

Benchmarking eGovernment Quality – Whose Quality Are We Measuring?

Arild Jansen¹, Svein Ølnes²

¹Section for eGovernment, University of Oslo, Norway

²Western Norway Research Centre), Norway

arildj@jus.uio.no

sol@vestforsk.no

Abstract. This paper analyses the results of several years of benchmarking of public online services in Norway. We compare these data, which are showing significant differences in measured quality between small and larger municipalities, with results from a comprehensive survey measuring citizens' satisfaction with public services. Finding that these observed differences are not supported by the user survey, we have to ask: whose quality are we really measuring? Many evaluation systems rely on similar heuristic methods, e.g. the EU's eGovernment benchmark 2012 framework, while the Danish benchmarking system has a different approach. The paper argues for a multi-dimensional approach to evaluation of public websites and gives some suggestions for this.

Keywords: Quality, Heuristics, Benchmarking, Evaluation, eGovernment

1 Introduction

Given the high priority of digital public service provision, their quality in general is important. Evaluating digital services is thus of crucial importance to assure that the overall goals have been reached. Several benchmarking systems have been applied; many of them are of great importance. Most notably, the EU's eGovernment Benchmark 2012 [Ca12] is an important tool for evaluating the effect of EU initiatives and the different national ICT policies. However, critics have also argued that the EU benchmarking system needs a thorough rework [Gr10]. Schellong [Sc09] in his evaluation of the EU's eGovernment benchmarking system, points to several troubling issues and suggests a number of improvements.

Norway is among the countries with a long history of evaluating public websites, which has been carried out annually since then 2001. The Norwegian framework is built on a set of general indicators common for all websites, using expert based heuristics, and there have been relatively small changes during the last 10 years. This framework is partly in line with the EU benchmarking system, but different from e.g. the Danish benchmarking system where domain specific indicators are used [Vi02].

Bannister [Ba07] points to several problems with benchmarking in general: Firstly, any ranking system needs a final single *scale* and the ability to compute a *score* on that scale. Secondly, if what one is ranking a concept or a mental state (e.g. attitude) rather than something concrete, it becomes necessary to use psychometric type tools. Bannister emphasizes that the answers to these question will vary with the context.

This paper compares the results from the Norwegian evaluations of public websites with a comprehensive citizen survey and finds that there are diverging conceptions of quality. Such findings provides arguments for asking whether the quality criteria's formulated in the Norwegian evaluations, and partly included in the EU eGovernment benchmarking, reflect a different conception of quality than the one that is perceived by the end users. Heuristic based evaluation systems for public websites are necessary to raise the awareness of common design principles, but may not be sufficient to assure quality to the end users. On the other hand, Denmark uses a combination of automated and manually assessed indicators [Dig12], and in addition a short user survey

Our main research question is accordingly: *What quality is measured in the various benchmarking systems and for whom are we evaluating?*

Furthermore, developing benchmarking systems must be according to some overall objective. A sub-question is thus: *Are the objectives of evaluations compatible, and are the methods used adequate?*

Brief on our research approach: This work is based on an inductive approach, in which we use data from various benchmarks and evaluations: i) the Norwegian evaluation of public web sites from 2007 to 2011 [Dif12], ii) a survey among users of Norwegian public services from 2009/2010 [Dif10]. Furthermore, we have carried out a limited literature review based on relevant literature from the eGovernment Reference Library (EGRL). The insights gained from these results are used as reference when analysing the criteria used the Danish "Bedst på nettet" [Dig12] and the EU's benchmarking system [Ca12]. Through these comparisons, we suggest some guidelines for developing more adequate frameworks for measuring quality of information and public e-services. Our focus is on the quality of available online services, and not on the broader issue such as benchmarking capabilities of eGovernment across nations.

The structure of the paper is as follows. In chapter 2 we discuss the quality concept as well as different heuristic models for evaluations. In chapter 3 we analyse the results from the Norwegian quality evaluations, fuelling our question of whose quality we are measuring. Chapter 4 conclude by suggesting a broader approach to quality assessment of eGovernment online services.

2 Can We Measure Quality?

Few words have been more used and misused than *quality*. Public as well as private companies emphasize the importance of increased quality of services, but very often they fail to define what quality means. Some definitions can illustrate this:

- a system's capability to satisfy needs, expectations and requests [DM93]
- the proportion between expected and experienced yield of a system [BO95]:

These definitions emphasize different aspects of quality. The first one looks at measuring the difference between what is specified and what is measured or registered through 'objective' criteria while the other is based on experienced properties, that is, '*a subjective evaluation from the individual user concerned*'.

