Benchmarking e-government

A comparative review of three international benchmarking studies

Lasse Berntzen
Faculty of Social Science
Vestfold University College
Tonsberg, Norway
lasse.berntzen@hive.no

Morten Goodwin Olsen
Tingtun AS
Grimstad, Norway
morten.g.olsen@tingtun.no

This paper makes a range of comparisons between e-government developments and performance worldwide. In order to make such comparisons, it is necessary to use a set of indicators. This paper examines the evolution of indicators used by three widely referenced international e-government studies, from the early days of e-government benchmarking until today. Some critical remarks related to the current state-of-the-art are given. The authors conclude that all three studies have their strengths and weaknesses, and propose automatic assessment of e-government services as a potential solution to some of the problems experienced by current benchmarking studies.

Keywords: e-government, benchmarking, indicators

I. INTRODUCTION

Electronic government is the use of technology to provide electronic services to citizens, businesses and organizations. Through such services, users can interact with government independent of time or physical location. For the government, provision of electronic services is effective, since most users serve themselves.

As electronic government has developed, benchmarking studies have been initiated to show how different countries compare to each other.

Benchmarking is a technique for comparing e-government performance, and is normally based on a set of indicators. Such indicators are used to calculate some kind of performance index. The performance index can be used to rank different governments or government agencies against each other.

A framework is a set of assumptions, concepts, values, and practices that constitutes a way of viewing reality. A framework may be used to derive a set of indicators suitable for assessment and evaluation. Figure 1 shows one popular framework for e-government service maturity [1].

The vertical axis represents complexity, while the horizontal axis represents level of integration. The boxes represent different levels of maturity: catalogue level, transaction level, vertical integration level and horizontal integration level. This framework can be used to assess the maturity of different e-government services.
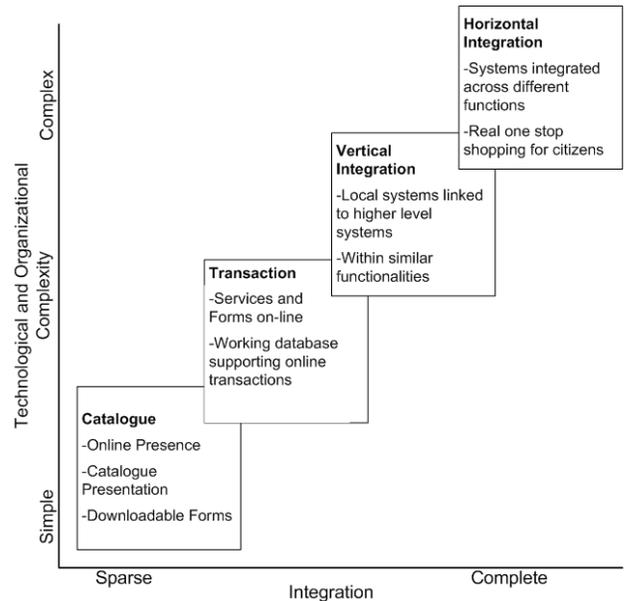

Figure 1. Example of e-government maturity framework.

Furthermore, benchmarking can help e-government development, by drawing attention to best practice elsewhere. The results of such benchmarking are often used in e-government strategy and planning processes.

II. RESEARCH SCOPE AND RELATED RESEARCH

The scope of this research is to compare three well-known, widely referenced, international e-government surveys and identify strengths and weaknesses of their respective methodologies. All three surveys have been repeated several times. As e-government services have matured, the methodologies have been expanded and refined.

First, this paper examines the evolution of the three benchmarking studies with respect to methodologies. Next, we compare the methodologies used, and discuss some problems with the current benchmarking efforts. Finally we introduce an ongoing research project aiming to overcome some of the challenges identified.

A. Related research

Other researchers have studied e-government benchmarking from other perspectives.

