese and English. Its definition and subdivision scheme (if any) will follow an internationally recognized scheme.

Through extensive trial and discussion, it was decided to use the X Mind: ZEN software as a general knowledge tree construction platform. This platform has a low learning cost, complete functions, and easy-to-understand pages, which meets the actual needs of the project's current stage.

In the construction of the sedimentary knowledge tree, four kinds of tools willused mainly to record four types of information. The 'Notes' tool will record internationally accepted definitions of a concept or indicator, the 'Summary' tool will record a reference to the indicator's subdivision or fill plan, the 'Label' tool will record a concept or indicator that lacks an internationally accepted definition, and the 'Relationships'

* Corresponding author. E-mail: huxm@nju.edu.cn

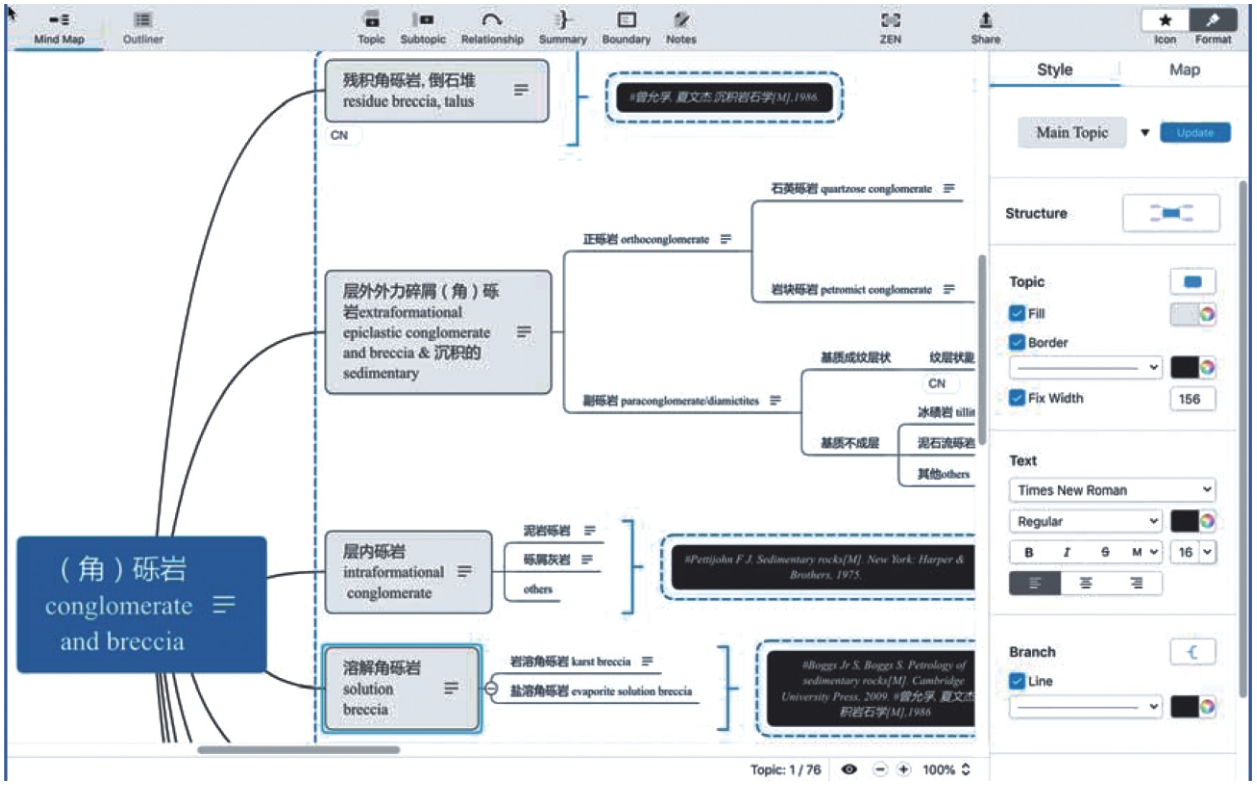

Fig. 1. XMind:ZEN software interface.
Upper part shows the tool bar, lower left displays the drawing area, and lower right is the format area.

