CHAPTER 14

BIG DATA TECHNIQUES AND RESEARCH METHODS IN POLITICAL SCIENCE

Rasheed Obasekore, GARUBA and
Olayinka Moshood, SALAHU

[1]Fountain University, Nigeria
[2]Kwara State University, Nigeria

Chapter Overview

Contemporary studies in Political Science are becoming a data field where fundamental questions in research are now answered by data scientists operating on large datasets. This has made the traditional approach to advance inquiry in political science face the threat of becoming marginalized for new methods of research where the issue of computational skills gains an additional advantage because of empirical data accessibility for research work. According to Niels (2016), behaviouralism in political science brings a complete idea of the behavior of political actors as a significant factor that has a greater influence on political events and institutions. With the big data model in political science research, the burdens of interdisciplinary collaboration with other social science disciplines have been demystified.

Sequel to the above, this chapter has come up with conceptual foundations for the use of big data in policy research toward advancing scholarship in political science. This comes as a result of a lack of excitement from many political scientists to take advantage of the rise of big data to advance reasoning in the subject matter. As argued by Niels (2016), big data offered sufficient information about context, processes, and interactions among variables and actors. Therefore, this chapter examines the applicability of big data techniques to research methods in Political Science in line with under listed learning objectives: understand the

Big data plays a si
nificant role in advan
ing political science r
search by providi
new tools, metho
and insights that e
hance our understan
ing of political dyna
ics, electoral beha
iours, and politic
trends. It allows r
searchers to analy
vast amounts of diver
data sources to ga
deeper insights into p
litical processes, pub
opinions and poli
outcomes.

methodology of gathering data in large quantities from online sources for empirical-based political science research; comprehend statistical analysis to gather and interpret survey data on political phenomena; identify appropriate techniques for proffering empirical solutions to challenges impeding governance and democratic sustainability in the state; appreciate the significant need for data scientists in the government and public sector; provide solutions to challenges of computational methods in political science research.

Introduction

Big data is a new technique of accessing information for systematic analysis of data in contemporary research work. It encompasses the volume of data in large quantities that are difficult to process using the traditional method. It is means of gathering information in larger volumes and higher velocity for research work. A researcher can utilize findings from big data to make an empirical-based research output. According to Segal (2022), big data is a combination of structured, semi-structured, and unstructured data collected by the researcher and can be mined for information purposes and useful in learning projects, predictive modeling, and another form of analytics applications.

Segal (2022) characterizes big data with three V's comprising *volume; variety* and *velocity* at which data could be generated, collated, and processed for empirical research. It can also be generated through shared comments on social media networks or from websites or be voluntarily gathered from electronic apps, online questions generated through Google Maps, or electronic check-ins. The data generated through big data techniques could be stored in a computer database and analyzed using software designed to handle large and complex data sets for information processes.

In political science research, the application of big data has become a major trend for behavioural science research. As argued by Niels (2016), behaviouralism in political science brings a complete idea of the political behavior of actors as a major factor and condition that has a greater influence on political events, and political institutions. With the big data model in political science research, the burdens of inter-disciplinary collaboration with other social science disciplines when the study of politics solely relied on the traditional model of inquiry have been demystified.

Anifowose (1999) had earlier seen political science as an art that limited itself to the traditional model of inquiry due to its emphasis on the adoption of common sense, intuition, imagination, and experiences. It is a study that promotes value-based research orientation and its adherents believed that the study of the subject-matter should not be based on a scientific approach because facts alone

cannot justify the cause of political phenomenon within the political system. This justifies the position of classical political philosophers that normative orientations rather than the behavioral approach where the big data technique is situated could not be enough to derive facts in political analysis.

As argued by Gauba (2010), the behavioral approach to the study of politics is in line with the systematic investigation because the behavior of several actors in a political situation is capable of scientific study. Researchers in political science could generate data in large quantities and make it coherent for the logical conclusion; this could make research in political science derive hypothetical relationships among political forces. They can be investigated by the researchers. Therefore, in political research of recent times, the big data technique has gained considerable traction to analyze causal relationships, make useful classifications, and forecasting. It is the latest strategy adopted to study political phenomena and criteria employed in selecting questions for the target audience.