Janssen et al. [2] analyzed 18 e-government benchmarking studies. Most of these were one-time, local studies. The research team grouped indicators into five different categories:

- Input indicators, the resources countries have invested in e-government.
- Output indicators, the amount of e-government applications realized.
- Usage/Intensity indicators, the actual use of e-government applications by citizens.
- Impact/Effect indicators, the impact e-government has had on e.g., processing time or waiting time.
- Environmental/readiness indicators, how 'ready' a country is to implement e-government.

The authors concluded that output and environmental indicators were most common in these benchmarking studies. These observations are also valid for the three benchmarking studies analyzed in this paper.

Ojo et al. [3] examined the same three international benchmarking studies used in this paper, and identified a set of core indicators for assessing e-government readiness. Their research focused on e-readiness, the capability to implement e-government solutions, and proposed a modified set of indicators for this particular purpose.

Other researchers like Germanakos et al. [4] studied benchmarking of e-government services provided by cities.

III. MAJOR BENCHMARKING STUDIES

There are currently three widely published international benchmarking studies:

- Accenture, a global management consulting, technology and outsourcing company, has done yearly assessments of e-government efforts in 20+ countries since 2000 [5,6,7,8,9,10,11].
- Brown University (Prof. West and his research team), has done annual e-government assessments since 2001 [12,13,14,15,16,17,18,19].
- The United Nations Department of Economic and Social Affairs have assessed e-government readiness among its member nations since 2002 [20,21,22,23,24].

A. Accenture

Accenture has published its e-government surveys every year since 2000. Data collection is done by local Accenture employees in the countries surveyed. Although the majority are European, the survey also includes non-European countries.

As shown in table 1, the survey reports have changed title every year, to indicate the current state of e-government development and implementation in the surveyed countries.

Accenture has consistently included a description of methodology used, either as a separate chapter or an appendix. The last report, however, provides fewer details than earlier reports.

In 2001 [5] the ranking of countries was based on two indicator sets: *service maturity* and *delivery maturity*.

Service maturity was calculated from the number of services implemented and their level of completeness using a framework not very different from the one described in the introduction. The maturity levels used were: publish, interact and transact. The first level, publish, signifies that information is available online. The next level, interact, signifies that citizens are able to submit information online, e.g., application forms. The government does not need to respond electronically. The third level, transact, signifies that the government responds electronically to information submitted.

Delivery maturity covers selected delivery aspects, such as single point of entry, design by customer intentions, CRM techniques, portal capabilities etc.

In the overall e-government maturity index, service maturity was weighted 70%, while delivery maturity was weighted 30%. The resulting index was used to group countries into *innovative leaders*, *visionary followers*, *steady achievers* and *platform builders*.

Table 1 shows how the Accenture surveys targeted more services and more sectors each year until 2006.

In 2002 [6], *delivery maturity* was substituted by *customer relationship management* (CRM). Customer relationship management covers five aspects: Insight, interaction, organization performance, customer offerings and networks.

Service maturity was still weighted 70%, while CRM was weighted 30%. The resulting index was used to group countries into *innovative leaders*, *visionary challengers*, *emerging performers* and *platform builders*.

In 2004 [8] a new indicator, *support*, was included in the CRM to measure integration with other channels (e.g., call center) to assist citizens in finding information and complete transactions.

TABLE I.   ACCENTURE SURVEYS

Year	Title	Countries	Services	Sectors
2000	Implementing e-government: Rhetoric or Reality	N/A	N/A	N/A
2001	E-government leadership: Rhetoric vs. Reality	22	165	9
2002	e-government Leadership: Realizing the Vision	23	169	9
2003	e-government Leadership: Engaging the Customer	22	201	11
2004	e-government Leadership: High Performance, Maximum Value	22	206	12
2005	Leadership in Customer Service: New Expectations, New Experiences	22	177	12
2006	Leadership in Customer Service: Building Trust	-	-	.
2007	Leadership in Customer Service: Delivering on the Promise	22	N/A	N/A

The 2005 survey [9] introduced some major changes in methodology and measurements, by including more facets of service delivery. The index was built from two components, each weighted 50%. The first component, *service maturity*, was the same as in earlier surveys.

