



Big data and public policies: Opportunities and challenges

Giovanni Azzone

Department of Management, Economics and Industrial Engineering, Politecnico di Milano, Italy

ARTICLE INFO

Article history:
Available online 19 February 2018

Keywords:
Public policy design
Public value
Big data
Personal policies

ABSTRACT

Big data are often presented as a strategic opportunity for the design of new public policies, improving the quality and effectiveness of public services and using resources more efficiently. The paper discusses such opportunities and identifies a few open questions.

© 2018 Elsevier B.V. All rights reserved.

1. Introduction

Secchi (2018) clarifies how the “big data revolution” is creating new challenges to statisticians: volume, velocity and variability of often unstructured data will (or could) require new theories, methods, and tools for data integration and visualization. However, such an effort will result in “a better society”, only if there are “problems” where the characteristics of big data can improve the quality of decision making and the decision makers themselves are aware of such opportunity and willing to take advantage of it.

The **design and implementation of public policies** is, from this point of view, a promising area for an extended use of big data in decision making (Maciejewski, 2017; Mercel, 2017). In this paper, I will try to clarify why there is a strong potential in supporting public policies through big data and what is the value added to our society. I will also argue that to turn such potential into reality, we must face new organizational and conceptual problems, whose solution requires the joint effort of statisticians together with scholars of many other disciplines, from ethics to law, from political science to management.

The paper is articulated in three main sections:

- Section two describes the key dimensions of the quality of a public policy, and the processes needed to create value through public policies and services;
- Section three discusses the interactions between some characteristics of “big data” and the quality of public policies and services;
- Section four presents a few open questions, whose answer is needed for an actual use of big data information in public policies.

2. Adding value to public policies

A **public policy** can be defined as *the set of actions that affect the solution of a policy problem, i.e. a dissatisfaction regarding a certain need, demand or opportunity for public intervention* (Dente, 2014). Its quality is measured by the capacity to create public value; Osborne (2017) describes public value as the combination of four different characteristics. Two of them are somehow “traditional”, as follows:

E-mail address: giovanni.azzone@polimi.it.

- The **capacity to solve individual problems**; we expect that the health system finds a therapy for our illness; that our wastes are collected, ...
- The **capacity to improve the quantity and quality of life in the Country**. Such dimension is more complex to assess in an objective way, as citizens could have different ideas about what “improves quality of life” in a Country. However, there are some “shared values” that can lead to “shared measures”. For instance, the increase in “expected life” can be used as a measure of the quality of health policy, or the percentage of recycled wastes as a measure of quality of an environmental policy.

Over the last years, however, two additional dimensions gained relevance, due to changes in society.

- The whole **life experience that individuals get from a public service**, i.e. its wide impact on the life of an individual. A university student assesses the “public service of higher education” according to the contents of its courses and the job opportunities he will find after graduation (i.e. measuring the *capacity of the university to solve the individual problems of the students*), but also rates the opportunity to have international experiences or to build an interpersonal networking, the quality of internships or the “passion” of his/her professors. Similarly, a patient evaluates hospital services also by the kindness of nurses or the quality of food.
- The **readiness for change**, that is assuming increasing importance as the world is becoming more and more dynamic. In such an environment, universities are required to educate students to learn how to learn (Van Vaught, 1999), or hospital to prepare ill people to live with a chronic pathology (Bodenheimer et al., 2002).

The two latter dimensions of public value increase the importance of some characteristics of policy design and implementation:

- The capacity to **provide a personalized service**, i.e. a service consistent with requirements and expectations of each citizen. Today, we get a “customized service” when we buy a car or choose foods in a restaurant, hence we expect the same approach from public services;
- The **involvement of final users in the process of designing and producing such personalized service**. A public service always involves the user, but sometimes the cooperation is unintended and at an individual level (we cannot educate a student without his/her active participation...); today, there is a growing claim for co-design (where the individual and the public administration jointly design the services) or co-innovation, where actions are taken at the level of groups of users interacting with the administration.
- The **capacity to change the characteristics of the service over time**, according to the evolution of personal and societal requirements and expectations.

Big data can improve the quality of public policies acting on all these characteristics.