ISO-8402 of the ISO-9000 standard is guidelines for quality management and quality assurance and uses this definition of quality: "*Quality is defined as the total sum of properties a unit carries and that concerns its ability to satisfy explicitly expressed or implied needs*".

Dahlbom and Mathiassen [DM93] held that quality is then often measured and analysed into a number of factors, such as correctness, reliability, efficiency, integrity, usability, flexibility, interoperability, portability etc. Each of these factors as more or less explicitly defined in order to give a precise meaning of the term used. However some of these factors may contradict each other [9, p140].

In general, we may use two different forms of evaluation, one based on metrics, measuring objective measurable attributes (criteria), and one based on perceived quality. This is rather depending on competence and individual judgement, which may also include aesthetic factor (e.g. how a webpage looks), and symbolic aspects as part of the overall quality; Symbolism has to do with its social use, e.g. as a mean for communicating the culture of an organisation. In this respect, understanding quality must also include cultural aspects and even a political dimension (interests and power), which is about why we evaluate and how the results will be used. Accordingly, quality criteria must conform to the overall goals, and furthermore the needs of important user groups and other stakeholders.

2.1 Heuristic Models for Measuring Quality

Measuring quality of websites does often rely on using heuristic methods, which have become the most used approach for expert-oriented evaluations [DJL10]. Kahneman [Ka11] offers this definition of heuristics: "*A simple procedure that helps find adequate, though often imperfect, answers to difficult questions*". A more practical definition concerning the use in website evaluations is "*all the sets of process guides, principles, criteria, tips and tricks, and guidelines that are available to support web designer*" [JG00]. It should be added that in the on-going process of developing and broadening the scope of public website evaluation the focus has shifted from almost solely assessing usability issues to add more and more governmental issues such as the level of digital service provisions.

The heuristic method for evaluating website quality was developed by Nielsen and Molich in the early 1990s [Ni94]. In a heuristic evaluation, one or more experts check a given website using a predefined set of evaluation criteria, the heuristics [DJL10]. The heuristics developed by Nielsen and Molich were primarily aimed at evaluating user interfaces, and consist of ten basic principles derived from studies of problems

found in dealing with user interfaces. However, although the heuristic model involving an expert evaluation is much used, we do not know very much about how heuristics function [DJL10]. Donker-Kuijer et. al. [DJL10] analysed five e-Government heuristics with respect to a) context of use, (b) the information they cover, (c) their validity, and (d) their presentation format (ibid). Their conclusions were that the government heuristics are very complex documents difficult for (end) users to read and comprehend. Also information about the foundations of the heuristics is often missing making it difficult to judge the quality of the heuristics. Compliance with the heuristics is also in many cases difficult to check because it requires an extensive amount of automated and (manual) expert evaluation techniques. All in all the authors seriously doubt if the examined heuristics aid the experts in their work. De Jong and van der Geest [JG00] thus distinguish between these four foundations for heuristics:

1. Standards-based heuristics
2. Theory-based heuristics
3. Research-based heuristics
4. Practitioners' heuristics

Heuristic models have their weaknesses and limitations, but for large scale screening of website quality there are hardly any alternatives. For measuring the usability aspects there are methods like user testing. In his book *Usability Engineering* [Ni93] Jakob Nielsen discusses the usability of a system and refers to concepts like user friendliness, usability and usefulness which can all be viewed as different dimensions of system acceptability. One of his main arguments is that different categories of users, different user situations and individually different preferences make usability testing difficult. He points to three main dimensions:

- experience with computers and relevant computer systems in general
- experience with the actual system (novice – expert)
- knowledge and competence in the actual domain where the system is used

The heuristic methods are especially suited for evaluating usability. Nielsen [Ni93] has formulated 10 heuristic principles for usability, derived from the usability properties listed above. Nielsen points to problems with user testing in general where the results will differ because of different user categories as mentioned above. We agree with Nielsen that user tests should take into account all the three dimensions.

