A Systematic Literature Review of the Relationships between Policy Analysis and Information Technologies: Understanding and Integrating Multiple Conceptualizations

Cesar Renteria[1] and J. Ramon Gil-Garcia[1]

[1] University at Albany, State University of New York, Albany, NY, USA

crenteria@albany.edu, jgil-garcia@ctg.albany.edu

Abstract. Researchers and practitioners are increasingly aware of changes in the environment, broadly defined, that affect the policy process and the current capabilities for policy analysis. Examples of these changes are emergent information technologies, big and interconnected data, and the availability of computational power to perform analysis at a very disaggregate level. These and other forces have the potential to significantly change multiple stages of the policy process, from design to implementation and evaluation. The emergence of this phenomenon has led to the use of a variety of labels to define it. Potentially, a variety of labels might contribute to some conceptual confusion, but most importantly to concept stretching. This article aims to provide a conceptual space by identifying the attributes that compose the phenomenon. Based on a systematic literature review, this paper identifies the terms that have been used to refer to this phenomenon and analyzes their associated attributes. Based on Gerring & Barrosi's Min-Max strategy of concept formation, we propose two sets of attributes to define the phenomenon.

Keywords: Policy informatics, policy analytics, IT-enabled policy analysis, e-policy-making, systematic literature review, policy analysis, policy process.

adfa, p. 1, 2011.
© Springer-Verlag Berlin Heidelberg 2011

1 Introduction

Recent technological and analytical developments have grabbed the attention of researchers and practitioners as potential innovations that could improve the quality and timeliness of the policy process, compared with more traditional methods and approaches to collecting and analyzing information for policymaking. In this paper, we broadly refer to this as the relationship between policy analysis and information technologies. This relationship involves, among other things, new data sources and structures, improved computational capacity, and new methods of analysis that could contribute to better address the increasing complexity, interconnectedness, and uncertainty of public problems [1–4].

Recent academic interest has been brought on the matter from diverse fields, such as health [5], policy analysis [6–8], statistics [9], electronic government [10], population studies [11], complexity science [2], computational science [12], and informatics [2, 3, 13]. There is a variety of terms of concepts that have emerged from these backgrounds to refer to that phenomenon. For example, Jassen and Wimmer [2] identified the following terms: e-policy-making, computational intelligence, digital policy sciences, and policy informatics. Other authors have associated the terms IT-enabled policy analysis, policy modeling, and data-driven decision-making as other forms to label the same phenomenon [2, 14].

The high number of different terms to refer, arguably, to approximately the same phenomenon is a problem of conceptual clarity [15, 16]. Because every term has associated certain attributes, a number of terms imply a loose constellation of attributes associated with the same phenomenon. This conceptual ambiguity (in the constitutive attributes) might limit the possibility to build knowledge on top of previous works. Conceptual clarity is not to prefer a label rather than others, but to provide insights to facilitate future research on the matter. For example, as a basis for case selection or comparative analysis, for the operationalization of measurements, or to undergo a revision of the conceptual definition of current terms. Conceptual clarity also contributes to mitigate conceptual stretching, which is defined by Goertz [17] as "concepts [that] are loosened up so that they apply to additive cases. Thus, we seek to contribute in the study of the relationship of policy analysis and information technologies by defining a conceptual space[1] for the phenomenon of interest, as well as to provide two sets of attributes that best suit the definition of the phenomenon and have clear conceptual boundaries. Based on this, our research question is: what are the common and distinct attributes of the terms defining the relationship between policy analysis and information technologies?

The article is structured in five sections, including the present introduction. Section two explains the methodological approach used to develop the proposed conceptual space. This is followed by a brief description of the terms' backgrounds and an assessment of their conceptual clarity (or ambiguity). Section four provides a description of the minimal and ideal-type definitions, as well as the set of the constitutive

[1] A conceptual space is defined here as the range of attributes extracted from the concepts or terms aforementioned.

attributes of the proposed conceptual space. In the final section, we provide some conclusions and discuss future research directions.