Big Data: Meanings and Definitions

Big data as a new technique for research work started in the mid-70s with the formation of relational databases in Europe. But in 2005, the growth in the use of social media and other online services made the use of big data for social and management research a major innovation in research work. As noted by Monogan (2015), the big data orientation has become a form of facts and statistics collected for referencing. It is huge in volume and growing exponentially with time. It is a type of data generation platform with so large size and complexity that none of the traditional data management tools can store or process efficiently.

Monogan (2015) posits that big data has intrinsic value to provide information about the process that requires insightful analysis, recognizing patterns, making informed assumptions of the phenomenon, and predicting the behavior of indi-

viduals, groups, and organizations. In terms of growth, big data makes it possible for researchers to gain more complete answers to many issues because of the availability of useful information in a larger quantity.

As noted by Taylor (2022), big data statistic indicates that *500+terabytes* of new data get ingested into the databases of the social media site Facebook every day and it is being generated in terms of message exchanges, comments, photos and video upload that contains greater variety, arriving in increasing volumes with more velocity. This is also known as the three Vs (Volume, Velocity, and Variety) where it can generate *10+ terabytes* of data in 30 minutes of flight time. With thousands of flights per day, the generation of data can reach up to many petabytes and this is more useful for data generation for research work.

Taylor (2022) avers that big data is a more complex data sets and so voluminous beyond the traditional data processing software package. The *first V is the Volume* of the three Vs showing how data can be processed in high volumes with low density where the unstructured data could be generated. It can also be processed through sensor-enabled equipment and might generate many terabytes of data in its processes. The second ***V is*** *Velocity:* this is the fast rate at which data is received and (perhaps) acted on. Normally, the highest velocity of data streams directly into the written disk. *The Variety* as the last V has to do with different types of data that are available for processes. Its traditional types are structured and fit neatly in a rational database system.

Conclusively, the trend of big data for research work has come far and it provides the opportunity for a researcher to analyze potential issues before the problems happen. The political system generates information in large quantities and at random. Political science researcher needs an analytical approach to making a coherent choice of data and arranging them logically for smooth research work.

involves the observation of research subjects. In some subjects that do not physically carry out experimentation, the originality of the research is in a particular way to aid understanding of the outcome of the research. Therefore, research is a process through which the researcher searches for definite or useful information from the number of respondents to evaluate or solve the problem-related questions under investigation.

Research is the art of empirical investigation that could generate a systematic way of gathering data and harnessing curiosity. This research provides scientific information and theories for the explanation of the nature and the properties of the phenomenon. It makes practical applications possible and can be subdivided into different classifications according to their academic and application disciplines. Scientific research is a widely used criterion for judging social phenomena.

Nature of Political Science Research

The post-World War II era ushered in a behavioral revolution that subjected political phenomena to scientific analysis. According to Eneanya (2012), the objective of political science research is to propound theories that can describe the phenomenon, explain, analyze, and predict the effect of political phenomenon. It is a modern method of political research that political thinkers of the later stage felt the necessity to study from a new perspective. The behavioral approach to research in political science aims to minimize deficiencies of the traditional model of research in political science and is characterized by the following:

i. The behavioral model tries to conclude empirical data.

ii. It goes beyond the study of political structures and their historical analysis.

iii. The model believes in inter-disciplinary study.

involves the observation of research subjects. In some subjects that do not typically carry out experimentation, the originality of the research is in a particular way to aid understanding of the outcome of the research. Therefore, research is a process through which the researcher searches for definite or useful information from the number of respondents to evaluate or solve the problem-related questions under investigation.

Research is the art of empirical investigation that could generate a systematic way of gathering data and harnessing curiosity. This research provides scientific information and theories for the explanation of the nature and the properties of the phenomenon. It makes practical applications possible and can be subdivided into different classifications according to their academic and application disciplines. Scientific research is a widely used criterion for judging social phenomena.