The second component, *customer service maturity*, measured how well governments addressed the four dimensions of customer service: *citizen-centered*, *multi-channel*, *cross-government service delivery* and *proactive communication* about the services towards users.

For the first time, citizens were asked their opinions as part of the data collection process. 400 citizens in each country (600 in U.S.A) were asked a series of 15 questions about their attitudes toward their countries' service delivery program, their use of different service channels and their interests in particular services. This practice has been continued in later surveys as shown in table II.

In 2006 Accenture decided to temporarily drop the ranking of individual countries [10]. Instead the study focused on those countries that have performed consistently well during the previous surveys, in order to give examples of best practice.

In addition, in-depth interviews were conducted with 46 high-ranking government executives. The research was also extended by a survey of 8600 citizens as explained above.

The 2007 survey [7] introduced a new indicator set, *citizen voice*, weighted 40%, to incorporate results from citizen interviews in the ranking. The weighting of the service maturity component was reduced to 10%. The reason for this reduction was explained in the following way [11]: "*This element of our rankings has decreased in importance over the years as e-government has become increasingly ubiquitous and less of a differentiator among countries*". The service maturity component included the same aspects as in 2005, and was weighted 50%. Interviews with high-ranking government executives were continued and used as input to the general discussion about e-government maturity.

TABLE II.   DATA COLLECTION THROUGH INTERVIEWS (ACCENTURE)

Year	Countries	Interviews
2004	12	5000
2005	22	9000
2006	21	8600
2007	-	-
2008	22	N/A

B. Brown University

Brown University has published a series of e-government surveys [12,13,14,15,16,17,18,19] since 2001. In 2001 the survey examined 2288 web sites in 198 countries.

The Brown University surveys examine a broad range of public web sites.

Among the sites analyzed are those of executive offices (such as a president, prime minister, ruler, party leader, or royalty), legislative offices (such as Congress, Parliament, or People's Assemblies), judicial offices (such as major national courts), Cabinet offices, and major agencies serving crucial functions of government, such as health, human services, taxation, education, interior, economic development, administration, natural resources, foreign affairs, foreign investment, transportation, military, tourism, and business regulation.

Each web site has been evaluated for the presence of 28 specific features dealing with information availability, service delivery, and public access. Features assessed included type of site, name of nation, region of the world, office phone number, office address, online publications, online database, external links to non-governmental sites, audio clips, video clips, non-native languages or foreign language translation, commercial advertising, user payments or fees, subject index, handicap access, privacy policy, security features, presence of online services, number of different services, links to a government services portal, digital signatures, credit card payments, email address, search capability, comment form or chat-room, broadcast of events, automatic email updates, and having an English version of the web site.

The survey only examined presence of features, with no efforts to measure maturity or depth of individual services.

Observations were done by native speaking researchers and in some cases by automatic translation tools.

Since 2001 the survey has been repeated, using the same research methodologies. Table III shows the number of services and countries assessed each year.

TABLE III.   BROWN UNIVERSITY SURVEYS

Year	No. of countries	Total no. of web sites
2001	198	2288
2002	198	1197
2003	198	2166
2004	198	1935
2005	198	1796
2006	198	1782
2007	198	1687
2008	198	N/A

In 2002, a simple e-mail responsiveness test was performed. After 2002, there have been some changes of features checked, like PDA accessibility.

C. UN Department of Economic and Social Affairs

The United Nations Department of Economic and Social Affairs (UNDESA) published its first e-government survey in 2002. From 2004 the title was changed to e-government

readiness report. In 2008 the title has reverted back to e-government survey. All reports give a detailed description of research methodology used.