2.2 Using Heuristics when Measuring Quality of Public Services

Norway has been evaluating public websites since 2001 and the number has risen from around 500 to more than 700, of which approximately 430 are municipality sites and the rest various governmental agencies' websites. This work is based on the following definition of quality: "*The quality of websites in this project is defined as that public information and services on the Internet must meet a predefined standard or level that can satisfy some central user needs*".[JO04]

The central guidelines for the development of quality assessment indicators are:

- the Governmental ICT policy [Mi12]
- Relevant laws, regulations and principles for public administration
- Widely accepted standards and guidelines on the web (= heuristics),

formulated by the W3C, especially their recommendations for web accessibility, expressed through the recent Web Content Accessibility 2.0 Guidelines (WCAG) [W08].

The overall structure of quality indicators has been based on these three dimensions: i) Accessibility, ii) Usability, iii) Useful services.

Similarly, the Danish benchmarking system (Bedst paa nettet) started out in 2001 as an expert evaluation based on a mainly heuristic set of criteria. Gradually it has shifted the focus to the citizens with the introduction of user surveys, after having tried both sector specific criteria sets and self-evaluations. The 2012 benchmarking consist of an accessibility evaluation, *Webtjek*, which is a combination of automated and manually assessed indicators [Dig12]. In addition to the screening it also consists of a user survey of eight questions which all municipalities and governmental agencies must carry out in order to participate in the contest.

EU's benchmarking system has been a cornerstone of the Commission's "open method of coordination" since the Lisbon meeting introducing the first eEurope plan [Sc09]. It has been regarded a success due to its influence on the eGovernment progress and politics in the EU. However, it has also been criticized for being too focused on the supply side of eGovernment and not really user oriented. The "Insight Report" [Se12] presents the findings of the 2012 eGovernment survey. The framework for this survey is rather complex and include three main areas: i) A demand-side citizen view of public services, ii) Three life-event assessments of very relevant customer experiences, iii) Assessment of five key technology enablers, the foundations on which services can be delivered in a more consistent manner. In each country two mystery shoppers assessed these life events against seven criteria.

In the demand-side citizen survey 28,000 internet-using citizens were asked 27 questions about 19 common citizen services. These questions includes eGovernment use and channel preferences, furthermore barriers, eGovernment satisfaction: and finally fulfillment and benefits: reasons for using eGovernment services and indicating whether governments are able to meet expectations citizens do have. It cannot be questioned that the results from this survey is very interesting and useful for the work to improve improving online services, both oat a national and an international level. However, whether these data should be used for comparison between countries is questionable, being based on subjective and (volatile) criteria.

2.3 Assessment and Benchmarking of e-Government Initiatives

Benchmarking of governmental websites and national e-government initiatives has been conducted in a number of years. There are several well-established surveys on e-government e.g. CapGemini [19], United Nations [UN08], and West [UN08]. These surveys employ different assessment models for e-readiness, digital divide and other relevant factors, leading to varying conclusions on the global state of e-government. The grounds for these efforts are well illustrated by a statement from the EU report [We08]:

"The ministerial declaration on the eGovernment conference, together with

benchmarking survey should give political momentum to the development of online public services and to the identification of the needs for these services at pan-European level. This will have to be complemented by a focus on back-office re-organization, the creation electronic marketplaces for public procurement and investment in new equipment in administration”.

Although eGovernment benchmarks are derived from political goals and decisions, their results generate in many cases political discussions and may lead to new political decisions. In itself it is not a problem; that is what benchmarking systems are developed for, but if the results are not comparable, it can be a serious problem. Schellong [Sc09] shows that the rankings in three major eGovernment benchmarks (EUeGovBe, UN and Brown/Brookings) differ rather remarkable, almost dramatically. This shows that comparisons of benchmarks e.g. between countries is not only difficult but often downright wrong, as Schellong [Sc09] and Bannister [Ba07] warn.

3 Results from the Norwegian Quality Evaluations

Below we present the results from two different studies; (a) the results from the expert evaluation of public websites 2007 to 2011 [Dif12], and (b) a user survey targeted at the users of public websites [Dif10].

3.1 Expert Evaluations of Websites 2007 – 2011

The set of indicators used for expert evaluation of public websites in Norway has only modest changes from 2007 to 2011. The indicators are divided into three subsets, as listed in the following table. The table also show the number of indicators and the maximum score in points per subset.