3. Big data as a way to create value in public policies

The term “Big data” is applied to data sources with different characteristics, moving from structured administrative data bases to unstructured weak signals from social media, leading to different possible uses of each data source in supporting public policies. For our goals, we can classify the data sources according to three dimensions:

- The “completeness” of data, i.e. their availability for the whole population interested by the policy;
- The “timeliness” of data, i.e. a limited time lapse between an event and the availability of information concerning the event;
- The “voluntariness of users to supply data and information” (for instance writing a post on Facebook or answering to a question on a web site), as opposed to cases when they are forced by law or they are unaware of providing data (as for instance when we monitor their “signals” through cellular phones and GPS).

Below, we independently analyze the impact of each of these characteristics on public policies; however, in general we can have combined effects, depending on the specific data sources we are considering.

3.1. Designing “personalized policies”

A first opportunity arises from the possibility to manage very large data bases, including complete and detailed personal information about policy users, hence supporting the design of “personalized” or “precision” policies. Traditionally, public policies were “standard”, providing the same opportunities and services to all users, or to wide groups of them (for instance, a social policy can be dedicated to all citizens with low income or and economic policy to small enterprises). But requirements and expectations are only partially shared by users in the same group: policies can only be designed for the “average user”, thus resulting in an excess of resources for some and in a lack of resources for others. Precision policies focus on the specific expectations of homogeneous groups of users (in principle, even the expectations of a single user, if complete and detailed data at individual level are available), thus allowing an increase in efficiency and effectiveness of public policies.

3.1.1. An example of improving efficiency and effectiveness of public policies: Casa Italia project

An example of such use of big data is the project Casa Italia, a risk prevention policy aimed at reducing the vulnerability of buildings in case of earthquakes in Italy (www.governo.it/approfondimento/piano-casa-italia/7093). Previously, the policy of the Government was based on a financial support granted to the owners of the building, when they invested to reinforce their houses. The incentive, designed as a tax deduction, was determined as a given share of the investment, regardless of the location of the building (that influences hazard) and of its vulnerability before reinforcement. Such generic financial support resulted to be ineffective in the most hazardous cities, where the value of the building is lower and property is fragmented. At the same time it resulted in a waste of public money in other places, as where the owners of the buildings would have been able to pay the investment back thanks to the increase in market value after reinforcement. A “big data analysis” of the “individual” characteristics of over 18 million buildings in the country, allowed to identify 550,000 “more risky building”, and to design for them a customized policy, where the first stage of the intervention (the diagnosis) is fully financed by the central government; such action is sustainable for a limited number of buildings (3% of the total), while it would not be, if applied in a standard way to all buildings in the Country. The next step will be to deepen the analysis, customizing the policy according to other characteristics of the building (for instance, increasing the public contribution when the building is adjacent to another or on a road for emergency vehicles) or of the owner (for some citizens tax reduction could be replaced by a grant component or other kinds of financial support).

3.1.2. Big data as enablers: the National infrastructure for electric vehicles charge

Sometimes, personalization does not only increase effectiveness and efficiency, but rather enables otherwise unfeasible policies. An example is the design of a National infrastructure for charging electric vehicles in Italy. Plan started in 2016, when first economic estimates, based on standard assumptions on “average behaviors”, forecasted an investment of about 1 billion €, unsustainable from the National Government (that did not consider it a priority) and economically unjustified for private companies, due to the limited number of circulating electric vehicles. A “chicken and egg” problem: without a charging infrastructure covering all main roads, the number of electric vehicles would remain very low; but such limited number would not encourage private investments. The problem was then analyzed through a “big data approach”, trying to understand how many people in each municipality in Italy would be interested in having an electric car, and for how many kilometers a day each one would use it. To this aim, demographic data on the residents of the 8,000 Italian municipalities and their personal mobility were analyzed, estimating the number of charging infrastructures needed in any place. In the end, the size of required investments resulted to be about 20% of what previously expected, hence enabling a private operator to proceed to the development of the infrastructure network.

3.2. Weak signals as a way for anticipating policy changes

A second opportunity arises from gathering the weak signals, that people “send” unintentionally, as those derived from monitoring cellular phones or tracking energy consumptions. Such information can anticipate socio-economic trends. Hence they represent “timely warnings” for policy makers, who can then activate a further robust statistical analysis, to understand if the weak signals are effective predictors of long term trends, requiring a policy redesign, or are biased on crude and premature analysis.

As an example, the project Urbanscope (www.urbanscope.polimi.it) collects such information in the urban area of Milan, Italy. Some of these data, as the number of phone calls to and from a specific country or the language used in social networks, could be used to monitor immigration flows in the different districts of the city. They are weak signals, as an increase in the number of phone calls to Philippines or in the number of tweets written in Tagalog language could depend both on the growth of the Philippine community in Milan or on a contingency, like a fair or a wedding. But they can represent a warning for administrators, who can use other tools to check the consistency of the signal and, if needed, anticipate possible actions.

