The data turn of scientific cognition and the research program of philosophy of data

Xinrong Huang

Research Center of Management Philosophy, Jiangxi University of Finance and Economics, Nanchang, China

ABSTRACT

We entered the age of big data characterized by data revolution at the beginning of the 21st Century. In this new age, data has become an economic resource, and an important component of scientific cognition as well. Compared with natural language and logic language, data is a scientific language with higher accuracy and convenience. With the advent of the age of big data, ways of scientific cognition are bound to be transformed from the language turn of the 20th Century to the data turn of the 21st Century. And the possible conditions for this data turn have been prepared in terms of scientific premise, technical conditions, social background, and philosophical basis. At present, the ways of scientific cognition have begun to change from language to data, from logic to algorithm, from analysis to synthesis, from proof to discovery, and from causality to relevance. The language turn in the 20th Century will be replaced by the data turn in the 21st Century, and efforts shall be made to construct the philosophy of data, in which data is recognized as the object, algorithm as the tool, and synthesis as the method, and a wide variety of issues such as data and the world, data and language, data and algorithm, data and knowledge, data and truth, and data and ethics will be explored comprehensively.

KEYWORDS

Scientific cognition; Data turn; Philosophy of data

1 Introduction

Philosophy is the quintessence of the spirit of the times, with the brand of its own age. Due to the rise of mathematical logic at the turn of the 20th Century, people were able to make symbolic expression and logical analysis of human daily language, of which mathematical analysis was unlikely to be carried out in the past. Therefore, analytical philosophy appeared in the early 20th Century, and then the linguistic turn of scientific cognition took place. The 20th Century is called "the age of analysis" because of the dominant position of analytical philosophy (White, 1955). Now that we have entered the 2020s, how can the philosophy reflect the characteristics of our time? Should the analytical philosophy of the 20th century be held on to or should efforts be made to construct the philosophy of our own times? Based on the digital technologies such as digital electronics, computer, and Internet in the latter half 20th century, a series of new information technologies such as mobile network, big data, artificial intelligence, and blockchain are now flourishing rapidly in the 21st Century, and the age of big data has arrived. Digitalization, big data, data mining, and

knowledge discovery have become the characteristics of our times. Therefore, the ways of scientific cognition have been gradually transformed from analysis and reduction of the 20th Century to induction and synthesis of the 21st Century. We have entered the age of comprehensive innovation characterized by revealing the relevance and natural order through big data. Big data, dubbed as electricity and oil in the new age, has become the most important cognitive source of our time. With the advent of the age of big data and the ideological revolution of data, mining, algorithm, and intelligence, will the philosophy of our time undergo a transformation, just as the logic and analytical thinking of the early 20th Century brought about a philosophical revolution and formed a turn of analytical philosophy and language? Data is more scientific and accurate than natural language and logic language in describing phenomena, discovering laws, and predicting. Therefore, should our scientific cognition turn with the advent of the data age? Should we keep pace with the revolution of data science and data technology and construct a philosophy of data reflecting the spirit of the time? In this paper, the proposition of the data turn of scientific cognition in the age of big data is advanced, the necessity, possibility and directions of data turn are elaborated, and a preliminary research program for the philosophy of data is constructed.

2 The Necessity of Data Turn

The literal meaning of "turn" is that something originally has its own fixed development direction, but for some reason it causes directional changes, so it develops in other directions. In recent decades, the saying of "turn" has been popular in the philosophy circle, so that some people say that "the turn of philosophy is one of the hot topics in the current philosophy circle" (Huang, 2004). As long as we look up the literature, we will find that there are various kinds of philosophical turn, such as "ontology turn", "epistemology turn", "linguistics turn", "practice turn", "technology turn", "hermeneutics turn" and so on. Some people even joke that the various "turns" are confusing and misleading.

As a matter of fact, despite the hyped false "turns", most of the "turns" in philosophy in history are true, with difference in significance. While some turns were revolutions in the general direction, some were just changes in small direction. There are two major turns generally recognized by the philosophical circles, namely, the turn of modern philosophy from ontology to epistemology, and the turn of modern philosophy from epistemology to linguistics (Huang, 2004).

The so-called turn from ontology to epistemology happened with the rise of modern philosophy, which has a history of more than 300 years. The research of ancient Greek philosophy and medieval philosophy focuses on the composition and essence of the world. Although different philosophers have different answers, they all explored the ultimate structure of the world and the essential problem behind it, which is later called "ontology". Because of the backwardness of science and technology and the low cognitive ability at that time, philosophers basically relied on their rich conjecture ability to guess the ultimate composition and reasons behind the world. With the development of modern science and technology and the improvement of human's cognitive ability, philosophers (also scientists) represented by R. Descartes and F. Bacon began to focus on the study of human's cognitive ability. This is the most famous and the first major turn in the history of philosophy, namely "epistemological turn". After that, the focus of western philosophy research turned from ancient ontology to epistemology. Modern philosophy research basically focused on the cognitive ability of mankind, and developed along the two different paths of rationalism and em-

piricism.