Nature of Political Science Research

The post-World War II era ushered in a behavioral revolution that subjected political phenomena to scientific analysis. According to Eneanya (2012), the objective of political science research is to propound theories that can describe the phenomenon, explain, analyze, and predict the effect of political phenomenon. It is a modern method of political research that political thinkers of the later stage felt the necessity to study from a new perspective. The behavioral approach to research in political science aims to minimize deficiencies of the traditional model of research in political science and is characterized by the following:

i. The behavioral model tries to conclude empirical data.

ii. It goes beyond the study of political structures and their historical analysis.

iii. The model believes in inter-disciplinary study.

iv. It emphasizes scientific methods of study and attempts to draw scientific conclusions in political science.

As earlier mentioned, research is a process to discover new knowledge. It is a systematic investigation designed to develop or contribute to generalizable knowledge. In political science research, the hypothesis is a foundation of research. It is an informed prediction or explanation of the phenomenon under investigation. Part of the research process in political science involves the generation of data, testing of hypothesis, and examining results as they relate to the hypothesis and political phenomenon. Political science research reveals the relationships underlying political events; it attempts to construct general principles about the way the world of politics works. Given the above template, research in political science needs to be conceptually precise, free from bias, and well-substantiated by empirical evidence with precise and persuasive theories.

As argued by Groh (2018), in conducting value-free research in political science, researchers must strive to emulate the objectivity as well as the conceptual and methodological rigor typically associated with the so-called "hard" sciences (e.g., biology, chemistry, and physics). Researcher in political science needs to adhere to a simple model of scientific inquiry because the key to building precise and persuasive theories in political science is to develop and test hypotheses. Hypotheses are statements that researchers construct to test whether or not a certain relationship exists between two phenomena.

When a researcher forms a hypothesis, it acts as a map through the research study and also tells the researcher which factors are important to study and how they might be related to each other or caused by a manipulation that the researcher introduces. With this map, the researcher can interpret the data generated and can make sound conclusions about the results. No matter what topic is being studied, Groh (2018) notes that the value of research in political science depends on how well it is designed and done. One of the most important considerations in doing good research in political science is to follow the design or plan that is developed by an experienced researcher who is called the Principal Investigator (PI).

The PI is in charge of all aspects of research and creates what is called a protocol (the research plan) that a researcher needs to be followed in the cause of the investigation. By doing so, the PI and the public can be sure that the results of the research are real and useful to other scientists. In a nutshell, research in political science serves as the cornerstone of theory-building in the subject-matter

Types of Political Science Research

Exploratory Research: Exploratory research is designed to investigate a problem that is not clearly defined. It is the focus on the discovery of ideas as opposed to collecting statistically accurate data. It is most commonly used for further defining political issues, areas for potential growth, alternative courses of action, and prioritizing areas that require statistical research. Under this method. surveys or polls can be used to gather large amounts of data, usually from a predetermined group of respondents. It is one of the most popular quantitative research methods. Surveys or polls are used in exploratory research to explore the opinions, trends, or beliefs of the target population.

Descriptive Research: Descriptive research is considered to be conclusive due to its quantitative nature. Unlike exploratory research, descriptive research is preplanned and structured in design so the information collected can be statistically inferred from a population. The main idea behind using this type of research is to better define an opinion, attitude, or behavior held by a group of people on a given subject. Consider the researcher's everyday multiple-choice questions. Since there are predefined categories a respondent must choose from. it is considered descriptive research. These questions will not give unique insights into the issues like exploratory research would. Instead, grouping the responses into predetermined choices will provide statistically inferable data.

Causal Research: Causal research like descriptive research is quantitative in nature as well as preplanned and structured in design. For this reason, it is also considered conclusive research. Causal research differs in its attempt to explain

the cause and effect of the relationship between variables. This is opposed to the observational style of descriptive research because it attempts to decipher whether a relationship is causal through experimentation.

Big Data and Research in Political Science: An Intersection

The issue of data relevance to political science research has become a new trend with a behavioral approach to the study of politics. The approach identifies the behavior of individuals or groups of individuals as the primary unit of analysis". It "seeks to examine the behavior, actions, and acts of individuals, rather than characteristics of institutions such as the legislature, executive and judiciary". As argued by Anastasopoulos and Poole (2019), the behavioral approach sees the application of big data techniques to research in political science as an avenue for the adoption of micro-targeting techniques for political campaigns by tracking down the known behaviors of an observed group and standardizes the model of generating the needed information for the research work.