Table 1: Indicator set in 2007 and 2011

Subset	2007			2011		
	No. of indicators	Max. points		No. of indicators	Max. points	
Accessibility	11	27		11	28	
Usability	14	37		12	30	
Useful services	7	28		10	35	
Total	32	92		33	93	

The usability indicators are mostly based on the heuristic principles for usability formulated by Nielsen [Ni93]. The last set of indicators, useful services, looks at the service provision from a user’s point of view.

The only weighting in the set is in the maximum number of points for each indicator. The table above shows that useful services have been given priority over accessibility and especially usability over the period of five years. This reflects the strategy at the national level where provision of digital services to the citizens has been given an

increasingly higher priority. The weight on accessibility has remained almost constant and only usability has lost weight compared to the other subsets.

The following table gives the results for the 2007 and the 2011 evaluations and the change in percentage points for this period. The municipalities are divided in three groups after their population: up to 5.000, from 5–20.000, and more than 20.000 people. [LA11]. The labels ‘medium’ and ‘large’ must be seen with respect to the general size of Norwegian municipalities. Most of them are small by any measures.

Table 2: Total score for municipality websites after size (percentage of max. points)

Municipality group	2007	2011	Change (% points)
Small municipalities (< 5 000) (N = 226 in 2007 and 229 in 2011)	46.6	58.4	11.8
Med. sized munic. (5 000 – 20 000) (N = 148 in 2007 and 2011)	50.7	66.1	15.4
Large municipalities (> 20 000) (N = 52 both in 2007 and 2011)	58.4	72.0	13.6
Average, all municipalities	49.5	62.7	13.2

Significant differences both between municipality sizes and years (95 % confidence interval)

The results in the table above show that the large municipalities scored better than small and medium sized municipalities and they have also had the greatest improvement the last five years in terms of quality of websites as measured with this evaluation system. The difference in quality of websites between small, medium sized, and large municipalities is increasing. Part of this difference can probably be attributed to the increasing weight put on useful services, which in 2011 counted for 37.6 % of the maximum score in 2011 compared to 30.4 % of the 2007 indicator set.

3.2 Survey among Users of Public Services

Norway also undertakes a comprehensive user survey of public services with 3-4 years intervals. The last published survey is from 2009/2010 and comprises questions from a range of governmental bodies and municipalities [Dif10]. It is a large survey with questions from many sectors and services. The survey consisted of two main parts, where the first part was sent to some 30.000 citizens above 18 years and the response rate was 42 %. The second part of the service was sent to those of the respondents of the first part that had some experience with any of the chosen public services during the latest year. Part two of the survey was sent to 11.135 of the originally 30.000 citizens, and the response rate was 60 % (ibid.).

Of special relevance for this paper is the question of satisfaction with digital municipality services grouped after size of municipality. The results from 2009/2010 show that there are no significant difference in satisfaction between citizens from small municipalities and citizens from larger ones. In services like planning and

building permissions and care for elderly people, citizens from small municipalities give higher score than citizens from larger municipalities. But in services like kindergarten and primary school the result is opposite; citizens from larger municipalities are more satisfied than citizens from smaller municipalities. All in all the users' satisfaction with digital services cannot help us explaining the differences in quality observed in the expert evaluations of the websites. This could very well be an example of what Jakob Nielsen calls the first rule of usability: "Don't listen to the users, watch them work"¹. These results make it necessary to ask: what quality we are evaluating, and for whom?

3.3 Better Quality for the Users?

The objective for evaluating public websites in Norway as well as in Denmark has been to stimulate quality improvement. The results from Norway presented above show that there have been improvements from the evaluation in 2007 to the last undertaken in 2011, in terms of overall score on the quality indicators. Analysis of the same evaluation for the first years 2001-2003 also show a significant improvement in quality [JO04]. Thus, main objective of the evaluation project seems to have been met.

We cannot assure that this really is an indication of better digital services for the users. The results from the user survey described above do not confirm these results; in particular they do not give support for the observed difference in measured quality between small and large municipalities. So what is virtually measured in the evaluations of public websites? The problems with the heuristics which the evaluations is built on, is that they do not necessarily coincide with the genuine user needs and their behaviour. An important aspect that is missing in the expert evaluations is their context, as these evaluations are all carried out through expert testing, which is clearly different from the context of a typical user.