2 Methodological Approach

2.1 Min-Max Strategy of Concept Formation

The Min-Max strategy of concept formation was proposed by Gerring and Barresi [15] as a mechanism to provide conceptual clarity, by uncovering the defining attributes of a concept. The strategy is particularly useful when uncovering these attributes across contesting defining terms. This is because the strategy focus on identifying the *non-idiosyncratic definitions* (those that are less dependent on particularities of certain field or period). Gerring & Barresi's strategy is based on Sartori's propositions of the "ladder of abstraction." This is in reference to the generality or specificity of a concept due to the augment or decrease of the concept's *intension*. The intension of a concept is the set of properties or attributes that determine the constitutive elements belonging to a concept [16]. The concepts are found to be more general by simply reducing the set of attributes, whereas the concept is more specific by adding or unfolding attributes [16]. These changes have a direct effect on the *extension*, which is the group of observations that have the attributes specified in the concept. Thus, the *extension* increases as *intension* decreases, and vice versa.

To define what attributes should be kept in a prototypical definition, Gerring & Barresi propose two strategies. The first strategy is a *minimal definition*. This refers to a set of necessary attributes that must be present in all terms or concepts. Identifying such attributes is an empirical endeavor, rather than theoretical. The goal is to identify the attributes that are present across all the concepts reviewed. This strategy aims to find a *non-idiosyncratic definition* (i.e. a set of attributes that will not vary across the terms used). The second strategy is an *ideal-type definition*. This strategy seeks to identify a definition that is "maximal" in that it includes all the attributes that could possibly compose the definition [15].

The empirical strategy proposed by Gerring & Barresi is unfolded in three steps. The first step is to gather a representative sample of the terms or concepts of interests. In this regard, our work departs from the lists presented in the introduction. Next, we did a systematic literature review to find the relevant manuscripts that use the terms or concepts of interest. The protocol of the systematic literature review is presented in the next section. The second step is to typologize the attributes from analyzing the manuscripts found in the systematic literature review. We built a typology of attributes by obtaining explicit referenced attributes (characterized here as "strong" attributes) and by interpreting implicit attributes (characterized as "weak" attributes). The third step corresponds to the organization of the attributes in two sets. The first set corresponds to a *minimal definition* and an *ideal-type definition*.

2.2 Systematic Literature Review

We conducted a systematic literature review of the relationship between policy analysis and information technologies in academic publications, following the widely used Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) statement. The PRIMSA protocol confers to the research process limited bias, transparency, and replicability [18]. We collected publications from three digital libraries to cover the publications in social sciences: Scopus, Web of Science, and JSTOR. These libraries combined offer the best coverage of publications in social sciences [19, 20]. In addition, we selected the digital library DBLP that accounts for the most extended coverage in computer science [21].

Our inclusion criteria were as follows. In terms of publications, we considered peer-reviewed articles, books or book chapters, and conference proceedings. All types of study designs were considered. We considered publications that: (1) provided or was related with a descriptive or conceptual discussion about policy analysis and information technologies; and/or (2) provided an application of an innovative approach, method or technology in the policy process. Our exclusion criteria were to limit the search to the following research fields: computer science, complexity science, health, informatics, and social sciences, which are the fields from which we had previous knowledge of being using related terms. We also discarded the publications that did not made a substantial reference to policy analysis, policy process or policy cycle.

Based on our previous knowledge of the topic, we considered the following search terms: 1) “policy informatics”, 2) “e-policy-making”, 3) “IT-enabled policy analysis”, 4) “policy modeling”, 5) “computational intelligence”, 6) “digital policy science”, 7) “policy analytics”, 8) “data science”, 9) “computational social sciences”, 10) “digital science”, 11) “data-driven decision-making”. For the database searching, we followed two rounds of search queries with clearly defined query rules.[2] The first round was based on a search of the exact terms, after the application of the inclusion criteria. In the first round, we noticed that our search strategy needed further specificity for some terms, since the publications retrieved were populous. For example, we retrieved 2,087 for “computational social sciences” (see Table 1). In the second searching round, all exact concepts were combined with common concepts in public administration literature ("governance", "public administration", "policy analysis", "policy-making", and "policy process"). The reason of this intersection was to automatically reduce the possibility of including articles that did not match the inclusion criteria, without scanning the titles or abstracts. This was a practical solution to address massive matches in broader concepts such as "computational intelligence", "data science" or "policy modeling".