3.3. Social data as a way to involve users in the design and creation of services

Data and information that users voluntarily “post” in social network can play a strong role to involve citizens in the design and production of public services, hence empowering their “life experience” (Meijer et al., 2012). The involvement can be aimed at (Osborne et al., 2016):

- Improving the performance of existing services, by asking users an assessment, testing their sentiment towards possible policy changes before introducing them, and more in general gathering suggestions (*codesign*); many public administrations are using their web site and their social network accounts to this aim;
- Designing brand new services; here, social network can be used as a way to interact with groups of actual or perspective users for gathering ideas for new services, whose creation could also involve the user group. Students' groups and students' representatives can often play such a role in universities, suggesting new services (as assessment centers, startup competitions, project based learning, social events...) that universities could put in practice with their active support (*coinnovation*).

In both cases, data analysis is needed to understand how individuals' and groups' ideas and sentiment are really relevant for larger sets of users or, on the contrary, could lead to choices that most would refuse.

4. New challenges

To turn all these opportunities in a real improvement of the value added by policies, we still must face several problems. Some are well known, but intensified by “big data”, others fairly news. I will focus on six questions:

- How can we synchronize evolution of social and individual expectations with evolution of legislation?
- How can we guarantee “equity” among all citizens?
- How can we manage fake news and opportunistic behaviors in providing information?
- How should we balance individual privacy and general interest?
- How should we adapt roles and competences of public managers?
- How can we adapt the attitude of public managers towards accountability and responsibility?

4.1. Environment evolution vs. legislation times

Public policies act through laws and norms, that should change according to the evolution of the society's values and expectations. Personal policies further enhance this issues, as they must be adapted not only to general or “average” changes, but also to specific requirements of single users, according to data and information that continuously evolve; hence, the available time to modify laws and norms decrease.

How can we accordingly increase the speed of the legislation process, without violating the rules of democracy? Could we imagine to create *parametric laws*, where the service provided by the administration is *ex ante* linked to parameters evolving over time? Or *contingent laws*, that explicitly define different scenarios for the future, anticipating how the law will evolve if one of them happens?

4.2. The concept of equity

Equity, i.e. fairness and impartiality towards all concerned, based on the principle of giving the same advantage to all users, is a key objective of public policies. In “standard” public policy, equity was often interpreted as giving to all users the same opportunity to access a service (all students with a given mark in high school can attend higher education, all people suffering from the same disease can access hospitals) and the same level of service.

Personal policies question such approach; if public value depends on the expectations of individuals, the same level of service can provide different values (hence, different advantages) to each of them. But how can we determine for each individual the “fair” level of service?

4.3. Fake news and opportunistic behaviors

Unexperienced people find it more and more difficult to distinguish between information based on real data and fake news. However, if we do not enhance such capacity in all citizens, individual expectations could be based on fake information, thus compromising the quality of public policies design according to these expectations.

The problem is farther enhanced by opportunistic behaviors, especially when dealing with data voluntarily provided by users. The web is a chaotic system, where the consistent messages of few people can give the impression of a general problem. Besides, companies and individuals can directly “manage” information flows, giving more emphasis to some posts, or artificially creating messages through “bot” software.

How can we distinguish real data and information from “noise”?

4.4. Balancing privacy rights and general interest

The increase in volume and detail of personal information, available for public policies, enhances the problem of which information should be made public and which should not. An example can clarify the point. Let us consider the consequences of the Casa Italia Project described in Section 3.1.1, where the Italian Government will take care of the diagnosis of the level of the vulnerability of some buildings, upon request from the owner of the building. Who should receive the resulting information about the actual vulnerability of the building? No doubt that the owner and the residents, or the public administration, must be informed. But what about people living in other buildings in the same road? On the one hand, a vulnerable building could increase the probability of a crash in adjacent ones, in case of earthquake, hence we should inform them of the problem. On the other hand, if the owner of a “potentially dangerous building” knows that all his/her neighbors will be informed, there is a risk that he will not apply to the program, hence reducing its overall impact on risk prevention.

