中国石油大学（华东）
CHINA UNIVERSITY OF PETROLEUM

Some thoughts on big data analytics in oil and gas field

J. Sheng & Z. Liu & Y. L. Bai

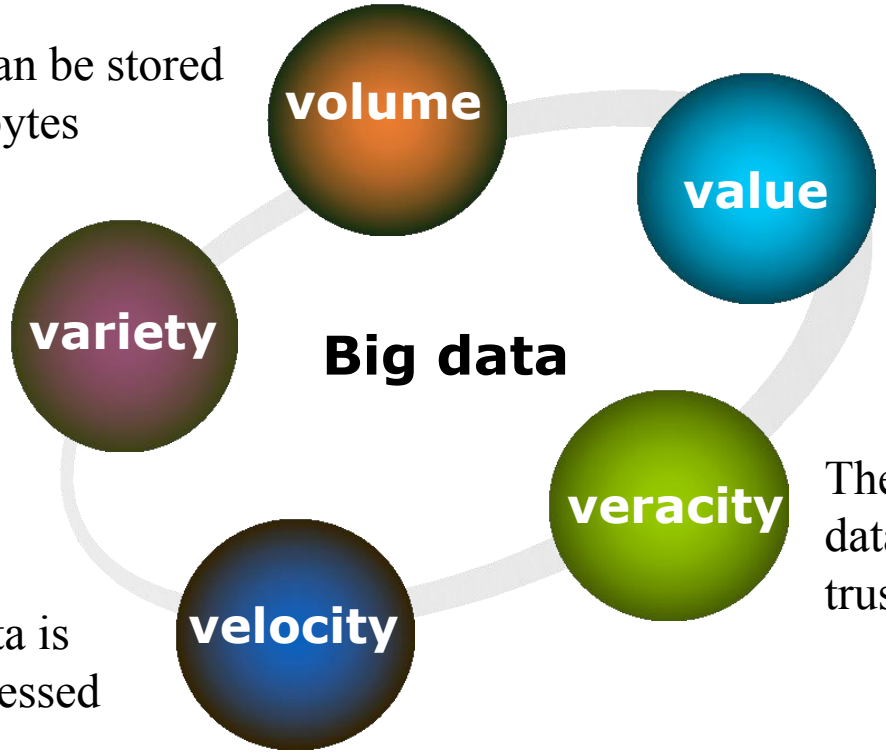
China University of Petroleum (East China)
Website: http://rg.upc.edu.cn/
E-mail:helloshengjie@163.com

Contents

How is big data different from what was done previously

5V's (volume, variety, velocity, veracity and value)

Large amount of data can be stored from terabytes to zettabytes



Extracting value from the data with the increased number of systems, size and storage.

structured, unstructured and semi structured data

The Integrity of data. data quality issues and trustworthiness of the information.