[2] Illustratively, the search query string for policy informatics in the first round is shown below. All the terms followed the same string. ALL ("policy informatics"). In the second round was: ALL("policy informatics") AND (ALL("governance") OR ALL("public administration") OR ALL("policy analysis") OR ALL("policy-making") OR ("policy process")).

Table 1. Publications retrieved by first and second rounds of refinement

Concept	First round	Second round
Policy informatics	121	31
E-Policy-Making	15	6
IT-Enabled Policy Analysis	10	9
Policy Modeling	18,979	1,301
Computational Intelligence	175,418	34,760
Digital Policy Science	52	12
Data Science	11,504	2,083
Computational Social Sciences	893	76
Digital Science	1,156	104
Data-Driven Decision-Making	920	303

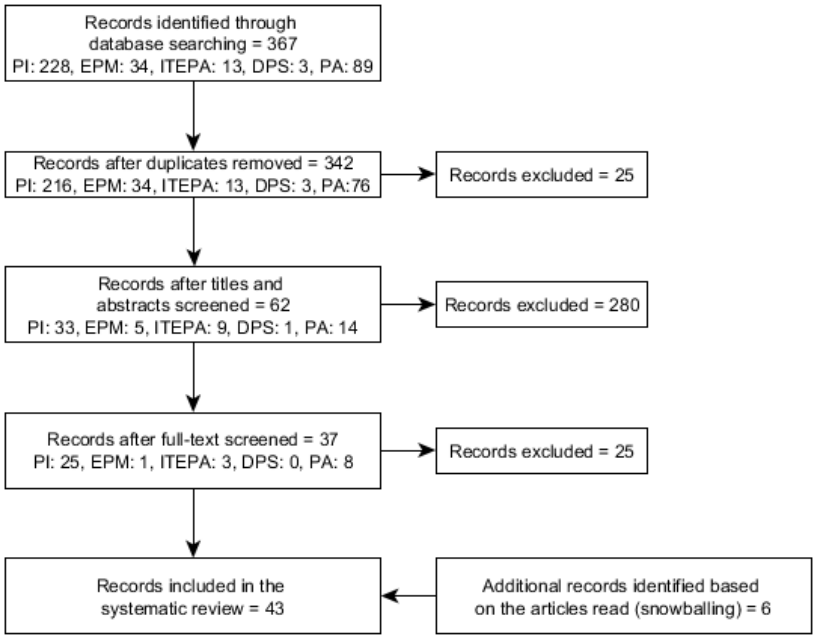
The search period of our review spans from 2000 to December 27th, 2016. By testing the search queries, we observed that most of the publications are not older than 8 years; however, we decided to extend the search period to 16 years to make sure we have a thorough coverage of research publications (thus, the lower bound is 2000).

Based on the results of the second round of search queries, we discarded some concepts from the study. The terms "computational social sciences", "data science", "digital science", "data-driven decision-making", "policy modeling", and "computational intelligence" were discarded as being considered too broad for the purposes of this research. We acknowledge that there is a chance that a subset of the articles retrieved with these terms might be related with the policy cycle and are not captured in the present study. In these cases, future research is needed in understanding whether these concepts could be associated with the policy analysis and how.[3] Finally, the concept "digital policy science", although we found publications in the search queries, this concept did not pass the fourth step in our PRISMA flowchart. Thus, we also discarded this term. The final list of terms was the following: 1) policy informatics, 2) e-policy-making, 3) IT-enabled policy analysis, and 4) policy analytics.

The search strategy consisted of five steps (see Figure 1). In the first step, we performed database searching following the inclusion and exclusion criteria defined above. The number of records identified in the first step was 367. In the second step, we extracted the reference metadata from the digital libraries and placed them into the reference manager software Mendeley, grouping the references by term. We then search for duplicates with the managing tools provided by Mendeley. As some records mentioned more than one term, we discarded duplicates within groups of references. We also grouped all records in a single file and searched for duplicates across groups, but keeping a research memo on the records that included many terms. The number of records after all duplicates were removed was 342. In the third step, we screened the titles and abstracts of the records to exclude the recrods that did not meet

[3] Furthermore, at the end of the literature review, we found three other related terms: "social data science", "data-centered policy-making", "data science for government and policy-making", and policymaking 2.0".

our inclusion. The number of records after screening titles and abstracts was 62. In the fourth step, we screened the full-text in the remaining records and removed those that did not meet the inclusion criteria. The third and fourth steps discarded mostly records that did not make any substantial reference to policy analysis, policy process or policy cycle. Finally, we included records identified while reading the selected records and were not identified through the search queries. This snowball sampling contributed to the identification of 6 relevant records that were included in the final sample. Based on that, the final number of records included in the systematic review was 43.