With I. Kant's integration of rationalism and empiricism, and the establishment of G. W. Hegel's comprehensive philosophy system, the epistemological turn of philosophical research has been basically completed. However, compared with the science and technology which just started in the Renaissance, after more than 2000 years of philosophical ontology in ancient times and more than 300 years of epistemological exploration in modern times, philosophical research was still bothered with endless arguments, without reaching consensus. Therefore, a group of philosophers with scientific literacy began to reflect on the questions of philosophy research, who believed that although the ancient ontology and modern epistemology had made great contributions to philosophy, they did not grasp the essence of the question. Without proper cognitive tools, we can neither explore the essence of things nor reflect on the cognitive ability of human beings. At the turn of the 20th Century, classical logic developed rapidly under the impetus of F. Frege, B. Russell, and many other mathematicians and philosophers. Attributing the slow progress of philosophy to the lack of appropriate expression tools, these philosophers introduced the newly rising mathematical logic into the study of philosophy, holding that mathematical logic should be taken as a strict expression tool, and that the daily natural language used in philosophy should be replaced with modern symbolic language such as mathematical logic.

At the beginning of the 20th Century, philosophical research turned from cognitive subject to cognitive tool, and began to enter the age of analytical philosophy and the linguistic turn (Dummentt, 2018). Of course, this is the first linguistic turn, the turn from daily language to logical language, among which the most representative are the philosophers F. Frege, B. Russell, L. Wittgenstein and the Logical Empiricism. Logical empiricism is the most representative. They advocated the use of mathematical logic to replace the vague daily language expression of traditional metaphysics. They thought that all knowledge must be logically clear and could be restored to testable propositions. R. Carnap even put forward the idea of logical construction of the world, trying to reconstruct the world with logic. However, the ideal of trying to reconstruct the world with mathematical logic soon proved to be so difficult to achieve that logical language turned to daily language and began the logical analysis of daily language in the 1950s and 1960s. This was the second small turn of language turn, which brought prosperity to the study of philosophy of language (Wang, 2009). After analytical philosophy and linguistic turn, there were hermeneutic turn, practical turn, and technological turn in 20th Century philosophical research, but these turns were actually small turns in the context of linguistic turn. The general feature of Philosophy in the 20th Century is the age of analytical philosophy with mathematical logic as the tool, natural language as the object, and logical analysis as the method. Philosophy of language has become an important field of philosophical research in the 20th Century (Jiang, 2009).

In the 21st Century, as the spirit of the times, will philosophy continue the analytical philosophy of the 20th century and take philosophy of language as its main field? Is there a similar situation at the beginning of the 21st century? Has the revolution of science and technology brought different characteristics from the past? Does the 21st century have its own unique spirit of the times, so that philosophy as the spirit of the times must be reconstructed to reflect the characteristics of the times? We will answer these questions from the dilemma of philosophy and the change of the times.

From the perspective of the development of philosophy itself, both analytic philosophy and linguistic turn have encountered insurmountable difficulties. The original ideal of analyti-

cal philosophy is to let philosophy, like science, use more accurate mathematical logic, a symbolic logic language, to replace fuzzy and polysemous daily language, so that philosophical views can be expressed clearly. Therefore, the task of logical positivism is to clarify the meaning of philosophy. They believe that the reason for the rapid progress of natural science lies in its use of a specialized scientific language. Although philosophy has lofty ideals such as ultimate concern, it has no appropriate language tools to accurately express its thoughts, and can only use vague and polysemous daily language, which has caused more than 2000 years of disputes and is difficult to make progress. The rise of mathematical logic at the end of the 19th century and the beginning of the 20th century brought hope to metaphysics. However, after decades of exploration in the early 20th century, we found that it is not as simple as we thought to make metaphysical language logical and symbolic. Due to the difficulty of logicalization and symbolization, the study of philosophy turned to the logical analysis of everyday language from the 1950s, began to pay attention to everyday language, and formally put forward the philosophy of language. It is a pity that after decades of attempts, apart from fragmenting everyday language, the logical analysis of everyday language has not made much progress. Since the 1980s, the movement of analytical philosophy has gradually declined, and the philosophy of language has almost fallen into a dilemma. Therefore, analytical philosophy or philosophy of language can no longer lead the trend of Philosophy in the 21st century (Rescher, 1994). Philosophy in the 21st century should have its own soul and tools, and should turn to a new philosophical direction.

From the perspective of the changes of the times, the new information revolution has taken place in the 21st century, and human beings have entered the age of big data and intelligence. Therefore, philosophy as the spirit of the times must also turn with the changes of the times. Although the 21st century is only 20 years old, it has shown its own characteristics of the times. Data, big data and intelligence have brought great changes of the times. The process of human digitalization has a long history. The emergence of electronic technology and computer in the 1940s and 1950s accelerated the pace of digitalization. Therefore, in the 1970s, some people claimed that human beings had entered the information age, and in the 1990s, Negroponte claimed that human beings had begun the age of digital existence. However, the real arrival of the age of data is at the beginning of the 21st century. A series of information technologies, such as mobile Internet, intelligent technology, smart phone, data mining and so on, have brought a sharp rise in the amount of data. Therefore, 2012 has become the first year of the age of big data. "Big data marks the beginning of a major transformation", just as the telescope enabled us to comprehend the universe and the microscope allowed us to understand germs, the new techniques for collecting and analyzing huge bodies of data will help us make sense of our world in ways we are just starting to appreciate." (Schoenberg & Cukier, 2013, p.7) "Big data marks an important step in humankind's quest to quantify and understand the world. A preponderance of things that could never be measured, stored, analyzed, and shared before being digitized. Harnessing vast quantities of data rather than a small portion, and privileging more data of less exactitude, opens the door to new ways of understanding." (Schoenberg & Cukier, 2013, pp. 17-18) With the advent of the age of big data, human beings can not only use natural language and logical language to understand the world, but also data will become the common language of human and machine. In other words, data language will become a universal language for humans and computers in the 21st century. Data language will be more accurate than natural language, more flexible than logic language, and more importantly, it can be

understood and calculated by intelligent machines. Therefore, philosophical research will develop with the development of big data age and data science and technology and turn from daily language and logic language to data language, so as to promote the data turn of scientific cognition.