High speed in which data is created, stored and processed

How is big data different from what was done previously

Fundamental shift in thinking about data quality and volume

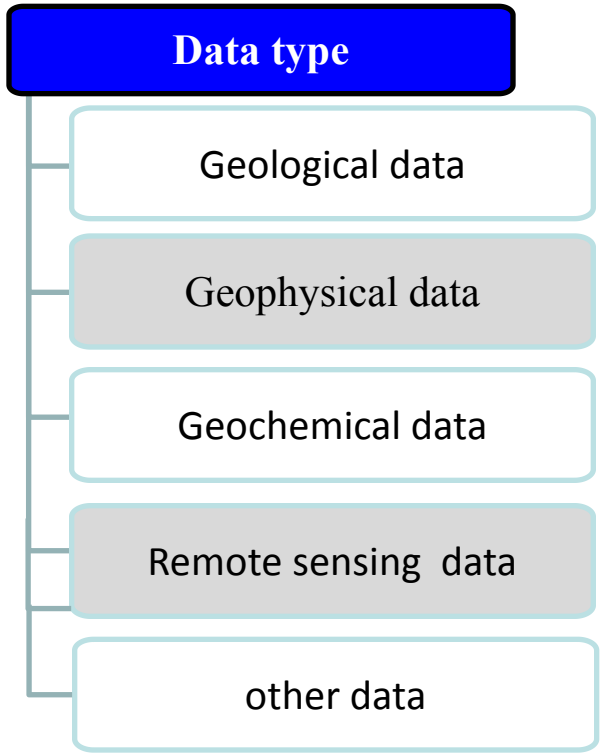
Traditional method → Big Data

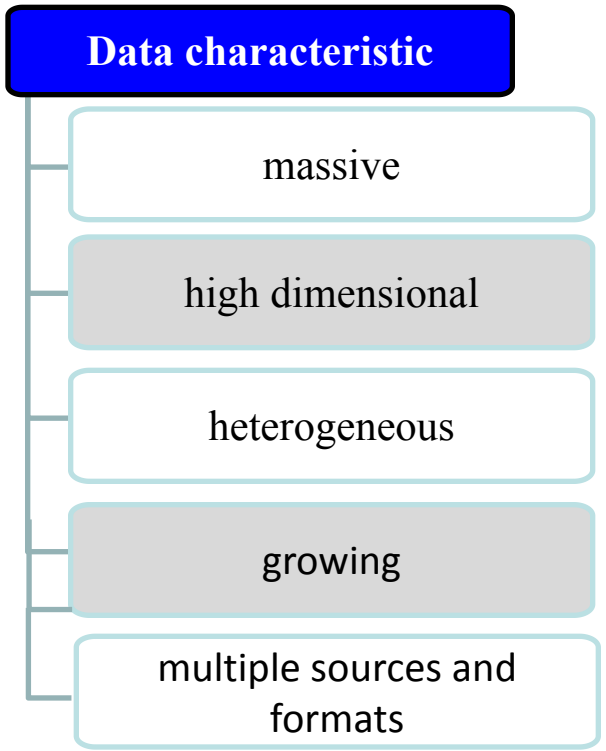
- The move towards Big Data has led to a much greater tolerance for messiness and imprecision.
- This more relaxed approach to Vagueness has been compensated, however, by much larger volumes of data.
- These much larger volumes of data are knowable to move with ever Increasing velocities such that a bewildering number of system variables can be monitored in nearly real-time.
- This data is coming from a wider variety of sources and in an increasingly broad array of formats.
- Big data methods pay more attention the correlation, rather than the causal relationship between variables.

Contents

What are the characteristics of big data in O&G

The oil and gas industry has it all — volume, velocity, variety, veracity and value