Notes: PI: Policy Informatics; EPM: E-Policy-Making; ITEPA: IT-Enabled Policy Analysis; DPS: Digital Policy Science; PA: Policy Analytics.

Fig. 1. Flow diagram of the search strategy and results

3 Understanding the Multiple Conceptualizations

3.1 "Insufficiently" Developed Terms

Aside from the terms that we considered not pertaining exclusively to policy analysis, policy process or policy cycle (computational social sciences, quantitative social sciences, data science, digital science, data-driven decision-making, policy modeling, and computational intelligence), we concluded that two of the terms reviewed lack of constituent attributes or they are vaguely defined. The first term that lies within this

vagueness is Digital Policy Science, from which although we found little evidence of its use in academic literature, there is neither an explicit definition, nor an implicit description of its constitutive attributes. The second term is e-policy or e-policy-making, which, although we found some attributes, we perceive them as weakly defined. In fact, the defining attributes found in the literature make direct reference to the defining attributes of other terms. As Hochtl [22] states, e-policy-making conceptually "shares many features of 'policy informatics,' such as analysis, administration, and governance [3]. and 'policymaking 2.0,' [...]" [22]. Sticking with the rule of discriminatory power in the intentionality of the attributes, we concluded that this term fails to set clearly some definitional boundaries as the definitional attributes are not described beyond these types of superficial descriptions.

3.2 Brief Presentation of Studied Terms

So far, the set of developments that have been perceived as useful for improving the policy process have been conceptualized in different terms. Janssen and Wimmer [2] include e-policy-making, computational intelligence, digital policy sciences, and policy informatics. Additionally, other authors have associated IT-enabled policy analysis, policy modeling, and data-driven decision-making as emergent concepts that also affect the policy process. Among these concepts, some have evolved recently as more complex conceptualizations than others, perhaps setting a framework for understanding the processes through which emergent developments could affect public policy.

For instance, IT-enabled policy analysis (ITEPA) is a framework that seeks to advance the study of policy analysis by integrating the views of Bardach's policy cycle with Sterman's system dynamics approach. This framework also expands the conceptualization of policy analysis as a task with a necessary combination of institutions, actors, data, and information technologies. ITEPA is a perspective primarily concerned with the relationship between government and citizens. Its background spans through the study of e-government and governance. In this perspective, the substantial change in policymaking is caused by the development of open government and open data initiatives [10, 14, 23]. In this sense, the deployment of open government policies, and the release of open data in particular can strengthen new mechanisms of government-citizen relationships such as co-production, collaboration, and participation of the citizens in public processes as they are provided with new sources of information on public issues. Because data is regularly an input for decision-making, the increased availability of public data in the public domain may shift the approach to policymaking from a top-down decision-making approach towards a networked participatory approach. Furthermore, potential changes in policymaking are not only driven by data and technology possibilities, but also by a set of governing principles on the rise: transparency, participation, collaboration, and empowerment.

E-Policy-Making (EPM) is a term less used and developed in its contents. This terms refers to the use of "e-governance processes" in policymaking. Hochtl [22] points out that the concept formation intersects the attributes of policy informatics and Policymaking 2.0. Furthermore, the authors imply that the study of e-policy-making encompass both the improvements of already existing structures in policymaking by

the incorporation of technology, as well the transformation of the policymaking structure itself [22].