The 2002 benchmarking study [20] was done as a collaborative effort between the UN Division of Public Economy and Public Affairs (UNPEPA) of the UN Department of Economic and Social Affairs (UNDESA) and The American Society for Public Administration (ASPA). The study examined government web sites of all 190 UN member states. At that time 169 member state governments showed web presence, while 84 member states had a national government web site.

Two methodologies were used. First, web sites were checked for content and services likely to be used by citizens. The results were then used to derive an index showing the sophistication of government web sites. Second, statistical information was collected on infrastructure and human capital. The resulting e-government index was used to rank the member states.

The study was primarily based on observation, but background information was also collected through on-site visits, interviews and questionnaires.

Countries were assessed using the following criteria:

- An official government web presence must exist.
- The type of service delivery available: basic, informational, interactive, or transactional.
- The presence of services in five critical sectors: education, health, labor/employment, welfare/social services and financial services.
- Use of single entry portals; and de facto portals (official national government web sites).
- To a lesser degree, fidelity to strategic plans, use of e-government teams.

The research is well documented, including the actual assessment form for government web sites and formulas for deriving the different indicators.

The 2003 e-government survey [21] assessed the 191 UN member states based on the same methodologies as in 2002. The Web Measure Index was revised and enhanced, and used a five stage model to assess sophistication of services. The five stages used were: *emerging presence*, *enhanced presence*, *interactive presence*, *transactional presence* and *networked presence*.

The web measure assessments are purely quantitative, and were based on a questionnaire that required researchers to assign a binary value to the indicator based on presence/absence of specific electronic facilities/services available.

The 2003 survey, however, also included a qualitative assessment of e-participation. A total of 21 public and participatory services and facilities were investigated across the 191 member states. A scale of 0-4 was used to assess the individual service.

The 2004 e-government survey [22] was based on the same methodology used in 2003.

The 2005 e-government survey [23] was based on the same methodologies as in 2004 with some enhancements.

The study included supplemental research on disability access of national web sites, using the tool WebXACT to measure compliance with current accessibility standards.

The 2008 e-government survey [24] was based on the same methodology as the 2003 and 2004 surveys. Some modifications were made to enhance the rigorousness of the methodology.

IV. COMPARISON AND DISCUSSION

A. Reproducibility

Reproducibility is one of the cornerstones of the scientific method, and refers to the ability of a test to be accurately reproduced, or replicated, by someone else working independently

Prerequisites for reproducibility are proper documentation of methodology and access to primary data used in deriving the results.

All three benchmarking studies include description of methodology used, but descriptions show variations in quality.

Accenture provides no detailed list of services investigated and their maturity scores. Therefore, the calculation of the indexes is not reproducible. The lack of detailed breakdown and assessment of services causes at least two problems. It is impossible to evaluate the relevance of the services selected, and it is also impossible to submit corrections. While the methodology has been described in detail in earlier reports, the last report does not include number of services examined, number of sectors included and number of citizens interviewed.

The Brown University survey describes categories of web sites searched for, and features examined. The list of web sites visited is not available. Again, it is not possible to reproduce results based on information published. Since this data is not available, it is not possible to submit corrections.

Recently, the UN launched a knowledge database (http://www.unpan.org/egovkb) to enable researchers and government officials to access the background information for the survey. UN also publishes a detailed description of their research methodology. By combining the methodology with data from the knowledge database, it is possible to reproduce results.

B. Observation

All three studies are based on observations. Accenture is using local employees. The Brown University survey examines almost 200 different countries, with a large variety of languages. At least, in recent years the research team claims to have employed foreign language readers to evaluate foreign language web sites. No description is given of how the research team maintains consistency when a large number of evaluators examine rather small samples. The UN reports lists research team and translators used, which makes the data collection process more trustworthy.

