If you measure it they will score: An assessment of international eGovernment benchmarking

Davy Janssen*, Sabine Rotthier and Kris Snijkers

Universiteit Antwerpen, Politieke en Sociale Wetenschappen, Steunpunt Bestuurlijke Organisatie Vlaanderen, Korte Sint-Annastraat 6, 2000 Antwerpen, Belgium
Tel.: +32 3 204 10 52; Fax: +32 3 204 10 80; E-mail: davy.janssen@ua.ac.be

Abstract. In this article we expound on the possible effects of eGovernment evaluation studies on countries’ eGovernment policy. Performance measurement and benchmarking are two well-known tools for policy evaluation and feedback. Performance information has a supply side as well as a demand side, which do not necessarily match. In the eGovernment sphere we notice a large supply of performance information. We ask ourselves to which degree this supply meets an actual demand from governments. We discuss the results of a thorough analysis of 18 international comparative eGovernment studies and warn for an unreflective feedback of these studies’ results in countries’ eGovernment policies.

Keywords: E-government, policy evaluation, performance measurement, benchmarking, indicators

1. Introduction

eGovernment has passed the agenda-setting phase of policy making. Programs and strategies have been discussed and the first projects have been set up, allowing the concept to get a foothold in government practice. The basic question for policy makers has consequently shifted from ‘what should we do’ to ‘how are we doing (compared to others)’. In an attempt to come to terms with this second question, government administrations, as well as international organisations and consultancy firms, have been busy designing international eGovernment benchmarking studies. The results of these studies, especially those that include country rankings, receive ample media attention and thus contribute to the general view of citizens on their country’s eGovernment. Headlines such as ‘country X trails behind for eGovernment’, or ‘country Y taking lead in eGovernment hit list’, can be found in all benchmarked countries. Strikingly, the boom in the amount of comparative studies produced in the years 2000–2001 has more than often resulted in a country scoring the high marks in one study, and ending at the bottom in another. In this article, questions concerning the validity of the measurements will be tackled. The result of a legitimate question of the Flemish government on how to interpret the results of the myriads of eGovernment rankings, an analysis of eighteen of those studies was made. The aim was to analyse what (mostly under the heading of eGovernment) was actually measured, in order to assist policy makers in making

*Corresponding author.
a well-informed judgement on the relevance of benchmarking results. As these studies are some of the main determinants of public opinion on eGovernment, as well as important drivers of eGovernment strategy, we believe a thorough analysis of what is being measured is crucial for the further development of eGovernment policy reform.

This article is the result of so-called ‘policy-relevant’ research, funded by the Flemish government [3]. Eighteen international benchmarking studies on eGovernment (and/or information society) have been selected and analysed. Studies were selected on the basis of their relevance for the research questions. On the basis of their experience with a recently finished comparative study on eGovernment implementation in seven OECD countries [4] the three researchers agreed on the selection of studies on the basis of their apparent relevance for policy makers. The main research question is aimed at explaining the often considerable variance in the results of a country’s eGovernment, as measured at $t_1$ in different studies. In the first section of this article we focus on the role of policy evaluation in the development of eGovernment policy, as well as on the impact of performance measurement and benchmarking. The second section deals with the variables that account for differences in the results of existing studies. We discuss differences in focus and in the eGovernment concepts used and look at the issue of comparability across time and space. A large part of the second section is reserved for a discussion of the nature of the benchmarking indicators that are used in existing studies. In a final section, we advise governments to evaluate the results of these studies carefully, instead of taking direct cues from them when devising an eGovernment strategy.

2. Performance measurement as an instrument of eGovernment evaluation

eGovernment stopped being a novel concept or development some years ago. Throughout the world, governments have started developing and implementing eGovernment initiatives and projects. eGovernment has passed the different phases of the policy cycle [6] for a first time: from agenda setting and problem definition to policy preparation, policy making, policy implementation and finally policy evaluation and feedback. These two final phases of evaluation and feedback are the focus of this article.

Policy evaluation as a phase in the policy cycle in public management has been thoroughly professionalized over the last decade. Performance measurement and benchmarking are some of the instruments that are used, often with different purposes. The Dutch report *Handreiking: prestatievergelijking binnen de openbare sector* mentions five of them:

- learn from and with each other;
- offer transparency;
- offer accountability;
- support external supervision;
- evaluate performances.

