When Reality Comes Knocking
Norwegian Experiences with Verifiable Electronic Voting

Ida Sofie Gebhardt Stenerud and Christian Bull

Norwegian Ministry of Local Government and Regional Development
P.O. Box 8112 Dep.
0032 Oslo
Norway
{ida.stenerud | christian.bull}@krd.dep.no

Abstract: This paper discusses the Norwegian experiences in piloting a verifiable, remote voting system in a legally binding, public election. First, we provide a high-level description of the system used. We then go into detail about the major challenges that were encountered in the implementation and execution of the system. In particular, the generation and printing of return codes and the key management are described in detail. We also discuss the relationship between the Norwegian Electoral Management Body and the system integrators, indicating how verifiability may enable new models of cooperation.

1 Introduction

During the municipal and county council elections in September 2011, Norway conducted trials using remote electronic voting. Ten municipalities participated in the trials, and the approximately 168.000 voters could vote online during the advance-voting period, lasting for 30 days. These trials were unique in that they – as far as we are aware– represented the first venture into coercion-resistant, verifiable, and remote electronic voting conducted by a national government. The Norwegian system is able to mathematically prove that recorded votes are counted correctly, and this is verifiable to independent third parties. In addition, voters get proof that their voting intent has been correctly recorded.

The purpose of this document is to provide a primary source of insight into the practical sides of piloting verifiable electronic voting. The intended recipients are the Electoral Management Bodies of other countries that may be considering piloting or implementing Internet voting. Some of the lessons learnt throughout the project have been painful, and by sharing them, we are hoping to make the road less rocky for the next country in line.

We also hope that these practical experiences are noted by academic protocol authors. Seemingly insignificant protocol design choices may have unexpected real-life consequences when implemented. Therefore, practical considerations need to be taken in protocol design.

In Norway, the Ministry of Local Government and Regional Development acts as the Electoral Management Body (EMB) and is responsible for electoral rules and regulations. While local authorities are usually responsible for actually carrying out the elections, the ministry took a more hands-on approach in the case of the e-voting pilot. Therefore, in this paper, the terms "EMB", "Ministry" and "e-vote 2011 project" will be used interchangeably.

2 Functional Overview of the Norwegian Electronic Voting System

From the voter's perspective, the Norwegian electronic voting system is fairly simple. The voter logs in using MinID, a widespread, well-known, and freely available two-factor authentication mechanism. Once verified, the voter is presented with a point-and-click interface showing the ballot. The voter makes her selections and submits them to a Java applet, which has already been downloaded to the voter client PC. The applet encrypts and digitally signs the vote and then sends it to the central voting servers.

Immediately after voting, the voter receives a text message containing a 4-digit number, from now on referred to as a *return code*. This return code can be compared to the voter's poll card. The poll card, which the voter receives by mail before the voting period begins, contains a list of all the available parties to vote for and their corresponding 4-digit code. The return codes are individually calculated per voter prior to the election. The return code in the SMS should correspond exactly to the chosen party printed on the poll card. This allows the voter to verify that the vote has been correctly received by the voting server, and is referred to as a cast-as-intended proof. If the codes do not match the option for which she voted, she will know that the vote has not been received correctly.

The voting process is illustrated in Figure 1 below:

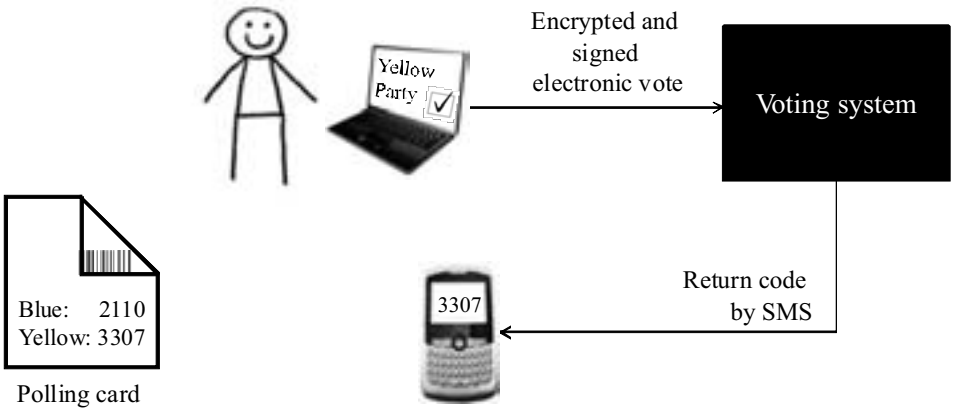

Fig. 1: A functional overview of the voting process

To mitigate the threat of coercion in Internet voting, voters are allowed to cast an unlimited number of Internet ballots, and even cancel the electronic ballot on by voting on paper. This feature is not discussed further in this paper. For more information, see [Gj10].