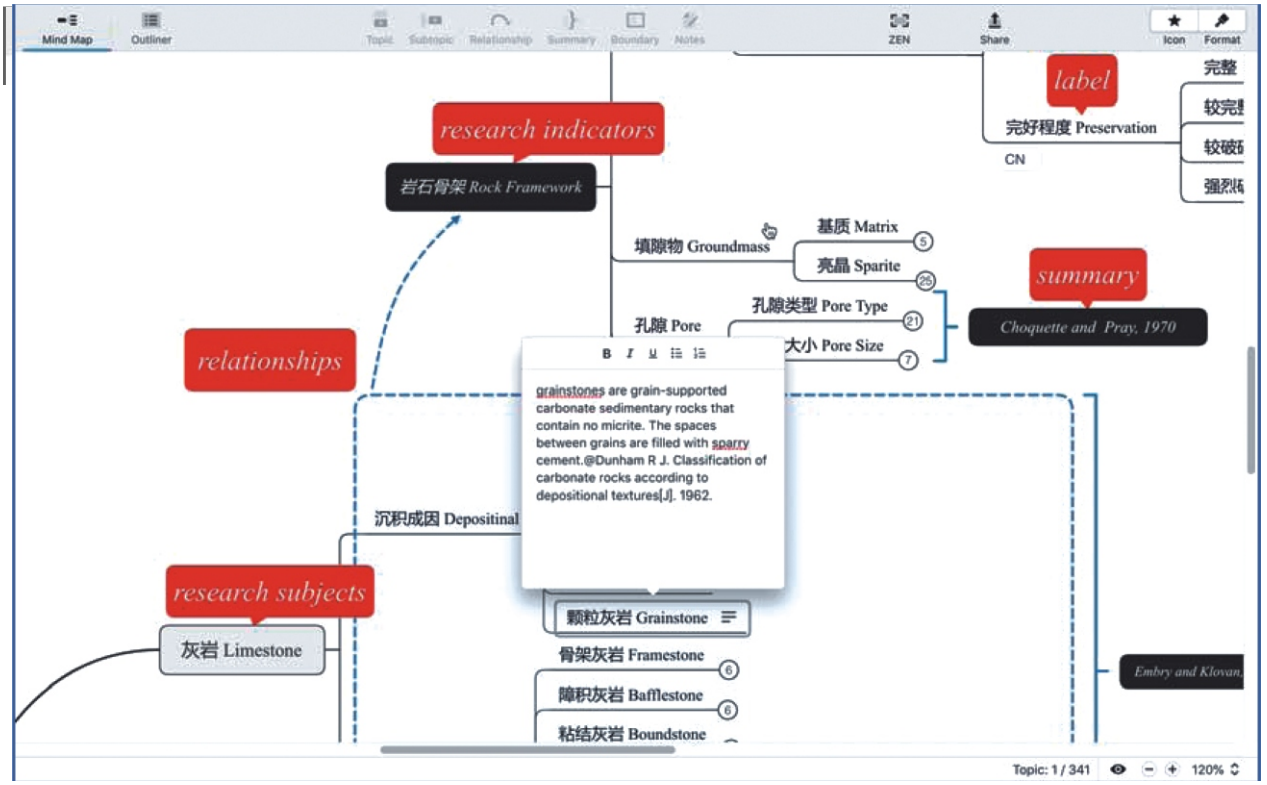

Fig. 2. Main tools and related information.
Summary tool, relationship tool and label tool are presented by red labels, and the note tool is represented by white editing box.

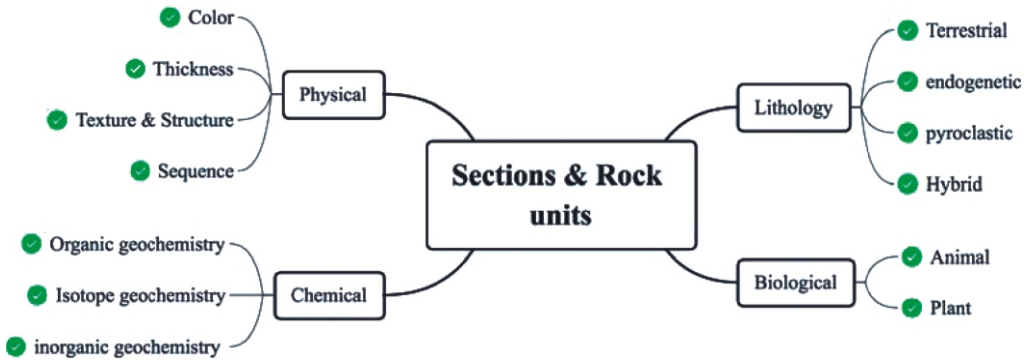

Fig. 3. Completion of sedimentary knowledge tree.

tool will establish a connection between research subjects and research indicators. Usage scenarios of the four tools are defined strictly, which is conducive to the formation of a general reading method for the tree.

3 Progress of the sedimentary knowledge tree construction

According to sedimentology's research directions, the construction task of the DDE sedimentary knowledge tree will be divided into four parts, namely physical data, chemical data, rock data, and biological data.

Physical data focuses on the physical properties of the rock formations (e.g., color, thickness, sedimentary structure and texture, sedimentary sequence). Chemical data focuses on the chemical properties of the sample (e.g., inorganic geochemistry, organic geochemistry, isotope geochemistry). Rock data focuses on the sample's lithologic features, including terrigenous clastic rocks, endogenetic sedimentary rock, pyroclastic rocks, hybrid sedimentary rocks, and sedimentologists' descriptions or measurements of various lithologies. Biological data focuses on fossils with sedimentary environmental implications and probably age determination.

After several months of discussion and work, the Sedimentary Data Group completed the construction of the preliminary plan of the sedimentary knowledge tree. Version 1.0 of sedimentary knowledge tree, both in Chinese and in English, is available for further development, suggestions, and comments.

Keywords: Sedimentology, Deep-time Digital Earth, Data Science, knowledge tree, tree diagram

Acknowledgements: Thanks to the DDE program for support, thanks to the members in DDE Sedimentary Data Group for their work and thanks to the valuable advice from sedimentary experts.

About the first author

LI Chao, male, born in 1990 in Baoding City, Hebei Province; Master; graduated from Chengdu University of Technology; Ph.D student in Nanjing University. He is now interested in the South China Block tectonics and big data analysis in Sedimentology. Email: lichaogeo@icloud.com; phone: 025-89680703.

About the corresponding author

HU Xiumian, male, born in 1974 in Nanchang City, Jiangxi Province; PhD; graduated from Chengdu University of Technology; Professor in Nanjing University. He is now interested in the study of sedimentary, tectonic and palaeoenvironmental evolution of the eastern Tethys ocean. Email: huxm@nju.edu.cn; phone: 025-89683002.