At the beginning of the 21st century, the revolution of data technology was in full swing, and data language has become a new language in the new age. Therefore, it has become an inevitable trend for philosophical research in the 21st century to turn to a new philosophy of data, which brings about a new movement of data turn. From the perspective of philosophy itself, analytic philosophy, as a symbol of the 20th century, has been declining slowly and unable to lead the development direction of Philosophy in the 21st century. From the perspective of the development of science and technology, the revolution of big data and artificial intelligence promotes the philosophy of the 21st century to carry out the turn of data-centered philosophy.

3   The Possibility of Data Turn

In the 21st century, scientific cognition will no longer be the analytical philosophy centered on natural language or logical language analysis, but should move towards the research of philosophy of data based on more accurate and convenient data language, so as to realize the data turn of scientific cognition. However, the necessary for data turn of scientific cognition is not enough to motivate a complete turn. To make the data turn into a reality, we should have a series of possible conditions, including scientific basis, technical conditions, social background and philosophical basis. Only with these four preconditions, data diversion will become possible.

3.1   The Scientific Basis of Data Turn

Since the turn of modern epistemology, epistemology has become the focus of philosophical research. Whether it is the turn of modern language or the turn of modern data, it is the continuation, deepening and development of modern epistemology. Scientific cognition is closely related to scientific development. The data turn of scientific cognition should be based on a solid foundation of scientific development. The development of digital logic, discrete mathematics, computer science, information science, computing science and data science has laid a solid scientific foundation for the data turn of philosophy.

Arithmetic is the oldest branch of mathematics, which focuses on measurement, data and its calculation. Therefore, although data science and philosophy of data are new trends, their origins are particularly old. Although paying attention to graph and function has become the focus of modern mathematics research, data and calculation have not been forgotten, especially with the rise of computer science and technology. Boolean algebra turns mathematical logic into data, while discrete mathematics turns mathematics from continuous quantity to discrete quantity, which provides logical and mathematical basis for the world's data. Computer science, especially data structure and database theory, provides scientific means for data representation, storage and processing. The rise of information science provides a relatively complete theory for the measurement, transmission, coding and decoding of information in the world, thus providing an opportunity for data science and philosophy to pay attention to information and data. Before the advent of the age of big data, computational science used data and computers to solve many large and complex computing problems, especially the nonlinear partial differential equations that are difficult to analyze and a large num-

ber of engineering problems that are difficult to obtain analytical equations, thus computational mathematics, computational mechanics, computational chemistry and so on were established. With the massive data brought by the Internet and intelligent terminals, data mining and processing has become an important task of scientific research, so the rise of data science has laid a solid scientific foundation for the data turn of scientific cognition and the rise of philosophy of data (Chao, 2020).

3.2   Technical Conditions for Data Turn

It seems that the philosophical turn has little to do with technology, but it actually involves the question of whether the turn can be realized successfully. One of the important reasons why the linguistic turn of scientific cognition in the 20th century was not successful in the end is that the logical analysis of language is difficult to realize mechanization and automation, and can only be limited in the use of mathematical logic analysis of language by hand, which affects the efficiency and thoroughness of logical analysis.

It has always been the dream of mankind to digitize the world and use more accurate data to represent and understand things. However, due to the limitations of technology, in the past, we could only have ideas but no way. Technically speaking, many new technologies in the 21st century promote the world's digitalization, including digital electronic technology, computer technology, network technology, cloud computing, data mining, artificial intelligence, etc. Digital electronic technology is the most basic technology to realize the digitalization of the world. Before digital electronic technology, data acquisition, storage, processing and transmission can only be realized by analog technology, which will lead to serious data distortion and difficult to achieve automation. Digital electronic technology opens up the way for data to be represented by two simple signals: 0 and 1, so that human beings can accurately express data with electronic signals and realize the electronic processing of data. Computer technology automates and programs data storage and computation. The rise of network technology makes data transmission more convenient, and makes data live and online. Cloud computing allows massive data to be quickly stored and parallel computing, and data mining technology allows massive data resources to be more fully utilized according to needs. Artificial intelligence technology enables data acquisition, storage, transmission, processing and utilization of all aspects to be automated, intelligent and machine learning, so that human beings can be liberated. In the 21st century, a series of technologies mentioned above have basically become mature, which has gradually opened the curtain of world data. The digitalization of the world enables the data representation and analysis of the world to be mechanized, automated and intelligent. Even the logical analysis of language will be transformed into the data analysis of language, which makes it possible for the data turn of scientific cognition to be realized in technology.