Why were the return codes sent via SMS and not just displayed on the screen? If a voter casts multiple votes, and the return codes were shown on the voter's computer, an attacker could learn the meaning of the return codes and replace the vote without the voter noticing. Therefore, the codes are delivered out-of-band.

Note that checking the return code is entirely optional and that the poll card is not used for authentication. Hence, a voter not in possession of the poll card can still vote, but will be unable to verify the SMS return code.

3 Return Codes Production: A Series of Unfortunate Events

The return codes form the first part of what is known as the Norwegian end-to-end[1] verifiable voting protocol (see Figure 2 below). Verifiability enables voters, election commissions, and election observers to verify the integrity of the election results and thus increase transparency and trust in the election [Ka11]. Such protocols are often seen as a measure to build voter trust.

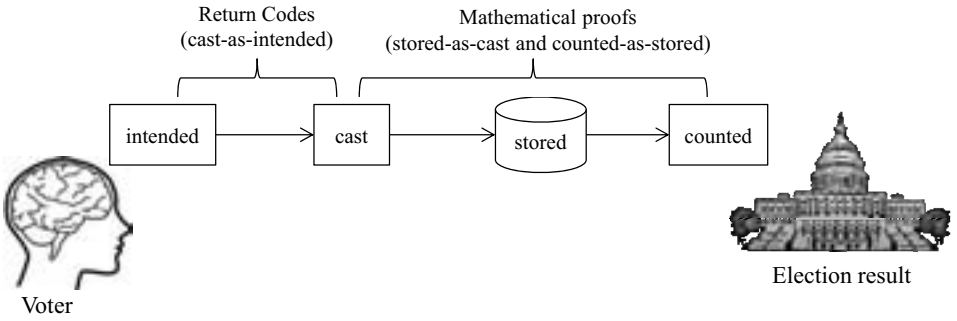

Fig. 2: The vote life cycle and the verification steps

The rationale behind implementing return codes in Norway was, however, somewhat different. The main purpose was to give the EMB the ability to detect systematic manipulation of client computers. In fact, the return codes were a solution to the requirement OS8.7 of the system requirement specification: *"Even though the e-voting client domain may be under outsider control, the e-voting solution shall be such that it is not feasible for an outsider to systematically manipulate the votes without detection"* [Ev09]. However, the fact that they also seemed to raise trust was a welcome side effect.

[1] The Norwegian use of the term "end-to-end verifiability" is somewhat controversial. However, the system enables verification of the entire life cycle of a vote, from end to end.

For the EMB to be confident that an attack would be detected, a certain percentage of voters would need to actually perform the check of their return codes. Though calculations of this percentage have not been published, they will most likely be similar to those published for the Pnyx protocol:

> *In an election with 40,000 ballots cast and a manipulation of just 1% of them, the chances of detecting the manipulation are more than 90% if just 230 voters verify. If 2% of the voters verify their ballots, the same manipulation is detected with a probability of more than 99.9%.* [Sc05]

At the time of writing, we do not have any estimates of the percentage of voters who performed the verification. However, to test the system prior to the pilots, the Ministry conducted several small-scale, non-binding test elections (so-called pre-pilots), with return codes used in two of them. According to data from a voter survey conducted by Synovate AS, an independent market survey provider, close to 90% report to have checked the return codes in these tests. Raw data can be found in [Ev11] (Norwegian only). Though one should be careful to generalize from this small sample, these are undoubtedly high numbers. Still, considering that return codes are pushed out to the voter by text messages, and require very little effort to check, the numbers are probably not so unrealistic when it comes to the actual pilot.

In general, return codes were well-received by voters. In-depth interviews indicated that voters found the return codes "confidence-inspiring", and some voters with disabilities mentioned how it gave them confidence that they had managed to cast their vote successfully. Interestingly enough, survey data from the pre-pilots that were conducted without return codes also showed that the majority of voters had high confidence in the solution. This is perhaps a symptom of the high level of trust in Norwegian elections.