Policy analytics is a concept formed to encompass different methods able to cope with the growing challenges that the rise of Big Data poses on data analysis [24, 25]. Its conceptualization is an adaptation of the idea of "business analytics" in the private sector, but applied to public policies. The focus of this conceptualization is to understand which methods or approaches could contribute to leverage the emergence of massive, complex and unstructured data production for the decision-making in the policy process. Thus, policy analytics is a field of study about how to adapt the set of skills, applications, methods, and technologies that lie within the field of data science to assist the construction of evidence for decision-making. Policy analytics is a perspective primarily concerned with the exploit of data quality, quantity and availability commonly known as the data revolution [24–26]. Overall, analytics are perceived as a way to solve the challenges associated with analyzing massive and unstructured data [25]. The term comes primarily from the field of operations research, as an attempt to understand the implications of adopting quantitative decision support methods in the private sector known as business analytics. Business analytics is a collection of innovative computational techniques to leverage and, at the same time, cope with the challenges of managing big data to inform decision-making.

Decision-making under this perspective pursues the ideal of evidence-based decision-making, where big data features and data science techniques are perceived as more accurate, rich, and timely, as well as less costly than traditional methods of collecting data [24, 25]. Furthermore, these are perceived as less biased in the collection and interpretation processes [24, 26]. As for the policy process, this perspective directly implies the incorporation of new sources of information for decision-making; remarkably, Daniell [25] has attempted to organize the blending of data sources and data science techniques and associate them as tools for policy analysis at different stages in the policy cycle. In addition, policy analytics indirectly implies the incorporation of predictive analysis to the toolkit of policy analysts.

Table 2. Summary of Definitions

Concept	**Definition**
E-Policy-Making	“the act of policymaking in e-government using e-governance processes, with the distinctive feature that evaluation happens as an integral part of all along the policy cycle rather than a s a separate step at the end of policymaking process”. [22]
IT-Enabled Policy Analysis	“The use of IT tools, mathematical modeling and analytical methods to take advantage of the available data to aid individuals and groups make policy options or solve policy problems.” [14]
Policy Informatics	“The study of how computation and communication technology is leveraged to understand and address complex public policy and administration problems and realize innovations in governance processes and institutions.” [3]
Policy Analytics	“The development and application of […] skills, methodologies, methods, and technologies, which aim to support relevant stakeholders

	engaged at any stage of a policy cycle, with the aim of facilitating meaningful and informative hindsight, insight and foresight." [27]

4 Integrating the Attributes from Multiple Terms: A Min-Max Approach

Even though the Min-Max strategy is generally used to set the conceptual space of a given concept (i.e. democracy or culture), here we used this method to ensemble the attributes of a constellation of concepts that we hypothesized belong to the same latent concept. The argument is that this latent concept has been characterized through different lenses, where each lens has its own background and thus would likely assign different attributes to the same phenomenon.

Each concept formation offers, explicitly or implicitly, a variety of definitional attributes, where some of them will intersect each other, whereas others are context-specific reminiscent or idiosyncratic attributes [15, 28]. Thus, the *minimal definition* comprises the attributes that are common to all perspectives, providing a more general and agnostic perspective on the phenomenon.

The relationship between policy analysis and information technologies, minimally defined, is a phenomenon composed by the development of methods, technology, and data. Adding the idiosyncratic attributes allows to broaden the definition to a more overarching idea without blurring the definitional boundaries. This *ideal-type definition* includes all the attributes that comprise the conceptual space of the phenomenon, regardless of the perspective. The *ideal-type definition* is equivalent to the conceptual space in that both comprise the full range of attributes. Thus, as will be described below, the conceptual space is composed by five attributes: human resources, governance, methods, technology, and data (see Table 3).

Table 3. Table of attributes, by term

	Human resources	**Governance**	**Methods**	**Technology**	**Data**
E-Policy-Making		Weak	Weak	Weak	Weak
IT-Enabled Policy Analysis	Strong	Strong	Strong	Strong	Strong
Policy Informatics		Strong	Strong	Strong	Strong
Policy Analytics			Strong	Weak	Strong
Typology of attribute	Ideal-Type	Ideal-Type	Minimal	Minimal	Minimal

Note: "Strong" stands for strongly defined attribute, whereas "Weak" means weakly defined attribute.

We also found that some terms had a stronger identification of attributes than others. For example, e-policy-making had a weak definition of attributes (i.e. subject to textual exegesis). Policy analytics has a strong emphasis on data as the fundamental attribute constitutive of the concept, whereas there is a weak connection with technol-

ogies. As for IT-enabled policy analysis and policy informatics, there is an evenly strong identification of attributes across the definitional works reviewed. Interestingly, although the IT-enabled policy analysis framework has more constitutive attributes than policy informatics, the authors consider this as a possible variation of policy informatics [10].