In the domain of eGovernment policy the attention for policy evaluation has been considerable and has mostly resulted in benchmarking exercises and international country rankings. The serious financial implications of implementing eGovernment policy, coupled with the uncertainty of (the timing) of benefits, have lead countries to legitimize their eGovernment spending by pointing at international studies that place them high on the eGovernment ladder or urge them to climb up a couple of steps.

Performance measurement can be seen as a cyclical process with a supply and a demand side [1]. On the supply side, the focus is on the supplier of performance information and the way in which information is generated. Important questions are: what occasioned the evaluation or measurement?
What is it meant to achieve? What are its targets? The occasion can be the need to learn or the need to account for investments. Other reasons can be thought of, though. It is conceivable, for example, that a commercial organization would offer performance information concerning the eGovernment policy of governments for commercial considerations. The presentation of country rankings could lead to increased pressure on governments to take immediate action, maybe in the form of private consultancy services.

We now arrive at the demand side, where performance information influences organization’s strategic choices. If some type of performance information – a specific mix of indicators- is recognized as the established criterion for the measurement of eGovernment success, governments might want to change their policy to make sure they score well on these indicators in a next evaluation round. If the performance information that is supplied corresponds with a government’s demand for evaluation information all is well. If there is no match between what is asked and what is supplied, strategic policy changes might not realize an authentic and qualitative eGovernment. In the field of eGovernment performance information, which we are looking at here, it should be noted that the majority of evaluation studies under consideration are developed and supplied by commercial organizations. The rest is either offered by an international organization or a national government. In the next section we will have a more detailed look at these studies

3. An analysis of several international comparative eGovernment evaluation studies

3.1. The focus of the studies

The studies could easily be divided into four groups with a clear difference in focus (see Fig. 1). A first cluster of studies can be called the supply-oriented eGovernment measurements. Half of the studies (1–8) inventoried fall into this group. The focus is on the supply of eGovernment applications, and the success of a country’s eGovernment is measured by counting the amount of visible applications. A second group of studies (9–11) takes an opposite approach and evaluates the demand side of eGovernment. These are studies that see eGovernment success in terms of actual levels of usage (‘take-up’ levels) or levels of customer satisfaction with online services. A third group of studies (12–16) has the Information Society as a focus. These studies benchmark countries, not only in terms of eGovernment efforts, but also according to their broader policy for (creating enabling conditions for) the Information Society and/or

---

1 An overview of the studies can be found at the end of the paper.
Table 1

<table>
<thead>
<tr>
<th>Supply</th>
<th>Information society</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) online service delivery</td>
<td>(12) Information Society: infrastructure + usage</td>
</tr>
<tr>
<td>(2) online service delivery + infrastructure component + human development component</td>
<td>(13) eEconomy: use of ICT to advance social and/or economic development</td>
</tr>
<tr>
<td>(3) online service delivery + eParticipation component</td>
<td>(14) Networked Readiness: ICT environment</td>
</tr>
<tr>
<td>(4) online service delivery + back-office (qualitative study, indicators)</td>
<td>(15) ICT test: ICT environment</td>
</tr>
<tr>
<td>(5) online service delivery</td>
<td>(16) eReadiness: ICT environment</td>
</tr>
<tr>
<td>(6) eGov = ‘when the public sector digitises its processes and interactions’</td>
<td></td>
</tr>
<tr>
<td>(7) online service delivery + back-office (qualitative study, indicators)</td>
<td></td>
</tr>
<tr>
<td>(8) online service delivery</td>
<td></td>
</tr>
<tr>
<td>Demand</td>
<td>EGOV Indicators</td>
</tr>
<tr>
<td>(9) service delivery + participation</td>
<td>(17) indicators to measure ICT environment</td>
</tr>
<tr>
<td>(10) service delivery + participation</td>
<td>(18) indicators to measure production and use of ICT</td>
</tr>
<tr>
<td>(11) service delivery</td>
<td></td>
</tr>
</tbody>
</table>

the Knowledge Economy. A fourth and final cluster (17,18) contains two meta-benchmarking studies in which criteria and indicators for the measurement of eGovernment are sought and proposed.

Although the segmentation of studies into four groups provides for a first categorization, differences in country rankings cannot be solely attributed to differences in focus, for even within each category country rankings vary considerably. Rankings should not necessarily be interpreted as ‘good-better-best’ qualifications. The aim of a benchmark is not always to see who has the ‘best’ eGovernment. Other rationales for eGovernment benchmarking include:

– finding out if lessons can be learned from other country’s eGovernment policies;
– measuring eGovernment progress compared to other countries;
– identifying and learning from best practices in other countries;
– discovering global trends in eGovernment;
– measuring of underlying eGovernment concepts to identify points of leverage.