3.3   The Social Background of Data Turn

Material, energy and information are the three elements of the world. In the past, people's cognition and application mainly focused on material and energy. It was not until the middle of the 20th century that the problem of information was gradually paid attention to. Due to the development of digital electronics, integrated circuit, computer, artificial intelligence and other technologies, the problems of information measurement, representation, storage and transmission have been gradually solved. C. E. Shannon and N. Wiener's research on information theory has directly promoted human cognition and application of information

With the development of information technology, in the late 1970s, American scholar Daniel Bell put forward the concept of post-industrial society. In 1980, Alvin Toffler directly put forward the concept of information society and called it the third wave after the agricultural wave and industrial wave. In the 1990s, with the rise of the Internet, human beings quickly entered the network age and began to live in a digital environment, which is also the further development of the information society. In the 21st century, the pace of information technology began to accelerate. First of all, people began to make full use of the Internet for all kinds of communication, and online social networking has become a new way of communication. Secondly, people began to use the network to engage in various economic and social activities, and the network economy has become a new way of economic development. Finally, human beings began to use the network for all kinds of learning, work and life, network life has become a new way of human life. Especially in the last decade, the original fixed network has been gradually replaced by various mobile networks, and various communication devices have been replaced by various intelligent terminals. People have gradually entered a new age of mobile. With the development of various smart terminals, Internet and IOT, all kinds of data show explosive growth, so 2012 is called entering the age of big data. All kinds of intelligent devices are mining and learning big data, and gradually become more and more intelligent, so in recent years, human beings have announced the new age of artificial intelligence.

With the advent of the age of big data and artificial intelligence, all kinds of intelligent sensing devices are gradually increasing, and all things are gradually converted into data. The rise of 5G network accelerates the formation of the Internet of things, so all things are gradually digitized, and gradually form a data world reflecting the physical world. Our society has gradually become a data society. The digital society and digital survival provide a real space for data cognition, data life and data experience, so that people can really feel the arrival of the data age. The formation of data world and the coming of data society provide a solid social foundation for the data turn of scientific cognition.

3.4   The Philosophical Basis of Data Turn

Any change in philosophical research, or the emergence of any new philosophical thought, can find the source in the philosophical thought of the past. Just as the linguistic turn at the beginning of last century has been brewing for decades or even hundreds of years, the data turn of scientific cognition is not imaginary or produced out of thin air. We can find its ideological source from thousands of years of philosophy history. The data turn of scientific cognition can be traced back to the exploration of data by Pythagoras, the ancient Greek philosopher. The analytical philosophy and the linguistic turn of the last century have provided many inspirations for the data turn. The information philosophy of recent decades has directly demonstrated the data turn and the philosophy of data, The latest big philosophy of data research and the concept of philosophy of data directly give birth to data turn and philosophy of data.

People invented data very early and used it to describe things and their states. In the process of tracing the origin of the world, ancient Greek philosophers also regarded abstract data as one of the origins of the world, among which the most representative is Pythagoras, the philosopher and mathematician. He believes that all other things are ever-changing, and only "number" can remain unchanged. Therefore, he proposes that "number is the origin of all things". In addition, he also reveals the essence and law of things through the quantita-

tive relationship between things, such as studying music rhythm through the length and thickness of strings. After the Renaissance, Galileo, Newton and other scientists and philosophers directly opened up the experimental tradition and began to understand the laws of things through experimental data. Descartes and other philosophers directly promoted the epistemological turn of modern philosophy. Leibniz not only put forward the idea of binary, but also put forward the idea of solving philosophical disputes through calculation.

At the end of the 20th century and the beginning of the 21th century, philosophers began to realize that language and its correct expression are the premise of human cognition. Therefore, the problem of language and its logic must be solved first after the turn of epistemology. More importantly, because the rise of mathematical logic has possessed the logical structure tool of analyzing language, analytical philosophy and philosophy of language have sprung up. It also holds that scientific cognition should realize the linguistic turn. The linguistic turn of philosophy provides beneficial enlightenment for the data turn: first, the linguistic turn provides an example of how to make a further turn for the data turn. Second, it is necessary to make a further turn to realize the great wish of the epistemological turn.

With the advent of information theory and the rise of information revolution, scholars at home and abroad have carried out research on information philosophy. In essence, the philosophy of data should belong to the category of information philosophy. Therefore, the construction of information philosophy lays the foundation for philosophy of data and data turn. With the development of big data technology and the advent of the age of big data, people have carried out all-round research on the ontology, epistemology, methodology and ethics of big data, so people have a clearer understanding of the nature of data and the role of data in human cognition, which has made a direct contribution to the study of digital turn and philosophy of data.

With the advent of the new generation of information technology revolution and the age of big data, data has become the most valuable resource in our age. As a scientific representation of information, data has become more and more influential. As a result, people pay more and more attention to data, and regard it as a new object of philosophical research. Pythagoras's idea of "all things in life" has been gestated and developed for more than 2000 years, and finally has been echoed in the 21st century. In recent years, Bogen & Woodward (1988), Stanton & Bunker (2009), Bogen (2011), Woodward (2011), Brooks (2013), Teboul (2017), Liu and Hu (2012), Ye and Zhang (2015), Furner (2017), Huang (2021) have put forward the concept of philosophy of data, which has made the most direct contribution to the construction of philosophy of data and data turn.

4   Where does the Date Turn

With the change of the times and the development of science and technology, it is necessary and possible to make the data turn of scientific cognition. However, where will the data turn? In which fields or directions will scientific cognition turn to data? This requires us to continue to explore how data diversion is possible.

The linguistic turn in the last century is mainly manifested in the shift from the physical world to human language, from the speculative method to the analytical method, and from the formal logic to the mathematical logic. What are the main aspects of this data turn? To sum up, the data turn will take place in the following five dimensions: from the perspective of cognitive objects, from language to data, from the perspective of cognitive tools, from logic to algorithm, from the perspective of cognitive method, from analysis to synthesis,

from the perspective of cognitive goal, from proof to discovery, from the perspective of cognitive results, from causality to correlation.