3.1 Return Code Printing

Even though we received positive feedback on the simplicity of the cast-as-intended verification process, this was anything but simple to implement. The return codes created significant challenges in the generation and printing processes.

During the configuration phase, two data sets are created.

1) The voter list, containing all eligible e-voters
2) The return code sets. Each set consists of a list of parties and their corresponding 4-digit return codes.

Initially, the contents of these files are not linked, and no secret can be learned by the possession of just one of these files. However, the *relationships* (henceforth called "bindings") between individual voters and return codes are very sensitive. An attacker in possession of the return codes, the voter list, and the bindings, plus the ability to monitor

the SMS gateway, will be able to breach voter privacy. For an outsider, this would be nearly impossible to achieve. However, as the EMB is essentially in possession of all this data, great care must be taken to ensure that the EMB is never able to break voter privacy.

To ensure that the Norwegian EMB is able to learn the meaning of the return codes, the return code generation process generates an output encrypted with the public key of the printer service. The key pair is generated by the printer service, and only the printer service is in possession of the decryption key. Therefore, the EMB cannot learn the return codes. In addition, the bindings are created by the printer services during the printing process. This process is open to observation and in 2011 was observed by representatives from the EMB and the OSCE.

While this procedure ensures that the EMB is not able to violate privacy, the printing service is now in possession of uncomfortable amounts of data. To make sure that no single person or component is in possession of sufficient information to violate privacy at any time, printing is divided into two separate phases, each performed in a physically and logically separate printer environment. Figure 3 illustrates the process of printing return codes on poll cards.

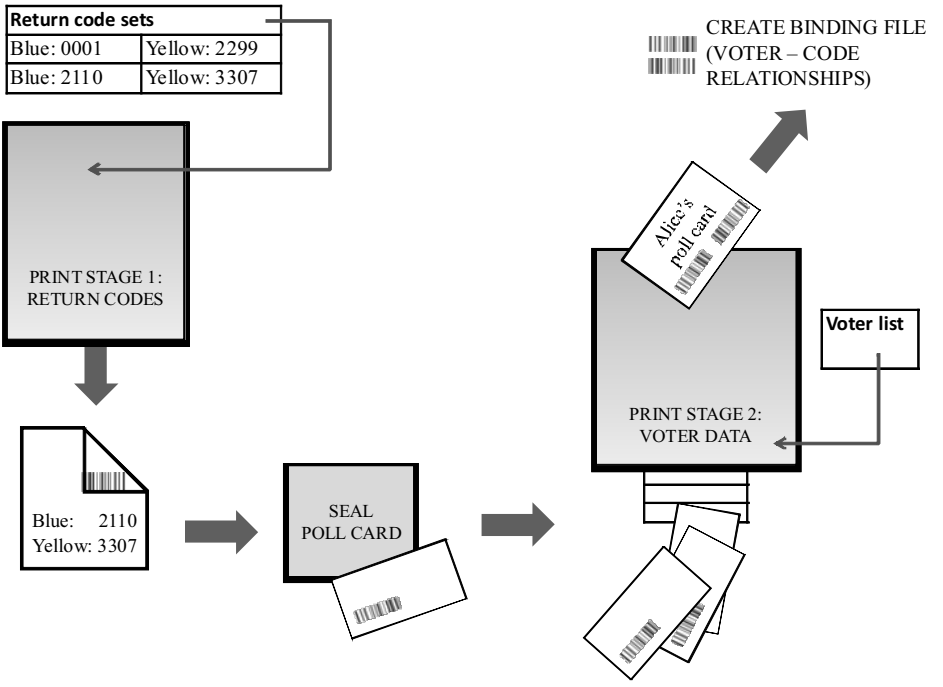

Fig. 3: The poll cards printing process

In print stage 1, the printer service randomly selects a return code set, and prints it on the inside of an A4 sheet. This sheet is then folded, sealed, and perforated so that the only thing printed on the outside is a bar code representing the ID of the return code set. During the 2011 pilots, in order to increase the opacity of the sealed poll card, the EMB used extra thick paper (120g) and coated the entire inside with yellow ink. The yellow ink also had the benefit of increasing contrast for improved readability; the thicker paper increased postage costs.

Once sealed, poll cards are manually shuffled and moved to print stage two, which is physically and logically separate from stage one and operated by different personnel. Here, eligible voters are picked at random from the voter list and their personal data printed on a poll card. The binding between voter and return code set is read from the bar code and subsequently written to file. This file is then uploaded by the EMB to the component responsible for sending out the return codes by SMS. This process ensures that no single person or component can ever know the meaning of the return codes relative to an individual voter.