Usability testing would be an obvious response to the problems of the expert-based heuristics and the context problem. But given the number of websites and the vast, accumulated amount of information on them, regular usability testing would not be feasible. Of the three main categories of indicators used in Norwegian evaluations, the accessibility category is the least difficult to assess given the general and widely used heuristics derived from W3C's WCAG work. The more we assess usability and usefulness of websites, the more difficult it gets because our expert-based heuristics have difficulties in capturing the needs and the experiences of a real user.

4 Evaluating Public Websites: What Answers Do We Really Get?

Our main research question was "*what quality is measured in the various benchmarking systems and for whom are we evaluating?*" The discussion above shows that the various evaluation and benchmarking frameworks that have been used in the past 10-15 years produce rather different, if not contradicting results in terms of

¹ <http://www.nngroup.com/articles/first-rule-of-usability-dont-listen-to-users/>

scores and ranking of the different countries. E.g. Schellong [Sc09] shows that the rankings in three major eGovernment benchmarks differ significantly. At a national level, both the two different evaluations of public websites in Norway as well the Danish are not compatible, but they all produce relevant results.

Such differences are not themselves problematic, as they can be attributed to different approaches (methods): the criteria are not compatible and the samples represent distinct universes. The first aspect concerns what *type of measurement method* do we use (as questionnaires, interviews or heuristics involving users or experts, etc.). Another aspect is *scope*: what is to be measured, which is determined by its purpose. The third one is *scale*, which implies limiting the socio-demographic reach of an evaluation, and what is to be assessed and compared. However, the way such results are used creates a lot of confusion and often misleading conclusion, when not taking into consideration the context of the evaluations. In this way they can be a misleading basis for political priorities.

The other research question was: Are the objectives of the evaluations compatible, and are the methods used adequate?

Our understanding of quality is closely linked to fulfilling overall goals and supporting (political) priorities, and has to include various perspectives and dimensions. However, there is often a lack of a clear connection between the purpose of an evaluation and what dimension(s) or perspectives that shall be evaluated. This is illustrated by the multi-functional character of a municipal website: both to serve the democratic ideals, to mainly focus on service-delivery (customer-orientation) in the service provisions and as well as to include emphasize the efficiency perspective. These goals are not necessarily compatible. In that respect, we believe that the last EU benchmark is an improvement, but it does still create some confusion. Our first suggestion is thus:

1. *The design of evaluation framework should be compatible with the specific goals and priorities that are defined.*

Measuring quality should not rely on one single method or approach, but cover different perspectives, and include both objective and subjective criteria depending on the purpose of the measurements. This requires different approaches including formal methods as measurements based on well-defined metrics along with more heuristic based evaluations and user testing etc. Important in this work is to design detailed user scenarios and different user settings in which the website is to be evaluated. These different perspectives do have important implications for how we define the quality requirements. This point is illustrated by the EU benchmark framework, based on a “eGovernment “Progress diamond“ including 4 dimensions: i) Better eGovernment, ii) Efficient eGovernment processes, iii) Egovernment Building blocks and iv) eGovernment empowerment. Our second suggestion is then:

2. *The selection of quality criteria set should reflect the perspectives that are the primary target for the evaluations.*

Is standardisation one way to go? We do agree that a “mild standardisation” in the benchmarking approach (criteria set, methods used, type of heuristics, etc.) can be an efficient way of help improving the quality of public web sites and it can also be an efficient instrument to ensure that public bodies follow standards; either formally approved standards or recommended standards. It presupposes, however, that such standards as well as the arguments for such standards are widely recognised, which also imply that the objectives are well defined and accepted.

However, we emphasize that the indicator sets should include more criteria than those the existing evaluations are based on. We furthermore claim that these evaluations should be supplemented with other types of testing in order to get a more comprehensive picture of a web site. The experiences from such testing can then be used to further develop these indicator sets. Usability tests of a selection of the evaluated websites would give valuable feedback to the development of indicators. Surveys among users and those responsible for the work with the websites are also valuable methods that can give a richer picture and complement the overall quality issue. It is then important to bring these methods together in a common framework, and not separating them in different projects and processes as is currently done in Norway. Our third suggestion is thus:

3. Evaluation or assessment frameworks should combine multiple methods and techniques, and should be used in ways that allows for learning and knowledge accumulation within this field.

There is clearly need for more research, e.g. in where different methods of quality assessment are combined and the effects ultimately measured on real users. There is a need to combine heuristic methods (expert evaluations), usability tests, and user surveys to try to find a link between these. Such research should also help inform practitioners and not least decision-makers (politicians, etc.) about the usefulness as well as the limitations of various benchmark approaches.