The results show that the attributes observed across the concepts are convergent towards a unified core of attributes (the minimal definition). All these perspectives are relatively aligned, with some idiosyncratic attributes that are likely to be explained by their intellectual background.

4.1 Minimal Attributes

Methods

In this context, methods refer to the analytical tools that people, primarily in the public sector, use to obtain useful information insights or knowledge from data for decision-making. These methods are generally suited for the analysis of quantitative data, although Puron-Cid, Gil-Garcia and Luna-Reyes [14] also recognize analytical methods for qualitative data. The perceived constitutive methods of policy analysis span from a wide variety of fields, such as mathematics, statistics, economics, operations research, psychology, sociology, management, finance, and political science [14, 24, 25, 27, 29], although the focus is stressed in emergent methods or techniques from computer science. For example, text mining, exploratory data analysis, support vector machines, spreadsheet models, and machine learning [25]. Other methods considered that are not yet part of the traditional policy analyst's toolkit are group model building, multi-criteria analyses, simulation and optimization modelling, participatory planning, resource allocation modelling, real-time operations optimization, remote sensing, smart metering, and participatory GIS/evaluation [25]; simulation modelling, and cognitive mapping [2, 3, 14, 25]. There is no clarity however, on whether the list of methods considered as part of this phenomenon is pertaining to a single technological innovation (i.e. data revolution) or the concept should be extended to all methods that eventually would fit technological innovations in the future.

Data

Data is an input for decision making in the policy process [10, 30–32]. This is regarded as sets of measurements of social activity that require analytical skills or methods to be transformed into useful information or knowledge for decision making. Data is also perceived as an element that could be easily spread into the public domain, contributing to re-shape the relationships between government and citizens. Under certain circumstances, data is also perceived as a potential driver towards networked governance, transparency, and other activities beyond policy decision-making [14]. Although this attribute also refers to traditional data (e.g. survey data), the primary focus is on big data, which is perceived as massive, usually costless, and unstructured. There are several types of such data, primarily organized by its source; for example, commercial data, administrative data, open data, electronic data, on-line

data, cellphone data, geospatial data, daily census data, as well as data from sensor readings and crowd computing sources.

Technology
Technology or IT tools refers to the computational infrastructure that contributes to increase government's capacity [14, 30, 33–35]. This increased capacity comes in a variety of activities linked with both decision-making and governing. These activities are the likes of visualization technologies for the communication of policy analysis or decisions, technologies to process and manage information overload and ambiguity, technologies to generate and collect data, technologies for understanding patterns and detect trends in data, to increase the reach of the policy discussion, to enable collaborative networks, or to crowdsource public policy analysis, policy monitoring, evaluation or implementation.

4.2 Idiosyncratic Attributes

Human Resources
This attribute refers to the stock of capacities and skills for policymaking in the human resources available in an organization [14]. Any given development in computation and communication in the organization requires a body of expertise and knowledge in the human resources to effectively accomplish any task. More specifically, human resources refer to the personnel in charge of tasks such as producing insights for or advising decision-making, as well as being responsible for decisions in any given stage of the policy cycle.

Governance
Governance refers to the technological infrastructure through which governance processes could occur [14, 36–42]. The governance platforms are perceived to have the capacity to improve the flexibility and responsiveness of bureaucracies [39]. In addition, improved computation and communication capabilities in government activities could improve the interaction between citizens and government, as well as among government agencies. As a constitutive element for decisions in policymaking, governance represents the institutional and social arrangements in which a decision-making process takes place. As such, governance is highly intertwined with the rest of the constitutive elements of policy analysis and information technologies, since it might be shaped by data or technologies or it might determine the use of types of data and technologies in the policy process.