Even though the print process was tested prior to the 2011 pilot, problems were encountered when it came to producing larger number of poll cards. While details are not entirely clear, we know that there were incidents where the actual poll card did not correspond to the information in the bindings file. This caused a few voters to receive the wrong return code after voting. Out of the approximately 168,000 poll cards that were produced, from which 28,001 voters actually cast an electronic vote, the support call centre received 74 reports from voters who received a return code that did not match their vote option [NS11].

While this might sound like a potential disaster, it did not cause any uncertainty in the integrity of the system. The EMB knew that if there had been any vote manipulation, the received return code would have corresponded to one of the other return codes on the voter's poll card. Anything else would have been mathematically impossible. Fortunately, for all the affected voters, the SMS return code never corresponded to anything printed on the poll card.

On a positive note, this provides a good indication that voters not only read and understand the return codes, but act as instructed when something seems amiss. If there was any sign of manipulation, the EMB would have encouraged the voter to cast a physical ballot and started an investigation. As electronic voting was only available in the advance voting period, any voters subject to manipulation would have had time to cancel their electronic vote by voting on paper on Election Day.

3.2 Challenges Posed by Security Controls

Running simultaneously with the e-voting system is an elections administrative system. Here, all the rules governing the election, such as municipal data, eligible party lists, and election opening hours are configured. The print files containing voter data and return codes are based on data from the administrative system. Because of late changes to the

administrative system, some eligible party lists were not included in the original print file. As these files were encrypted with the printer service public key, the Ministry was unable to check their contents for correctness. The missing data were discovered in an extraordinary check of the administrative system. At this time, the return code printing was going on, causing the entire first batch of poll cards to be discarded.

Before printing could be resumed, the Ministry had to re-generate return codes, a challenge in itself, as the infrastructure was unavailable due to the terrorist bombing only nine days earlier. The building in which the return code generation servers were housed was a crime scene and thus inaccessible to the Ministry. After a few days, the Ministry was granted special permission to evacuate the servers. When printing was finally restarted, there was only a matter of days before the opening of polls. At this point there was not enough 120g perforated paper available, so paper thickness had to be reduced to 90g.

In addition to the delay caused by the re-generation of return codes, the printer company had also discovered that the printing process was significantly slower than expected. All this leads to a mad rush in the printing of poll cards, with three shifts working around the clock for several days. On the morning when the system was to be made available to the public, printing was still underway for the two largest pilot municipalities. As the generation of the bindings file is part of the printing process, voting cannot commence before printing is finished. This led to a few hours delay in making the system available for voters in the two affected municipalities.

In addition to the 74 reports on incorrect bindings, the support call center received another 35 return code related calls.

- 11 voters reported not having received a poll card
- 5 voters who voted online reported not receiving a return code
- 4 voters received a poll card with the return codes smeared
- 1 person received two poll cards, one with the correct binding and one incorrect
- 2 callers reported having received return codes without having voted

Upon receiving the first reports on incorrect return codes, the Ministry conducted an investigation into what had happened. As part of this investigation, representatives of the Ministry personally called several affected voters. Interestingly, the voters reported not having lost trust in the system. Rather, they felt that it was their duty to do as instructed and inform the authorities of the incident. When informed of the problems with the printing, all affected voters appeared assuaged.

All in all, while there were certainly problems related to the return codes, the Ministry is very happy with its first experience in using them. If the piloting of Internet voting is continued in Norway, our advice to the Ministry is to continue the use of return codes even where they, from a security standpoint, may not be strictly required (for example, for expatriates or low-value elections).

As should be evident from the preceding text, the return code solution piloted in 2011 was not entirely perfect. For instance, the printing process definitely needs re-working. In addition, both the voter information material and the user interface must be improved in order to better educate voters.

4 Verifiability by Proxy

In Figure 2, the return codes only form the first part of the Norwegian verifiable protocol. The second part is performed without any voter involvement. This is an extremely important feature as the return codes only verify to the voter that her intent has been correctly captured. They do not verify whether the vote has been correctly stored in the database or that it will be counted.