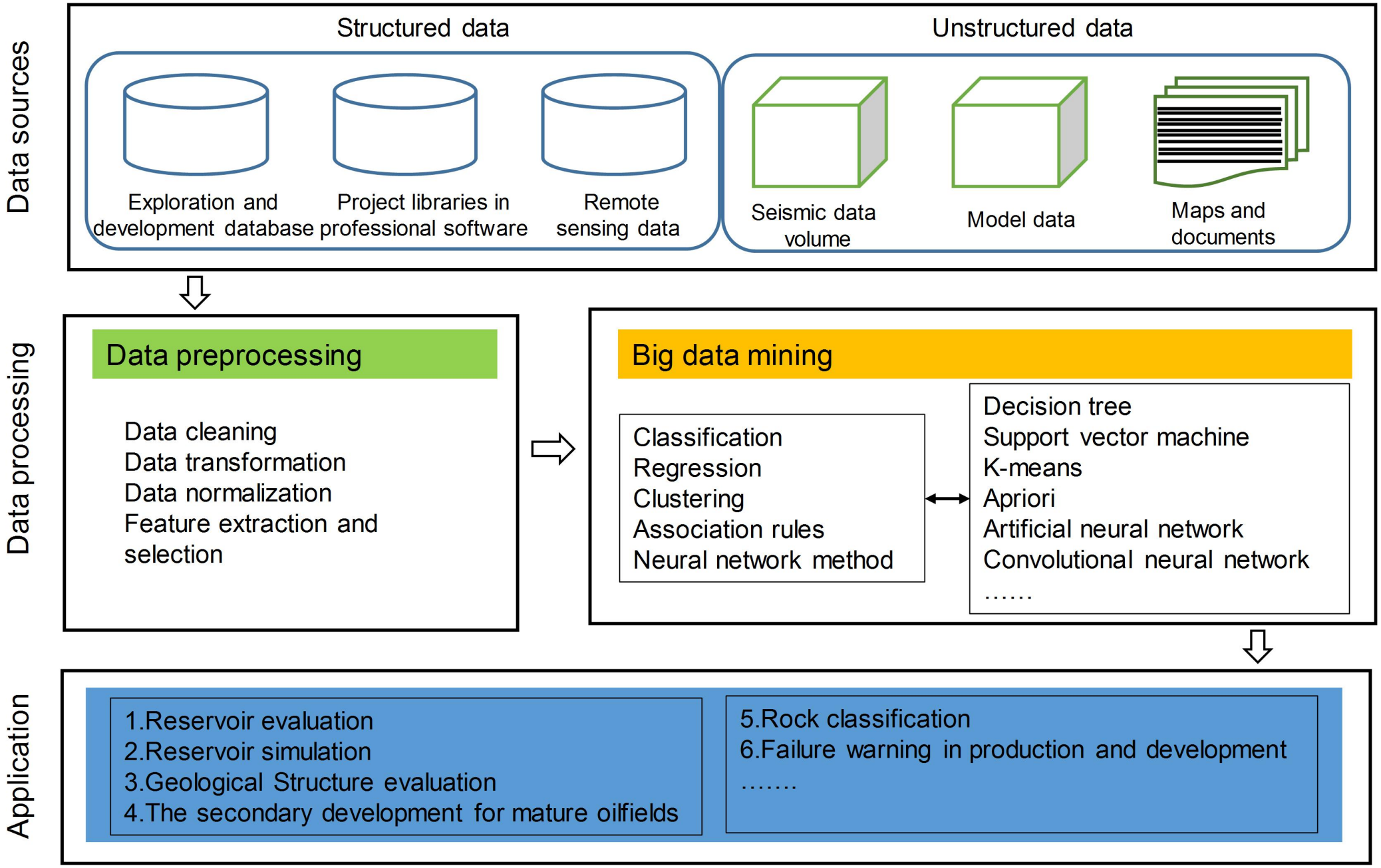


The difference between big data in O&G and other social data:

- Highly specialized data
- Multi-scale feature
- Qualitative description
- ……

This requires professionals to deal with big data in oil and gas field to solve different geological problems.

Contents



A general big data infrastructure in oil and gas field.

Contents

Key questions and research strategy

Key questions

1. • how to represent the domain knowledge for its application in oil and gas field?
2. • What kind of learning model to choose?
3. • how to get the big data ready for mining algorithm?

Research strategy

Ontology for knowledge representation: A formal, explicit specification of a shared conceptualization

Biomedical & medical fields

- Biomedical ontologies for data integration
  Smith et al ,2007
- Ontology rules for exploring medical and oral health data
  Shah et al ,2015
- Ontology approach to investigate toxicity mode of action
  Vitcheva et al,2013
- Using the Gene Ontology for Microarray Data Mining
  Pavlidis.et al ,2004

Web service

- An ontology-based framework to capture information from different data sources
  Izhar, et al, 2017
- Automatic ontology based user profile learning from heterogeneous web resources
  Hoppe.et al ,2013
- Service Discovery using Social Service Network with Domain Ontology and Workflow Awareness
  Siriweera et al, 2016
- Ontology-based service discovery for intelligent Big Data analytics
  Akila.et al ,2016