In academics, big data has made it possible the adoption of *machine learning tools*, *clustering algorithms, and Bayesian models* for political science research. The concept of big data also pertains to accumulating, combining, analyzing, and application of large-scale data for various purposes in political science inquiry. The application of big data techniques is also useful in governance by enabling the administrator of the public establishment to make a strategic decision based on evidence and insights thought related to a large number of employees in the organization (Anastasopoulos and Poole, 2019).

Indeed, big data applications to political science have enabled the researcher to regard the techniques as a growth engine for the solution to socio-political and economic problems confronting a political system. As averred by Christenson (2018), big data research can support the integration and correlation of variables in large volumes of diverse datasets. The science of Visual Analytics can provide interactive visualization of the data to support evidence-based policy. The establishment of data-driven policies using big data can be useful in public administration and it has become an innovation for research work that can easily

generate information about the operational efficiency of employees in a larger percentage.

Christenson (2018) avers that big data techniques have opened up avenues for researchers to glean insights into the modern way of research conduct and gained considerable recognition to analyze causal relationships and make useful classifications possible in political science analysis. The relevance of big data to contemporary political analysis is ubiquitous. It pertains to a myriad of political, social, and economic activities that can arrive with great frequency, and exist in vast quantities with the ability of the researcher to collect and analyze this data like never before. This combination of big data and access to it holds great promise for learning about politics.

In considering what it means for the mobile phone, email, blogs, and social media can be and are recorded for analysis. That the thoughts and reactions of unprecedentedly large swaths of the public to political and economic events are available like never before and this makes the ability to understand public opinion more beneficial to the political actors. Likewise, seeing how individuals in a society are networked may lead to a richer understanding of the diffusion of (mis) information and motivations for political participation.

In short, big data has the potential to provide some opportunities for political scientists to improve their measures of political events, actors, and activities. The increasing presence of big data and the ability to collect has led to a greater understanding of politics (Christenson, 2018). Using big data, political scientists can engage with some of the biggest questions in a variety of social science disciplines and it improved the speed of capturing data in larger quantities for political analysis.

How Big Data Advances Political Science Research

According to Niels (2016), big data advances political science research in the following ways:

Measuring Political Ideology: Through computational text analysis, big data can be applied to large sets of speeches to measure the political ideology of individuals, groups, and society. This can be achieved by using exploratory factor analysis. Big data will enable political scientists to rely on vote records to estimate ideal points in the cause of measuring the ideology of a political party. Debates can also be made through big data information gathering. This can yield more significant textual data than how it can be used to estimate ideological positions. The digitization efforts of some parliaments in the United Kingdom and the United States of America have made huge sets of speech data available to researchers. With existing algorithms, millions of speeches can be scaled in a matter of hours.

Measuring Political Sentiment: Political scientists use this approach by gauging how people feel about a particular phenomenon based on tweets; a researcher can similarly use tweets to measure the peoples' opinions and sentiments toward some candidates contesting for election. Political scientists can also analyze data gathered through Twitter and other social media cheaply with Supervised and Aggregated Sentiment Analysis (SASA). A large subset of the texts (in this case, tweets) are analyzed by human coders and scored based on the sentiment they convey. Usually, coding distinguishes between positive and negative sentiments, but a more complex scheme can be used as well. Then, this data is fed into an algorithm that "learns" what a positive is and what is a negative text, and finally, that algorithm is applied to the whole dataset.

Identification of Instrumental Variables: When political scientists cannot measure the phenomenon of interest directly, they sometimes use a proxy (or "instrument") that is closely correlated with the variable of interest. According to Clark and Golder (2015), "Big data can help to the extent that it makes previously unobservable variables observable, thereby reducing the need for an instrument, or by making new potential instruments available".

Big Data Allows Political Scientists to Scale Research up Effectively: Big data techniques have made it possible for political scientists to design experiments

on a scale previously impossible in the social sciences discipline while expanding hand-coded material into larger datasets with machine learning. At the same time, with more data across a greater array of contexts, political scientists can now make use of big data to formulate and test hypotheses at a more detailed level.