4.1 Cognitive Object: From Language to Data

After philosophy has changed from ontology to epistemology, people's focus has shifted from what the world is to how we know the world. Therefore, our knowledge must first have the object of knowledge. Ultimately, the object of our understanding is of course the material world and the human spiritual world evolved from the material world. At the beginning of the epistemological turn, philosophers thought that people's cognition was the direct cognition of various phenomena in the world. However, it was not until the beginning of the 20th century that philosophers discovered that human cognition was not really directly oriented to the phenomenal world, but the linguistic world represented by linguistic symbols, which led to the turn of language and the rise of analytic philosophy and linguistic philosophy in the 20th century. The reason why we can realize the cognition of the world through the cognition of language is that language and the world form a kind of mapping relationship, that is, people's cognition of the world must be described by language. However, the fuzziness, arbitrariness and polysemy of natural language make the relationship between language and the world not one-to-one. Analytic philosophy tries to introduce symbolic mathematical logic to improve the limitations of natural language, but mathematical logic is difficult to realize the calculation and automation of analysis, so the attempts of language turn, analytic philosophy and philosophy of language have not fully realized their original aspirations.

With the development of data technology and perception technology, it is more accurate to describe the world with data than with language, that is, data and the world can achieve one-to-one mapping relationship. Therefore, data language is more accurate than natural language, and more convenient to calculate and model than logic language. Therefore, with the advent of the age of big data, the digitalization of all things enables all things in the world to be mapped into data, and the material world or the spiritual world can be transformed into the data world. Through the cognition of the data world, we can recognize the material world or the spiritual world more accurately, conveniently, intelligently and automatically. Therefore, scientific cognition in the 21st century is about to move from language analysis to data analysis, from philosophy of language to Philosophy of data, and from cognitive objects to realize the transformation from language to data.

4.2 Cognitive Tools: From Logic to Algorithm

With the change of cognitive objects from language to data, cognitive tools will also change accordingly, that is, from logic to algorithm. Historically, the results of cognition have been recorded and represented in the form of language, and then the rules contained in the language world have been discovered. How to find rules in language? Aristotle's formal logic is the best tool for language analysis, but it lacks symbolic expression and calculus, so it is difficult to analyze all languages logically. Because of this, when mathematical logic appeared in the early 20th century, philosophers eagerly introduced it into language analysis, and even tried to construct a purely symbolic language of mathematical logic. However, neither the logical analysis of natural language nor the construction of symbolic mathematical logic language can be realized in practice or technology, resulting in the turn of language and the frustration of philosophy of language and analytical philosophy.

Since the Renaissance, scientists did scientific exploration by means of data collection, data

modeling, data calculation and data verification. Data is the most important tool to depict world phenomenon and essence. That is to say, data is a more suitable scientific language than natural language and logical language. With the advent of the age of big data, everything can be mapped to data. Big data uses bits to represent things and their characteristics, transforms everything into data, and forms a data world outside the natural world and language world. For the data world, we can't use formal logic or mathematical logic to analyze, we can only use calculation and algorithm to find the rules contained in the data. Therefore, from the perspective of cognitive tools, logical analysis after language turn will give way to data algorithm and calculation after data turn, that is, from logical analysis to algorithmic calculation.

4.3 Cognitive Approach: From Analysis to Synthesis

Scientific cognition has its own cognitive method system. Different times and different objects must use different cognitive methods. In essence, there are two basic methods of scientific cognition: analysis and synthesis. The so-called analytical method is to dissect and decompose a whole thing, turn the whole into its parts, that is, decompose it into its constituent elements, and then study the nature and mechanism of the constituent elements to achieve the purpose of understanding things. The so-called synthesis is to assemble and integrate the elements of things into an organic whole, and understand things through the nature or function of the whole.

In the process of linguistic turn, the basic cognitive method of analytic philosophy and linguistic philosophy is analysis. Analytic philosophy can know the position of "analysis" in methodology from its name. Philosophy of language also takes language as the object and analysis as the means, that is, to analyze the logical relationship between the meaning of words, concepts and propositions, so as to realize the understanding of concepts and propositions.

After entering the 21st century, big data technology has fragmented everything into data encoded by 0 and 1, that is to say, big data has turned the world into data fragments. To find rules and discover knowledge from these data fragments, we have to re integrate and reconstruct these data fragments, and realize the comprehensive integration of data through data cleaning, classification, association, aggregation and other processes. Therefore, in the age of big data, we mainly find the correlation between data through comprehensive methods and discover new knowledge through comprehensive integration. Although it is a typical process of analysis and decomposition to turn the world into data fragments, the comprehensive method will be the main cognitive method after the shift. An important sign of the shift from language to data is that the cognitive method is mainly from analysis to synthesis.

4.4 Cognitive Goal: From Proof to Discovery

There are two basic goals of scientific cognition. One is to discover regular knowledge from complicated phenomena, the other is to test and prove the discovered knowledge in order to prove its reliability and truth. This is the so-called problem of scientific discovery and scientific proof. Scientific discovery enables us to acquire new knowledge, while scientific proof makes new knowledge reliable.