C. *Qualitative assessments*

As a supplement to quantitative indicators, Accenture is also doing qualitative assessments based on knowledge of the individual countries, and such information is given in individual country sub-reports.   Accenture also started interviewing citizens in 2005. However, these results were not incorporated as an indicator set until 2007. Some of the qualitative remarks attack the problem of web content (what users see) vs. back office solutions (what users do not see).

D. *National scope*

All three surveys target national services only, excluding services provided by lower levels of government. Accenture specifically mentions that they do not penalize services delivered on lower levels.

E. *Other issues*

In contrast to the two other benchmarking studies, the UN surveys focus on additional aspects such as infrastructure and human capital, and also include indicators for assessment of citizen participation.

V.   CHALLENGES RELATED TO CURRENT INDICATORS

One problem with the benchmarking studies described above is that they are supply oriented. All surveys count services, and assess their sophistication. The services are not evaluated based on their usage or impact, as seen from the citizen point of view.   Only Accenture have used interviews to find out what citizens think of services. These results have only recently influenced the final index.

Another problem is that all three studies target electronic services at national level. In practice, many services are the responsibility of lower levels of government. Even if such services are removed from the analysis, this introduces a considerable source of error in the assessments. The later Accenture reports have discussed this as a problem.

All studies are primarily based on observation. Observation does not reveal what is behind the facade. A service may be poorly integrated with back-office systems, and still get a high score.   Another system with less functionality may be well integrated with back-office systems and get a low score. Assessment of e-government should certainly also take into consideration the level of integration. As earlier stated, Accenture does a qualitative assessment including such aspects in their descriptions of the individual countries.

Observation requires fluency in the languages used. A large number of observers may also cause problems, since observers may have different conceptions of what they observe.   In countries with more than one language, observations should include all languages. Accenture have solved this by using local staff in each country being studied. UN has at least some transparency regarding use of experts for local assessments.

A relevant question is what services to include. All three benchmarking studies use a checklist approach to look for existence of certain e-government services, but such checklists need to be transparent in order to evaluate relevance. A survey based on a checklist of possible services does not take into consideration the relevance of each service in relation to local circumstances, so a detailed breakdown of results should therefore be given.   UN is providing such breakdown through their recently launched knowledge base.

Observers may be unable to uncover "hidden" services. One example is when the government sends the web-address together with a user name and a password to its citizens to enable access to an application requiring some kind of authentication.

Another problem emerges from changes in research methodology. When services are added or removed from examination, the results are no longer consistent with earlier years.

The UN studies assess e-participation, but their assessment raises some important questions of what should be measured.   From the description of the research methodology, the UN examines the government web site for evidence of participatory applications, but is this the right place to look for participation? For countries based on indirect democracy, it would be more natural to see how citizens participate towards the parliament, not the government itself.

None of the three studies provide designated feedback channels to report on incorrect information.

VI.   THE FUTURE: AUTOMATED ASSESSMENT

Governments are introducing new e-government services every day and benchmarking is an important mechanism for keeping track of developments and a source for identifying best practices, but as the number of e-government services increases, data collection becomes more challenging. We have earlier discussed the problem of benchmarking studies only targeting national e-government services, but an expansion of benchmarking studies to include services provided at lower organizational levels, e.g., municipalities, using the traditional approach of observation, would be a very challenging and resource intensive task.

An ongoing project, eGovMon [25], co-funded by the Research Council of Norway, aims to partially automate assessment of e-government services. Performing parts of the evaluation automatically frees up resources and allows for a much wider range of web sites to be part of the evaluation. In practice, this approach allows web sites at all levels to be part of the evaluation, and additionally, it facilitates the presentation of results at more frequent intervals compared to traditional studies.

However, automated tests cannot evaluate with the same level of detail as manual evaluations.

Therefore, automated assessment will be supplemented by expert evaluations using a specialized tool for manual verification of collected information. The project will also include a feedback channel giving service provides the possibility to submit

The eGovMon project builds on experiences from a recently completed EU project to develop an online observatory for automatically collecting information on accessibility [26,27].