References

- [Ba07] Bannister F., “The curse of the benchmark: an assessment of the validity and value of e-government comparisons,” *International Review of Administrative Sciences*, vol. 73, no. 2, pp. 171–188, 2007.
- [BO95] Braa K. and L. Øgrim, “Critical View of the Application of the ISO Standard for Quality Assurance,” *Information Systems Journal*, vol. 5, no. 4, pp. 253–269, 1995.
- [Ca12] Cap Gemini, IDC, Sogeti, IDC, IS-practice and IndiGov, RAND Europe, and Danish Technological Institute, “eGovernment Benchmark Framework 2012-2015 Method Paper,” European Commission DG Communications Networks, Content and Technology, Brussels, SMART 2012/0034-1, Jul. 2012.
- [Dif10] Difi, “Innbyggerundersøkelsen Del 2,” Agency for Public Management and

- eGovernment (Difi), Oslo, 2010:14, 2010.
- [Dif12] Difi, "Kvalitet på nett," Kvalitet på nett. 2012.
- [Dig12] Digitaliseringsstyrelsen, "Bedst på Nettet - Vurderingsgrundlag 2012," Digitaliseringsstyrelsen, Copenhagen, 2012.
- [DJL10] Donker-Kuijjer M. W., de Jong M., and L. Lentz, "Usable guidelines for usable websites? - An analysis of five e-government heuristics," *Government Information Quarterly*, vol. 27, no. 3, pp. 254–263, 2010.
- [DM93] Dahlbom B. and L. Mathiassen, *Computers in context*. Basil Blackwell, 1993.
- [GESD10] Cap Gemini, Rand Europe, Sogeti, and DTi, "Method Paper 2010," European Commission, Directorate General Information Society and Media, 2010.
- [Gr02] Grönlund Å., *Electronic government: design, applications and management*. IGI Global, 2002.
- [Gr10] Grönlund Å., "Ten Years of E-Government: The 'End of History' and New Beginning," in *Electronic Government - 9th IFIP WG 8.5 International Conference, EGOV 2010, Lausanne, Switzerland, August 29 - September 2, 2010, Proceedings*, Springer Berlin/Heidelberg, 2010, pp. 11–24.
- [JG00] de Jong M. and T. van der Geest, "Characterizing web heuristics," *Technical Communication*, vol. 47, no. 3, pp. 311–326, Aug. 2000.
- [JO04] Jansen A. and S. Ølnes, "Quality assessment and benchmarking of Norwegian public web sites," in *4th European Conference on E-Government*, Dublin, 2004, pp. 17–18.
- [Ka11] Kahneman D., *Thinking, fast and slow*. Farrar, Straus and Giroux, 2011.
- [LA11] Langørgen A. and R. Aaberge, "Gruppering av kommuner etter folkemengde og økonomiske rammebetingelser 2008," *Statistics Norway*, Oslo, 8/2011, 2011.
- [Mi12] Ministry of Government Administration, Reform and Church Affairs, "På nett med innbyggerne - Regjeringens digitaliseringsprogram," Ministry of Government Administration, Reform and Church Affairs, Oslo, 2012.
- [Ni93] Nielsen J., *Usability Engineering*. San Diego, CA: Academic Press, 1993.
- [Ni94] Nielsen J., "Heuristic evaluation," in *Usability Inspection Methods*, New York, N.Y.: John Wiley, 1994.
- [Se12] See Public Services Online 'Digital by Default or by Detour? Assessing User Centric eGovernment performance in Europe – eGovernment Benchmark 2012
- [Sc09] Schellong A., "EU eGovernment Benchmarking 2010+ General remarks on the future of benchmarking Digital Government in the EU," Dec. 2009.
- [UN08] United Nations, "UN eGovernment Survey 2008: From eGovernment to Connected Governance," United Nations, New York, ST/ESA/PAD/SER.E/112, 2008.
- [Vi02] Videnskapsministeriet, "e-Service på borgernes premisser? Statusrapport for Bedst på Nettet 2002."
- [W08] W3C, "Web Content Accessibility Guidelines (WCAG) - W3C Recommendation." W3C, Dec-2008.
- [We08] West D., "State and federal electronic government in the United States, 2008," *The Brookings Institution. Governance Studies*, 2008.