Linguistic turn mainly relies on mathematical logic and its causal inevitability, so the emphasis of analytic philosophy and linguistic philosophy is to prove the reliability of existing knowledge through logical relations. For example, logical positivism, an important school of

the analytical movement, tries to prove the logicality and reliability of the words in the proposition through the logical analysis of the proposition, so as to completely eliminate the logical contradictions in the proposition. On the other hand, it tries to test the truth of propositions through empirical facts. Philosophy of language, on the other hand, eliminates all kinds of logical contradictions in language propositions through the logical relations between concepts and propositions, so that language propositions can accurately represent their correct meanings.

Both analytical philosophy and philosophy of language ignore the important issue of scientific discovery. They avoid the important issue of scientific discovery on the pretext that scientific discovery has no logical inevitability. Popper, the philosopher of science, put forward that "induction does not exist" and completely rejected induction, thus rejecting the problem of scientific discovery, and replaced the process of scientific discovery with bold hypothesis and bold trial and error. With the advent of the age of big data, we see the hope of exploring scientific discoveries. Big data technology finds the correlation between data by mining massive data and by means of data clarity, classification, association and aggregation, so as to discover the hidden regularity of things. Although there is no logical inevitability in the correlation between data, it can provide the direction of guessing the answer and play the role of "helping to find". The data mining of big data is to discover the possible rules contained in massive data by means of massive data and algorithm calculation, and then prove its reliability by means of knowledge proof. Data mining is also known as knowledge discovery because it may discover new knowledge in massive data.

4.5   Cognitive Outcome: From Causality to Relevance

The results of scientific cognition are basically expressed in the form of theory and law. From the traditional philosophy of science, any theory or law must have causal inevitability, that is to say, the result of scientific cognition is a series of laws or theories with causal inevitability. However, in practice, a lot of empirical knowledge is not necessarily casual, only relevant. The so-called relevance is that one event or phenomenon and another event or phenomenon have the following and related relationship. The appearance of one event may bring about the appearance of another event.

The movement of analytic philosophy in the 20th century is to build a knowledge system with strict causality. This kind of knowledge system based on strict causality existed in ancient Greece. Euclidean geometry was an example, and modern Newton's mechanics and Spinoza's ethics were examples. In the early 20th century, B. Russell and A. Whitehead's the principles of mathematics and L. Wittgenstein's Tractatus Logico-Philosophicus carried forward this spirit, it has become the ultimate pursuit of analytical philosophy and linguistic philosophy. "After all, as humans, we desire to make sense of the world through causal links; we want to believe that every effect has a cause, if we only look closely enough." (Schoenberg & Cukier, 2013, p.63)

The advent of the big data age has broken the traditional scientific cognition that takes causality as the ultimate goal. Big data emphasizes experience, phenomenon and practice, and believes that "Knowing what, not why, is good enough." (Schoenberg & Cukier, 2013, p. 52) it is more important to find the correlation law through the correlation presented by the data itself than through the causality law based on hypothesis, "In place of the hypothesis-driven approach, we can use a data-driven one. Our results may be less biased and more accurate, and we will almost certainly get them much faster." (Schoenberg & Cukier, 2013, p. 55) Therefore, after the shift of scientific cognition data, our scientific cognition will shift

from the pursuit of strict causality between things to the pursuit of correlation between data, "These non-causal analyses will aid our understanding of the world by primarily asking what rather than why." (Schoenberg & Cukier, 2013, p.63) From the pursuit of strict causality to the pursuit of correlation between data is an important direction of scientific cognitive data turn in the results of scientific cognition.

5   How to Construct Philosophy of Data

The age of data calls for the data turn of scientific cognition and the construction of philosophy of data. But how to construct philosophy of data? What is the research object of philosophy of data? What is its connotation and relationship with other disciplines? What is its basic research framework? These are the basic issues that we need to care about in the construction of philosophy of data.

5.1   Research Object of Philosophy of Data

Of course, the research object of philosophy of data is data, but why didn't data play a leading role until the 21st century? Why will data play a leading role in the 21st century? This is because different times have different spirits of the times, so the research focus of philosophy is not the same, and its research object will be different.

According to materialism, the world is made up of matter, but matter itself contains three elements: material, energy and information, which are related to the problems of time and space, power and cognition of matter. The most important concern of ancient Greek philosophy was the problem of material. Modern philosophy turned to focus on power (energy), that is, on the subject, focusing on epistemology. At the beginning of the 20th century, people began to focus on information and cognition. But at the beginning of the 20th century, the cognition involved mainly remained in the external form of information, that is, the problem of language and logic, so language turn, philosophy of language and analytical philosophy became the focus of Philosophy in the 20th century.

But with the development of science and technology, people's cognition of information has made great progress, and began to use more accurate and scientific digital form to represent information, that is, people pay more attention to the internal forms and content of information. So data has become the common concern of science, technology and philosophy in the 21st century. More importantly, natural language is only the language between human beings, while data language is the common language between human beings and machines (even everything). That is to say, with data language, communication between human beings and machines (even everything) can be realized. Therefore, after the turn of scientific cognitive data in the 21st century, philosophy of data takes data as the research object to study a series of problems related to data, such as the nature of data, the relationship between data and the world, the nature of algorithms and so on.

5.2   Two Meanings of Philosophy of Data

Just as philosophy of science has two meanings, philosophy of data also has two meanings and two ways of understanding. In fact, this involves the two meanings of philosophy of data.