Oil and gas field

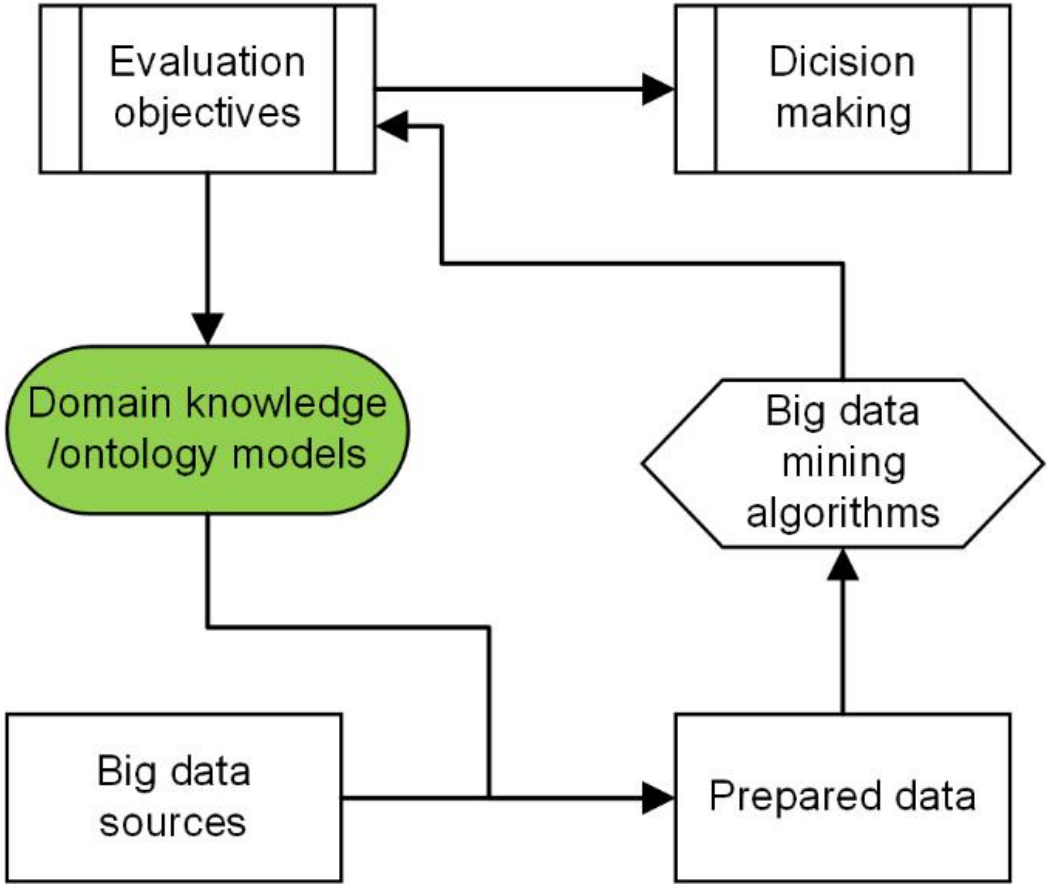
- Domain ontology for constructing knowledge base
  Izhar, et al, 2017
- Application of ontology in exploration well location deployment
  Yu,2013
- Ontology-based data warehousing and mining approach for PDE in basin scale
  Nimmagadda et al, 2015
- Integrated E&P business and knowledge management using ontology-based warehousing
  Nimmagadda et al, 2017

- Define the concepts and their relationships so as to remove any heterogeneity, allow syntactic and semantic interoperability
- Support knowledge sharing and reuse between different systems
- It is easy to expand ontologies.

Key questions and research strategy

Research strategy

(⑴)Determine evaluation objectives
(⑵)Construct ontology models
(⑶)Prepare data
(⑷)Mine big data
(⑸)Support decision making



Determine the score based on the value of each indicator according to prior knowledge

Objective	First level	Second level	score
Structure evaluation	Source rock	Thickness	1
		Organic matter abundance	0.8
		Organic matter type	0.6
		Organic matter maturity	0.4
	Reservoir	Sedimentary facies	0.2
		Lithology	0
		Thickness	……
		Physical property	……
	Caprock	Lithology	……
		Thickness	……
		Regional distribution	……
		Fault	……
	Trap	Type	……
		Area	……
		Closure	……
		Buried depth of high point	……
	Hydrocarbon migration	Migration channel	……
		Type of migration channel	……
		Time-space matching of source rock and trap	……
		Combinations of hydrocarbon generation and migration	……

Contents

Conclusion

I. Big data technology brings opportunities and challenges for the oil and gas field.

II. Based on the discussion on key issues and its research status in big data analysis, this paper proposes a knowledge-driven methodology for solving the key issues.

III. Domain knowledge and big data sources are integrated through ontology models. Future works would focus on case studies and parameter instantiation of ontology model.

Thank you for your attention !