5 Conclusion

Rising literature on this topic suggests that information technologies and the availability of new types of data are already affecting the policy process and the way people

think about policy analysis. The increasing computational power and alternative analytical methods also add to this situation and make it complex and not easy to conceptualize. In this paper, we have shown that there are many labels to refer to this phenomenon, including policy informatics, IT-enabled policy analysis, and E-Policy, among others. Despite the very diverse labels there are some important commonalities that should be part of our understanding of this phenomenon and a more comprehensive, but concise definition. There are aspects related to methods, data, technology, human resources, and governance and all of them contribute to a rich conceptualization of the relationships between information technologies and policy analysis.

In addition, this analysis identifies some terms with insufficient conceptual development, and other that are not clearly related to the policy process. In contrast, there are a core set of conceptualizations that help to identify the common attributes and specific attributes of the phenomenon in our approach. There are also references to the policy cycle, but almost no strong links to theories of the policy process. This is still a limitation of the existing terms and their respective conceptual definitions. Finally, there is also the challenge of identifying and integrating new technologies or analytical methods when they emerge. The usefulness of a concept related to these important and frequently rapid changes is also related to its capacity to leverage past research and being useful for future research. This paper aimed to contribute to conceptual clarity by uncovering some underlying attributes and structure them into two sets of definitions. Since ideas are not set in stone, it may well make sense to use these insights as material to reassess the conceptual definition of current terms.

References

1. Desai, A., Kim, Y.: Symposium on Policy Informatics. J. Policy Anal. Manag. 34, 354–357 (2015).
2. Janssen, M., Wimmer, M.A.: Introduction to policy-making in the digital age. (2015).
3. Johnston, E.W., Desouza, K.C.: Governance in the information Era: Theory and practice of policy informatics. Routledge (2015).
4. Lazer, D., Pentland, A.S., Adamic, L., Aral, S., Barabasi, A.L., Brewer, D., Christakis, N., Contractor, N., Fowler, J., Gutmann, M.: Life in the network: the coming age of computational social science. Sci. (New York, NY). 323, 721 (2009).
5. Martin, E.G., MacDonald, R.H., Smith, L.C., Gordon, D.E., Tesoriero, J.M., Laufer, F.N., Leung, S.J., O'Connell, D.A.: Policy Modeling To Support Administrative Decisionmaking On The New York State Hiv Testing Law. J. Policy Anal. Manag. 34, 403–423 (2015).
6. Blume, G., Scott, T., Pirog, M.: Empirical innovations in policy analysis. Policy Stud. J. 42, 33–50 (2014).
7. Cook, T.D.: "Big data" in research on social policy. J. Policy Anal. Manag. 33, 544–547 (2014).
8. Pirog, M.A.: Data will drive innovation in public policy and management research in the next decade. J. Policy Anal. Manag. 33, 537–543 (2014).

9. Crosas, M., King, G., Honaker, J., Sweeney, L.: Automating Open Science for Big Data. Ann. Am. Acad. Pol. Soc. Sci. 659, 260–273 (2015).
10. Puron Cid, G., Gil-Garcia, J.R., Luna-Reyes, L.F., Puron-Cid, G., Gil-Garcia, J.R., Luna-Reyes, L.F.: IT-enabled policy analysis: new technologies, sophisticated analysis and open data for better government decisions. In: DG.O. pp. 97–106. ACM (2012).
11. Kum, H.-C., Krishnamurthy, A., Machanavajjhala, A., Ahalt, S.C.: Social genome: Putting big data to work for population informatics. Computer (Long. Beach. Calif). 56–63 (2014).
12. Zeng, D.: Policy Informatics for Smart Policy-Making. IEEE Intell. Syst. 2–3 (2015).
13. Longo, J., Wald, D.M., Hondula, D.M.: The Future of Policy Informatics. In: Johnston, E.W. (ed.) Governance in the information Era: Theory and practice of policy informatics. pp. 335–352. Routledge (2015).
14. Puron-Cid, G., Ramon Gil-Garcia, J., Luna-Reyes, L.F.: Opportunities and Challenges of Policy Informatics: Tackling Complex Problems through the Combination of Open Data, Technology and Analytics. Int. J. PUBLIC Adm. Digit. AGE. 3, 66–85 (2016).
15. Gerring, J., Barresi, P.A.: Putting ordinary language to work: a min-max strategy of concept formation in the social sciences. J. Theor. Polit. 15, 201–232 (2003).
16. Sartori, G.: Concept misformation in comparative politics. Am. Polit. Sci. Rev. 64, 1033–1053 (1970).
17. Goertz, G.: Social science concepts: A user's guide. Princeton University Press (2006).
18. Petticrew, M., Roberts, H.: How to appraise the studies: an introduction to assessing study quality. Syst. Rev. Soc. Sci. A Pract. Guid. 125–163 (2006).
19. Albarillo, F.: Language in social science databases: English versus non-english articles in JSTOR and Scopus. Behav. Soc. Sci. Librar. 33, 77–90 (2014).
20. Norris, M., Oppenheim, C.: Comparing alternatives to the Web of Science for coverage of the social sciences' literature. J. Informetr. 1, 161–169 (2007).
21. Cavacini, A.: What is the best database for computer science journal articles? Scientometrics. 102, 2059–2071 (2015).
22. Höchtl, J., Parycek, P., Schöllhammer, R.: Big data in the policy cycle: Policy decision making in the digital era. J. Organ. Comput. Electron. Commer. 26, (2016).
23. Sandoval-Almazán, R.: Open government and transparency: Building a conceptual framework , (2015).
24. De Marchi, G., Lucertini, G., Tsoukiàs, A., Marchi, G. De, Lucertini, G., Tsoukiàs, A.: From evidence-based policy making to policy analytics. Ann. 236, 15–38 (2016).
25. Daniell, K.A., Morton, A., Ríos Insua, D.: Policy analysis and policy analytics, (2016).
26. Larsson, A., Taylor, S., Wandhöfer, T., Koulolias, V.: Exploiting online data