The first meaning focuses on the philosophical issues contained in the data. Data is the common language of humans and machines, which is more accurate than human natural language and more flexible than logic language. As a kind of accurate information mapping

of nature, society or thinking, data depicts the phenomenon and state of the world and forms the data world corresponding to the real world. In this world composed of data, there are profound philosophical problems. Therefore, the first meaning of philosophy of data is to use philosophical vision and methods to mine philosophical problems contained in the data world. Just as we need to study the philosophical problems in other sciences and technology, it is also necessary to study the philosophical problems in the data world in the 21st century.

The second meaning focuses on the problem of digitalization and data turn of philosophy. A series of unique ideas and methods in data science and data technology, such as digitization, data mining, data cognition, data modeling, algorithm, calculation and so on, will have a great impact on philosophy, especially cognitive philosophy. Positivism, logical positivism, falsificationism and analytical philosophy all criticize traditional philosophy for it self-referential and lack of experience, analysis, positivism and falsification. Therefore, these philosophical schools reject metaphysics and try to make philosophy learn from science. The so-called language turn and analysis movement means that philosophy should use the logic language of science like science, but the essence of scientific method is "speak with data", and data language is the real scientific language. Therefore, the rise of data science and data technology will have a great impact on philosophy, especially cognitive philosophy, which is the second meaning of philosophy of data.

5.3 The Subject Orientation of Philosophy of Data

From the two meanings of philosophy of data, we can see the narrow and broad disciplinary positioning of philosophy of data.

The so-called narrow sense of subject orientation is to regard philosophy of data as a branch or a field of philosophy of science. An important research content of philosophy of science and technology is the philosophical problems in science and technology, that is, to explore the philosophical problems contained in various science and technology from the perspective of philosophy. The corresponding philosophy of data is to explore the philosophical problems contained in data science and data technology from the perspective of philosophy.

The so-called generalized philosophy of data is the philosophical thinking of the relationship between data and reality. It is not only concerned with the ontological data view of what data is, but more importantly, how data relates to the objective world or reality on which we live. In the past, data was only used as a tool to describe phenomena, and the discussion of data was only an accompanying interest or a means to discuss other issues. Philosophy of data studies the data problem as the central problem of understanding and interpreting the world, or philosophy of data originates from the relationship between data and the world. Philosophy of data in a broad sense no longer stays in the orientation of department philosophy, but in the orientation of data turn in philosophy research, which is a change in the general direction of philosophy research. Philosophy of data in a broad sense is similar to analytic philosophy and linguistic philosophy in the 20th century. It is a great change of Philosophy (especially cognitive philosophy) in the 21st century and a philosophy of data of scientific cognition.

5.4 The Problem Space of Philosophy of Data

Philosophy of data is the philosophical generalization and the highest level of data science

system from the perspective of philosophy. Therefore, the philosophy of data is related to data science, data technology and data application, Stanton and Bunker (2009) think that philosophy of data is "a multi-disciplinary problem space": information philosophy, information system, information science and technology, semiotics, philosophy of science, philosophy of technology, information theory and so on all contribute to the foundation of philosophy of data from their own unique perspective.

Why does the construction of philosophy of data involve so many subjects? Stanton and Bunker made a comprehensive analysis. Semiotics studies words and symbols as data. Information philosophy discusses the polysemous relationship between data and information. Information systems, information science and technology explore the formal operation of data, such as abstract symbol sets, relational modeling and functional data modeling. Semiotics is an important part of semiotic philosophy and its interpretation. It formalizes some symbols into facts and becomes the field of philosophy of science, which attempts to reveal the inherent nature of facts. At the same time, philosophy of technology provides ethics and value systems for data, and also provides tools for us to explore the basic philosophical essence of data. Information theory explores the essence of data mechanical transmission and imposes physical constraints on our constructed reality. In addition, pedagogy, psychology and biology are exploring the neural basis of our brain's understanding of data. Since the discipline of data science and data technology was not as prominent in 2009 as it is today, it seems that Stanton and bunker did not particularly emphasize it. In fact, they have become the most important disciplinary basis for the construction of philosophy of data.

5.5 Research Framework of Philosophy of Data

Every subject has its own terminology. If the key words of analytic philosophy and linguistic philosophy are a series of words related to language and logic, such as language, logic, reference, symbol, proposition, meaning and reality, then the key words of philosophy of data become a series of new words related to data, such as information, data, mining, representation, algorithm, calculation, simulation and reality.

Philosophy of data, in short, is a new discipline that takes data as the object, algorithm as the tool, and synthesis as the method to conduct philosophical research on the data world mapped by the objective world. Analytic philosophy and philosophy of language mainly focus on the problem of scientific proof, but they don't care much about the problem of scientific discovery, that is to say, they only care about how to express and deduce the knowledge that has been discovered more in line with the logical rules. Philosophy of data pays more attention to scientific discovery, which focuses on knowledge discovery. That is to say, the research focus of philosophy of data is how to transform the phenomenon world into the data world, and then find the correlation contained in the data world through data mining methods, so as to summarize, synthesize and extract the algorithm and model of data, and explain the past and predict the future through algorithm calculation.

The main issues in philosophy of data are: data and world, data and structure, data and algorithm, data and model, data and calculation, data and language, data and knowledge, data and truth, data and ethics. If we classify these problems, we can find that they involve a series of problems such as ontology, epistemology, methodology, axiology, ethics and sociology.

Data Ontology: mainly discusses the nature of data, data and phenomenon, data and reality, data and reference, the relationship between data and the world, the data structure of the

world, data phenomenology and other issues.