This diversity of underlying goals and focus has its effect on the approach and the outcome of each study. One of the issues that has to be decided from the outset is the construction of a definition of eGovernment

3.2. The concept of eGovernment

A crucial step in the activity of measuring eGovernment is a clear demarcation of the concept itself. EGovernment definitions abound, so it is crucial to realise that country rankings from different benchmarking studies are probably based on different definitions of what is being measured. It is not our intention to formulate yet another definition of eGovernment. We merely give a descriptive account of how eGovernment is defined in the studies themselves. These definitions can be placed on a continuum with the following extremes: ‘service delivery on the internet’ (narrow definition) and ‘the use of ICT in the public sector’ (broad definition). Table 1 shows that the supply studies indeed mostly define eGovernment quite narrowly as online service delivery. The demand studies mostly take a somewhat broader perspective, often including aspects of eParticipation. The Information Society studies do not measure eGovernment in a narrow sense, but in some way look at the enabling environment for ICT, thereby
Table 2
Number of countries compared in benchmarking studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Number of countries compared</th>
</tr>
</thead>
<tbody>
<tr>
<td>Europe’s readiness for eGovernment.</td>
<td>15</td>
</tr>
<tr>
<td>Global eGovernment survey</td>
<td>196</td>
</tr>
<tr>
<td>Balanced eGovernment</td>
<td>Separate cases from Canada, USA and Europe</td>
</tr>
<tr>
<td>The current state of play. Australia’s scorecard.</td>
<td>14</td>
</tr>
<tr>
<td>Benchmarking E-government: A Global Perspective. Assessing the</td>
<td>190</td>
</tr>
<tr>
<td>Progress of the UN Member States</td>
<td></td>
</tr>
<tr>
<td>Online availability of Public Services: How does Europe progress?</td>
<td>17</td>
</tr>
<tr>
<td>eGovernment leadership: engaging the costumer</td>
<td>22</td>
</tr>
<tr>
<td>International ICT benchmark 2002.</td>
<td>6</td>
</tr>
<tr>
<td>International e-Economy benchmarking</td>
<td>9</td>
</tr>
<tr>
<td>The networked readiness of nations.</td>
<td>83</td>
</tr>
<tr>
<td>The 2003 e-readiness rankings.</td>
<td>60</td>
</tr>
<tr>
<td>Government Online: an international perspective.</td>
<td>31</td>
</tr>
<tr>
<td>E-government: de vraagkant aan bod.</td>
<td>n.a. (intra country demand)</td>
</tr>
<tr>
<td>Burgers aan het woord. Oordelen en klachten over de elektronische</td>
<td>n.a. (intra country demand)</td>
</tr>
<tr>
<td>overheid.</td>
<td></td>
</tr>
<tr>
<td>Eindrapport van een haalbaarheidsstudie ICT-monitor Vlaanderen.</td>
<td>n.a. (metastudy)</td>
</tr>
<tr>
<td>Benchmarking Ireland in the Information Society.</td>
<td>n.a. (metastudy)</td>
</tr>
</tbody>
</table>

often including eGovernment as one aspect of a broader scale of policy measures. The two eGovernment indicator studies, finally, are looking at indicators to monitor broader aspects of ICT development.

Studies that limit their view to online service delivery will obviously take into account different indicators than studies that include aspects of process change, back-office development and eParticipation. It seems that the inclination to limit the definition of eGovernment might partly be explained by the difficulties of gathering the information necessary for a broader conception of eGovernment. The following quote from Yankovitch and Kazner, cited in the benchmarking study Europe’s Readiness for eGovernment, warns of the dangers of this approach:

*Measuring and targets: four easy steps to disaster*

“The first step is to measure whatever can be easily measured. That is OK as far as it goes. The second step is to disregard that which can’t be measured, or give it an arbitrary quantitative value. This is artificial and misleading. The third step is to presume that which cannot be measured easily really isn’t very important. This is blindness. The fourth step is to say that what can’t easily be measured really doesn’t exist. This is suicide.”

Concerning data gathering strategies, it has to be said that most of the studies under consideration use existing, secondary, sources such as studies, national statistics, country reports, website analyses etc. Again, it seems that – besides the choice of a definition of eGovernment – also the choice of eGovernment indicators is often based on the information sources that are easily available.