4.5. The role and competence of public management

Personal policies, policy changes based on weak signals and involvement of citizens through social media strongly redefine the role and competence of public managers (Dunleavy, 2017). In general, public managers contribute to policy design and implementation:

- Defining the decision space, i.e. the boundaries that must be respected in policy design;
- Anticipating future impacts of alternative solutions, helping decision makers to identify the best one;
- Monitoring the environment, to suggest possible needs of new policies.

All these activities are influenced by big data:

- The **decision space** is not only limited, as in the past, by legislation and economic boundaries. We must also verify available information, checking if it is consistent with the requirements of personal policies, in terms of complete coverage of potential users, data reliability, privacy issues, etc.
- The expected **impact of alternatives** cannot be determined, as in the past, through a collection of publicly available information, usually prepared by public managers themselves or commissioned to external experts. Collecting and analyzing big data in a way able to support policy design requires strong and continuous interactions between public managers – who know the objectives of the policy and the boundary conditions – and data experts, who can “read” data.
- The problem of **monitoring the socio-economic environment** requires that public managers are able to collect weak signals – not underestimating their relevance – and to divide useful information from “noise”, taking advantage of new technologies and methods. The issue is very important, as individual and lobby groups could attempt to use social media to influence policies, playing with the system;¹ the public manager should hence be able to down weight some information to avoid being influenced by unfair actions.

In summary, big data require fairly new competences and skills in public managers. Will the public administration be able to do such a change, without limiting, as in the past, to a “formal” adaptation?

4.6. Accountability and responsibility of public managers

Finally, designing and implementing public policies through big data requires a change in the attitude of public managers. In “standard” policy, managers must “apply” a law; often the norm is a justification towards the users: *although I do not agree, I am obliged to follow the rules*. Personal policies open a space of autonomy for managers; they are in charge of understanding and deciding the best way to answer to the expectations of individuals, within “wide” boundaries. Hence, the public manager becomes responsible for the results of the policy. The difficulty of this change of attitude must not be underestimated, as in the past it was a main reason of failure in the innovation of public policies.

5. Conclusions: an ambitious research agenda

The six open issues I listed above represent an ambitious research agenda. Dealing with these problems do not merely imply the promotion of coordinated studies in different disciplines, but rather to establish a long term transdisciplinary research program, where even the definition of each question requires the strong integration of theories, methods and languages of different disciplines. Such program will be successful if it leads to feasible solutions, able to modify the way institutions design “real” policies and recruit, train and assess their management, hence requiring a continuous interaction between researchers, politicians and administrators.

Is this too ambitious? Should we accept that the use of big data in public policy remains a “fashionable idea”, but unable to change the quality of life of citizens?

I do not think so, as the stake is too high to be lost without trying. Personal policies, involvement of citizens in policy design through social data, and early warnings to enhance timeliness in policy redesign are a unique opportunity. They can increase the efficiency and effectiveness of public policies, balancing the crisis of the “Welfare State” – due to a lack a financial resources and an increasing need of public services – that most countries, especially in Europe, are experiencing, effectively counteracting populisms, which represent the most visible and dangerous consequence of such crisis.

References

- Bodenheimer, T., Wagner, E.H., Grumbach, K., 2002. Improving primary care for patients with chronic illness. *JAMA* 288 (14), 1775–1789.
- Dente, B., 2014. *Understanding Policy Decisions*. Springer.
- Dunleavy, P., 2017. Digital change, modern bureaucracy and big data, In: EGPA 2017 Conference, Milan.
- Maciejewski, M., 2017. To do more, better, faster and more cheaply: using big data in public administration. *Int. Rew Adm. Sci.* 83 (1), 120–135.
- Meijer, A.J., Curtin, D., Hillebrandt, M., 2012. Open government: connecting vision and voice. *Int. Rew Adm. Sci.* 78 (1), 10–29.
- Mercel, I., 2017. Big data in public affairs, In: EGPA 2017 Conference, Milan.
- Osborne, S.P., 2017. Delivering public services. Adding value to citizens through co-production. In: EGPA 2017 Conference, Milan.
- Osborne, S.P., Radnor, Z., Strokosch, K., 2016. Co-production and the co-creation of value in public services: a suitable case for treatment? *Public Manag. Rev.* 18 (5), 639–653.
- Secchi, P., 2018. On the role of statistics in the era of big data: A call for a debate. *Statist. Probab. Lett.* 136, 10–14. Special Issue on “The role of Statistics in the era of Big Data”.
- Van Vaught, F., 1999. Innovative universities. *Tertiary Educ. Manag.* 5, 347–354.

¹ See also Section 4.3.