in the policy making process, (2015).
27. Tsoukias, A., Montibeller, G., Lucertini, G., Belton, V.: Policy analytics: an agenda for research and practice. EURO J. Decis. Process. 1, 115–134 (2013).
28. Dreyer, D.R.: Unifying conceptualizations of interstate rivalry: A min–max approach. Coop. Confl. 49, 501–518 (2014).
29. Scharaschkin, A., McBride, T.: Policy analytics and accountability mechanisms: judging the "value for money" of policy implementation, (2016).
30. Barrett, C.L., Eubank, S., Marathe, A., Marathe, M.V., Pan, Z., Swarup, S.: Information integration to support model-based policy informatics. Innov. J. 16, (2011).
31. Henman, P.: e-Government, public policy and the growth of conditionality, (2005).
32. Kokkinakos, P., Koussouris, S., Markaki, O.I., Koutras, K., Psarras, J.E., Glickman, Y., Loehe, M., Lee, H.: Open Data Driven Policy Analysis and Impact Evaluation. In: EEPM@ eGOV (2015).
33. Tait, E.: Web 2.0 for eParticipation: Transformational tweeting or devaluation of democracy? (2013).
34. Charalabidis, Y., Loukis, E.: Participative public policy making through multiple social media platforms utilization. Int. J. Electron. Gov. Res. 8, 78–97 (2012).
35. Lampe, C., Zube, P., Lee, J., Park, C.H., Johnston, E.: Crowdsourcing civility: A natural experiment examining the effects of distributed moderation in online forums. Gov. Inf. Q. 31, 317–326 (2014).
36. Johnston, E.: Governance Infrastructures in 2020. Public Adm. Rev. 70, S122–S128 (2010).
37. Kim, Y., McGraw, C.: Use of agent-based modeling for e-governance research. In: ACM International Conference Proceeding Series (2012).
38. Krishnamurthy, R., Desouza, K.C., Johnston, E.W., Bhagwatwar, A.: A glimpse into policy informatics: The case of participatory platforms that generate synthetic empathy. Commun. Assoc. Inf. Syst. 33, (2013).
39. Wachhaus, T.A.: Governance as a framework to support informatics. Innov. J. 16, (2011).
40. Prpić, J., Taeihagh, A., Melton, J.: The fundamentals of policy crowdsourcing. Policy and Internet. 7, (2015).
41. Lampe, C., LaRose, R., Steinfield, C., DeMaagd, K.: Inherent barriers to the use of social media for public policy informatics. Innov. J. 16, (2011).
42. Kelley, T.M., Johnston, E.: Discovering the appropriate role of serious games in the design of open governance platforms. Public Adm. Q. 36, 504–554 (2012).