3.3. Comparing across time and space

Most of the studies under consideration were commissioned out of a concern for comparisons, be it a comparison with one’s own position at a previous point in time or be it a comparison with another country or organisation. The studies under consideration are mostly cross-country comparisons that also have a temporal character because of the recurrent (yearly or bi-yearly) replications of the research. A view on Table 2 quickly shows the great variety in the number of countries included. Apart from the two meta-studies and the two demand-studies, all studies compare countries, with a pool of countries varying from 6 to 196.
3.4. Indicators for the measurement of eGovernment

For each of the studies under consideration an inventory of the indicators used was made. A global comparison of indicators led to the following categorisation of indicators:

1. Input indicators;
2. Output indicators;
3. Usage/Intensity indicators;
4. Impact/Effect indicators;
5. Environmental/Readiness indicators.

Input indicators try to measure the resources countries have invested in eGovernment. Output indicators do not measure financial resources but instead measure the amount of eGovernment applications realised. Usage indicators do not measure the amount of applications but their actual usage by citizens/businesses. Impact indicators then, try to measure the impact eGovernment has had, for example concerning changes in processing time or waiting time. Finally, environmental indicators try to assess the degree in which a country is 'ready' for the Information Society and its consequences.

3.4.1. Input indicators

Examples of input indicators
- Amount of financial resources devoted to eGovernment. Absolute figures, per capita figures.
- IT/e-Government spending as % of GDP.
- Amount of resources devoted to Research and Development.
- Amount of public resources devoted to internet infrastructure.

In the studies under consideration input indicators seldom get a lot of attention. Most studies limit themselves to a statistic of public IT spending, per capita or as a percentage of GDP. It is often not quite clear how these statistics emerge: how can one separate IT and eGovernment spending? How can one take into account countries where most IT spending is done in decentralised governments? How can one compare accounting systems of countries that deal differently with IT investments (or instead see them as costs)? One can only conclude that if a statistic is found, it is never entirely comparable with statistics found in other countries.

3.4.2. Output indicators

Examples of output indicators
- Number of online services for citizens;
- Number of online services for businesses;
- Percentage of government departments that have a website;
- Percentage of government websites that offer electronic services.

Studies that make use of a broad set of output indicators are mostly those with a limited definition of eGovernment as online service delivery. The indicators used try to measure the online presence and complexity of services. Complexity is often measured with the categories information, interaction, transaction and integration. Electronic service delivery indeed is one of the most salient features of eGovernment, so the output indicators are in no way unimportant. There is a danger though that governments that base their strategy on studies that only include output indicators tend to forget that eGovernment is more than online service delivery. When governments try to score on those studies
they can often do so by ‘digitalizing’ as many existing services as possible, thereby neglecting the more fundamental process of redefining service delivery in an online environment: you might be better off with less but better services. Pro-active service delivery and so called zero-stop government might be ingenious ways of approaching government in the information age, they are not valued by output indicators as they are used in the studies under consideration. A country that has a nice website where citizens can apply for some document online gets higher scores than a country that has improved its back-office and was thereby able to abolish the document (and the need for citizens to apply for it).

3.4.3. Usage indicators

Examples of usage indicators

- Number of individuals that have made use of electronic services offered;
- Number of businesses that have made use of electronic services offered;
- Percentage of citizens that has visited government websites to search for information;
- Number of businesses that have made payments online;
- Percentage of internet traffic that pertains to electronic service delivery.

Usage indicators try to measure the actual usage or ‘take-up’ of electronic services offered. In more recent studies, there seems to be an acceptance of the critique on output indicators. The main critique concerns the fact that countries get good grades for making lots of applications but that it does not matter if these applications are actually used by citizens. This is being corrected more and more by the use of usage indicators and by weighing them together with output indicators. This seems to make sense as the result is an evaluation of both the supply of and the demand for eGovernment in a country.

The usage indicators furthermore, provide for a good monitoring instrument for governments to evaluate the success of different applications and make corresponding strategy decisions. To arrive at a nuanced view of usage, there are often indicators for information seeking, information provision, and transactions.

3.4.4. Impact indicators

Examples of impact indicators

- reduction of waiting time at government counter $x$ by $y$ %;
- decrease in case processing time at government organisation $x$ by $y$ %;
- citizen/business satisfaction levels concerning eGovernment;
- survey-type questions, e.g.: ‘do you feel more positive to your government, now that you can contact it by email?’ ‘has your government become more efficient, now that you can perform services online?’.