VII. CONCLUSIONS

This paper has discussed major e-government benchmarking studies. The three studies use different methodologies and indicators. The methodologies used have evolved over time. All benchmarking studies have become more complex, by collecting information from multiple sources and by expanding the number of indicators being used for assessment.

Still, there are several problems related to the use of indicators, and some important issues are not covered by these benchmarking studies, e.g., accessibility, transparency, efficiency and impact. Automated assessment may be one strategy to improve future assessments.

ACKNOWLEDGMENT

REFERENCES

[1] Layne, K. and J. Lee, Developing fully functional E-government: A four stage model, in Government Information Quarterly. 2001. p. 122-136.

[2] Janssen, D., S. Rotthier, and K. Snijkers. If you Measure it They will Score: An Assessment of International e-government Benchmarking. in 4th European Conference on e-government. 2004. Dublin, Ireland: Academic Conferences.

[3] Ojo, A., T. Janowski, and E. Estevez. Determining Progress Towards e-government: What are the Core Indicators? in 5th European Conference on e-government. 2005. Antwerp, Germany: Academic Conferences.

[4] Germanakos, P., E. Christodoulou, and G. Samaras. Towards the Definition of an e-government Benchmarking Status Methodology. in 6th European Conference on e-government. 2006. Marburg, Germany: Academic Conferences.

[5] Accenture, e-government Leadership: Rhetoric vs. Reality – Closing the Gap. 2001.

[6] Accenture, e-government Leadership - Realizing the Vision. 2002.

[7] Accenture, e-government Leadership: Engaging the Customer. 2003.

[8] [Accenture, e-government Leadership: High Performance, Maximum Value. 2004.

[9] Accenture, Leadership in Customer Service: New Expectations, New Experiences. 2005.

[10] Accenture, Leadership in Customer Service: Building the Trust. 2006.

[11] Accenture, Leadership in Customer Service: Delivering on the Promise. 2007.

[12] West, D. M., WMRC Global E-government Survey, October 2001, Taubman Center for Public Policy, Brown University, 2001

[13] [West, D. M., Global E-government 2002, Center for Public Policy, Brown University, 2002

[14] West, D. M., Global E-government 2003, Center for Public Policy, Brown University, 2003

[15] West, D. M., Global E-government 2004, Center for Public Policy, Brown University, 2004

[16] West, D. M., Global E-government 2005, Center for Public Policy, Brown University, 2005

[17] West, D. M., Global E-government 2006, Center for Public Policy, Brown University, 2006

[18] West, D. M., Global E-government 2007, Center for Public Policy, Brown University, 2007

[19] West, D. M., Improving Technology Utilization in Electronic Government around the World, Brookings, 2008

[20] United Nations, Benchmarking E-government: A Global Perspective. 2002, United Nations Division of Public Economics and Public Administration.

[21] United Nations, UN Global E-government Survey 2003. 2003.

[22] [United Nations, UN Global E-government Readiness Report 2004: Towards Access for Opportunity. 2004.

[23] United Nations, UN Global E-government Readiness Report 2005: From E-government to E-Inclusion. 2005.

[24] United Nations, UN Global E-government Readiness Report 2008: From E-government to Connected Governance. 2008.

[25] Berntzen, L., M. Snaprud, A. Sawicka, L. S. Flak, Towards Automatic Assessment of e-Government Services. in Expanding the Knowledge Economy : Issues, Applications, Case Studies, IOS Press, p. 223-229.

[26] Bicharra Garcia, A.C., C. Maciel, and F. Bicharra Pinto. A Quality Inspection Method to Evaluate E-government Sites. in Electronic Government, 4th International Conference. 2005. Copenhagen, Denmark: Springer Verlag.

[27] Choudrie, J., G. Ghinea, and V. Weerakkody, Evaluating Global e-government Sites: A View using Web Diagnostic Tools. Electronic Journal of e-government, 2004. 2(2): p. 105-114..