Data Epistemology: This paper mainly discusses a series of data cognitive activities based on Boolean algebra, such as logical structure of data language, data mining, data description, data analysis, data modeling, data calculation, data simulation, data creation (Wang, 2021), especially the problem of data intensive scientific discovery, and finally explores the relationship between data and truth.

Data methodology: mainly discusses data-driven methodology, especially the popular data-intensive scientific methods and computational scientific methods in recent years. Data intensive methods use massive data to describe problems and use data to find and solve them. In the past, computational science methods were mainly used for complex problems that were difficult to solve analytically, such as computational mathematics, computational physics, computational chemistry, etc. In recent years, with the rise of big data technology, computational science methods have been widely used in Humanities and Social Sciences. In particular, computational social sciences have derived from computational communication, computational sociology, computational politics, computational ideological and political education and so on.

Data axiology: mainly discusses the value of data. As a new resource, data will derive various values. It is necessary to discuss the value of data from the perspective of axiology.

Data Ethics: mainly discusses ethical issues caused by data opening and sharing. After the big data revolution, the ethical and moral issues caused by data have attracted great attention, so this issue is currently the hottest topic and has been studied widely.

Data society theory: data is closely related to society, which has been accompanied by the progress of human society. The changes of data collection, storage, transmission, calculation and utilization have caused great social changes in history. Therefore, data is closely related to a series of social phenomena such as social economy, politics, culture, thinking, life and work. Therefore, the philosophy of data must pay attention to data and social problems.

With the development of the new generation of information technology revolution and the further opening of the curtain of the age of big data, the content of philosophy of data research may be more colorful, and many problems we may not have thought of at present.

6 Conclusion

At the end of the 19th century and the beginning of the 20th century, the rise of mathematical logic brought about a linguistic turn of scientific cognition and a revolution of analytical philosophy in the 20th century. Since then, analytical philosophy and linguistic philosophy have occupied a dominant position in British and American philosophy for nearly a century. At the end of the 20th century and the beginning of the 21st century, the rise of data science and technology will also bring about a data turn and philosophical revolution of scientific cognition in the 21st century. Big data technology restores all things in the world in the form of bits of data fragments composed of 0 and 1, while data mining and machine learning reintegrate data fragments. Data mining and data algorithm are used to discover rules, and data calculation is used to predict the future. Therefore, in the 21st century, there will be a new revolution in the philosophy of scientific cognition. This revolution will take data as the basis, algorithm as the tool, and synthesis as the method. New knowledge will be discovered from data fragments with synthesis method and algorithm tool, and human beings will enter a new age of knowledge discovery, and gradually build the philosophy of data for new age.

Reference

Bogen, J., & Woodward, J. (1988) . Saving the phenomena. *The Philosophical Review, 97* (3) , 303–352.

Bogen, J. (2011) . Saving the phenomena´ and saving the phenomena. *Synthese, 182* (1) , 7–22.

Brooks, D. (2013, February 5) . The philosophy of data. *New York Times*. https://www.nytimes.com/2013/02/05/opinion/brooks–the–philosophy–of–data.html

Chao, L. (2020) . *Theory and practice of data science.* Beijing: Tsinghua University Press.

Dummentt, M. (2018) . Origins of analytical philosophy. In B. Chen & Y. Jiang (Eds.) , *Analytical philosophy: Review and reflection (I)* (pp.4–14) . Beijing: China Renmin University Press.

Furner, J. (2017) . Philosophy of data: Why?. *Education for Information, 33* (1) , 55–70.

Huang, S. (2004) . On the turn of western philosophy. *Zhejiang Academic Journal, 2004* (1) , 69–77. (In Chinese) .

Huang, X. (2021) . From linguistic turn to data turn: New changes in scientific cognition. *Theoretical Exploration, 2021* (02) , 32–39. (In Chinese)

Jiang, Y. (2009) . Course of analytical philosophy. Beijing: Peking University Press.

Leonelli, S. (2015) . The philosophy of data. In L. Floridi (ed) . *The Handbook for the Philosophy of Information.* Routledge.

Liu, H., & Hu, X. (2012) . Preliminary exploration on the construction of philosophy of data. *Philosophy Trends, 2012* (12) , 82–88.

Rescher, N. (1994) . *American philosophy today, and other philosophical studies.* Savage, MD : Rowman & Littlefield.

Schoenberg, V. M., & Cukier, K. (2013) . *Big data: A revolution that will transform how we live, work and think.* John Murray.

Stanton, B. B., & Bunker, D. (2009) . Philosophy of data (PoD) and its importance to the discipline of information systems. In *15th Americas Conference on Information Systems.* AMCIS–0193–2009. R1. https://www.researchgate.net/publication/220892532

Teboul, B. (2017, July 27) . A philosophy of data. 163. Retrieved May 17, 2021, from https://www.163.com/dy/article/DJ59BAJP0511FBRQ.html

Wang, L. (2009) . *Into analytical philosophy.* Beijing: China Renmin University Press.

Wang, T. (2021) . Big data and creative epistemology. *Journal of Shanghai University, 2021* (1) , 105–113. (In Chinese) .

White, M. (1955) . *The age of analysis: Twentieth century philosophers.* New American Library.

Woodward, J. F. (2011) . Data and phenomena: Arestatement and defense. *Synthese, 182* (1) , 165–179.

Ye, S., & Zhang, Y. (2015) . Data, phenomenon and theory: A review of Bergen and Woodward´s philosophy of data. *Dialectics of Nature, 2015* (6) , 106–111.