The use of usage indicators described above already resulted in an overview of actual usage patterns per online service. Impact indicators go even further down the demand side and are used in studies that measure end user satisfaction, but also in studies that evaluate government organisation’s efforts. They try to establish some form of impact, be it citizens that are ‘happier’ or waiting lists that are shorter because of the introduction of eGovernment. Only a few studies deal with these kinds of indicators. They are of course also the hardest to operationalise and require primary data gathering: interviews with citizens, overall evaluations of organisation’s efforts. They do measure in a much more direct sense than usage indicators (which represent ‘consumer power’) the actual satisfaction of end users, or more generally the way that things have been improved because of eGovernment.
3.4.5. Environmental indicators

Examples of environmental indicators

- ICT penetration rates (pc, internet, mobile phone) private households, work, schools;
- Indicator that measures ‘fear of invasion of privacy’;
- Online shopping rates as an indicator of trust in online environments;
- Indicator that measures ‘quality of legislation concerning the information society’;
- Telephone tariffs, GSM tariffs, Internet access tariffs.

The environmental, or ‘readiness’-indicators do not measure eGovernment as such, but instead measure some of the preconditions of a successful eGovernment. They are indicators of the eSociety that is the surrounding environment of eGovernment, and mostly have to do with ICT infrastructure, ICT skills, trust in ICT and the legal environment. ICT infrastructure is one of the basic requirements of online government and can be measured by indicators such as internet penetration rates, broadband penetration, internet access tariffs, amount of public access points, etc. ICT skills have to do with the way a country’s population is able to handle computers and ICT. A further categorisation here distinguishes ICT skills among citizens, businesses, and civil servants. Another indicators that is sometimes used here concerns the presence of scientific or academic institutions that excel in ICT knowledge.

A third group of environmental indicators indirectly measure trust on online environments by measuring the presence and success of eBusiness and eCommerce. A final group then, focuses on a country’s legal environment and assesses this in the light of the requirements of the information society, dealing for example with the juridical value of an eMail and with the issues of online identification, online safety and online privacy.

4. Conclusion

The evaluation of eGovernment has become a booming business, as testified by the numerous benchmarking studies offered by commercial organizations (often consultants), international organizations and national governments. The different motives and targets of these studies result in different approaches to performance measurement. This performance information is interpreted by governments that try to give a strategic direction to their eGovernment policies. Because of this, these studies are much more than mere theoretical exercises without any real policy consequences. A crucial effort in the interpretation of studies thus has to be made to determine what it is that is actually measured.

In 2003 we analysed 18 international eGovernment benchmarking studies. These studies differ in focus, in scope (definition of eGovernment) and in the type of measurement criteria used (input, output, usage, impact and environmental indicators). The output and environmental indicators are being used most extensively. Especially those studies that define eGovernment as online service delivery limit themselves to measuring output (the amount of eGovernment applications). It is these same studies that mostly include a competitive eGovernment country ranking. Countries that score the high marks are obviously relieved and will be able to ‘objectify’ and legitimate their success. Countries that do not score well basically have two options. They can either try to ignore the results of the study and hope that nobody will remind them of this, or they can aim at scoring better next time around. In this second option, the (poor) results of the study will be taken seriously and will guide future eGovernment policy decisions. Now, if the studies that get most attention are the ones that measure eGovernment output (as we have tried to show they did), and if policy decisions are made to score better on these output criteria, there is the danger that the policy as a whole will be overly focused on front-office realisations.
However, when the authentic aim for the demand of performance information is to learn from those who are really doing better, the focus will have to be on the processes (back-office integration, intra- and intergovernmental information sharing, database development etc.) instead of on the results (applications in the front office). The studies under consideration here do not offer this kind of information as the key to eGovernment success. We therefore conclude that there is a gap between the demand for and the supply of eGovernment performance information. Information is supplied, but it often focuses in a superficial way on the amount (and not the quality) of eGovernment information as the crucial criteria of success. Therefore, there does remain a third option for governments to respond to the results of benchmarking studies. Instead of ignoring the results or trying to conform policy to the criteria used, a sensible third option would be to evaluate the performance information offered. This might lead to the conclusion that it would not be very sensible to change policy in the direction advocated by the study. In this way, eGovernment policies that are carefully constructed from the bottom up can further flourish.

References

List of benchmarking studies
Supply studies

Demand studies

Information Society studies


eGov indicator studies