Big Data Aids Hypothesis Generation: With the availability of masses of new data and the newfound ability to manipulate and investigate it quickly and cheaply, political scientists can observe patterns that have not been observed before. From this improved ability to describe and explore data comes the possibility of generating new and interesting hypotheses.

Conclusion

An attempt has been made in this chapter to outline the nature of big data and its usefulness in political science research. The chapter revealed that with a behavioral approach to political science inquiry, big data techniques have gained considerable traction to analyze causal relationships, make useful classifications, and forecasting possibilities in political science. It has made it possible for the researcher to investigate lengthy trends that political scientists are usually ill-equipped to map with traditional data sets. Instead of focusing on the assessment of the institution of the state and dealing with the normative way of generating data for political science research.

The chapter also notes that the big data techniques have made it possible for a researcher to monitor campaigns in a different location at the same time by assessing millions of speeches over hundreds of years to show how political language changes over time, or how a specific kind of contentious issue develops. If applied properly, with millions of observations and information generated through big data techniques, political scientists can observe previously unseen patterns; analyze political behavior at both the aggregate and capture previously unexplored phenomena.

It is therefore not surprising that political scientists use combinations of approaches and methods in their investigations, and this has made it possible for methods of political inquiry to consist of the various processes of acquiring data and treating data at the same time. It is to be stressed that while traditional approaches help to underline restricted features of the political process, the big data techniques have been found useful in generating data in larger quantities for political science research.

References

Anifowose, R. (1999). The nature and scope of political science. In Remi Anifowose and Francis Enemuo (eds). *Elements of Politics*. Lagos: Malthouse Press Limited.

Anastasopoulos, L. J and Poole, K. T. (2019). Big data in political science research.
Retrieved from *https://www.oxfordbibliographies.com/view/document/*

Christenson, D. (2018). Big data and political science. Retrieved from *https://www.bu.edu*

Clark, W. R., and Golder, M. (2015). *Big data, causal inference, and formal theory: contradictory trends in political science?* PS: Political Science and Politics 48(1): 65 – 70

Creswell, J. W. (2008). Educational research: Planning, conducting, and evaluating quantitative and qualitative research (3rd ed.). Upper Saddle River, NJ: Pearson Education, Inc. Retrieved from *https://www.scirp.org*

Eneanya, A.N. (2012). *Research methods in political science and public administration*. Lagos: University of Lagos Press Ltd.

Garuba, O.P. (2010), *An Introduction to political theory*. Macmillan Publishers, Delhi
http://www.kkhsou.in/main/polscience/approaches_polscience.html

Groh, A. (2018). Research methods in indigenous contexts. New York: Springer. ISBN 978-3-319-72774-5.

Monogan, J. E. (2015). *Political analysis using R*. New York: Springer, 2015. DOI: 10.1007/978-3-319-23446-5

Nachimias, C.F and Nachimias, D. (2009). *Research methods in the social sciences*. (5th ed.). London: Hodder Education.

Niels, G. (2016). What big data can teach political scientists. Retrieved from *https://blog.politics.ox.ac.uk/big-data-can-teach-political-scientists/*

Segal, T. (2022). What is big data? Definition, how it works, and uses. Retrieved from https://www.investopedia.com/terms/b/big-data.asp

Snelson, C. L. (2016). Qualitative and mixed methods social media research. *International Journal of Qualitative Methods*. 15 (1): Doi:10.1177/1609406915624574.

Taylor, D. (2022). What is big data? introduction, types, characteristics, examples. Retrieved from *https://www.guru99.com/what-is-big-data.html*

Trochim, W. M. (2005). *Research methods: The concise knowledge base*. Publisher: Atomic Dog Pub. Retrieved from https://www.researchgate.net/publication/51993096_

Snelson, C. L. (2016). Qualitative and mixed methods social media research. *International Journal of Qualitative Methods*. 15 (1): Doi:10.1177/1609406915624574.

Taylor, D. (2022). What is big data? introduction, types, characteristics, examples. Retrieved from *https://www.guru99.com/what-is-big-data.html*

Trochim, W. M. (2005). *Research methods: The concise knowledge base*. Publisher: Atomic Dog Pub. Retrieved from https://www.researchgate.net/publication/51993096_