An in-depth description of this last part of verification is beyond the scope of this paper but can be found in [Gj10]. In sum, the system allows a verifier to independently verify

1. That return codes have been sent for all received ballots
2. That all received ballots have been stored
3. That all stored, valid ballots have been included in the tally

The Norwegian voting infrastructure must provide these proofs of correct operation to the verifier. This ensures that neither malfeasance on part of the EMB, nor any software error (intentional or unintentional) will undetectably alter the vote once cast. The fact that these measures were implemented to form a verifiable system ensured a lot of good-will in the academic community and among IT experts. We strongly believe that this academic support was important in achieving wide-spread trust in the technical solution.

4.1 The Effect of Verifiability in Trusting Infrastructure

As ever, the advantages of verifiability were not only apparent in building trust. An extremely positive side effect of verifiability was the fact that the EMB did not have to put complete trust in the counting infrastructure: the integrity proofs of the cleansing, mixing, and decrypting would reveal any irregularities.

Counting of electronic votes is extremely critical and even small errors can have dramatic consequences. It therefore seems common practice in electronic voting to use new servers for counting. Configuration and use of these is then performed under strict supervision. Considering the extensive number of certificates, keys, and passwords that need to be correctly in place for the Norwegian counting infrastructure to even operate, an untested infrastructure was unlikely to work on the first go. However, since the verifiable properties of the system allow, without any risk, the re-use hardware, the Ministry was able to perform test counts on the production system as late as Election Day to ensure that all components were functioning correctly.

In other words, the EMB itself has a clear self-interest in, and much to gain from, implementing verifiability in the system it deploys. This does not appear to be a motivation for most academic protocols, but has been a boon for the Norwegian government. On the other hand, verifiability is both computationally expensive and complex to implement. Though it is difficult to give an estimate of the extra development effort, it obviously raises the price.

4.2 The Legal Impact of Verifiability

Verifiability means that any manipulation or system error related to the processing of votes will be discovered. However, one can only know this once the election is finished. An obvious question is how to proceed if the proofs indicate irregularities. In the Norwegian e-voting pilot, the protocol would have been the same as in any electoral irregularity: the government would conduct an investigation. If the problems were shown to possibly have affected the election outcome, an option would have been to invalidate the results and call a second ballot. Note also that not all verification is performed after the e-voting period is over. As cast-as-intended verification is performed during the voting period, this would allow the EMB to detect irregularities during the advance voting period and act accordingly.

Even though an invalid proof would certainly have been unpleasant, it is still better than the worst-case outcome – an illegitimate winner of the election.

5 The Challenges of Key Management

Though not strictly related to verifiability, it's safe to say that one of the major challenges for the e-vote 2011 project was key management. To ensure integrity of the information flow, all communications between the different components were signed by the originating server and the signature verified by the recipient. The configuration phase creates, among other things, 15 different key pairs per election event, each consisting of a private key, a public key, and a password for the private key. Ensuring that each server had the correct files, when each component consisted of up to 10 servers, was a complex task.

For increased security, the passwords protecting the cryptographic keys were only held in the memory of the server. This means that restarting a server, or just the application, would require the passwords to be re-uploaded. If any one server lacked just one password, it would not have been possible to cast a vote using this server. For instance, if one of the ten RCG servers lacked a password, voters would have experienced intermittent failure when casting their votes (approximately one in ten votes).

This creates an additional challenge: How to gain 100% confidence in the correct functioning of the system before the opening of the election? The answer is that although the system vendor developed sophisticated "health checks" for the infrastructure, it was not, strictly speaking, possible. As one of many controls to assure that no one could cast a vote before the actual opening of the voting period, the system had a built-in scheduler that prevented this. It was therefore not possible to verify that votes would be accepted by the system before opening the election and the correct return codes calculated.

This was a typical paradox encountered several times: the strict security controls gave great confidence that no malfeasance could occur, but at the same time they also reduced the ability to test the system. This is one of the great dilemmas of secure electronic voting, and even within the e-vote 2011 project group there has been some disagreement on which property is more important.

5.1 Key Management and Separation of Duties

Cryptographic key management is a very challenging undertaking. One thing is the secure storage of secret keys; another is access control to those same keys. Typically, a small number of people both create the keys and have access to critical infrastructure. The only remedy for this is the separation of duties on the organizational as well as the technical level. In a small and fast-paced pilot project, this is, for all practical purposes, impossible to implement but will be a vital development in more mature electronic voting.

As part of the system design, a significant amount of separation of duties was implemented to ensure that critical secrets were kept apart. For instance, 4 laptops, 10 servers, 45 hard drives, and countless USB flash drives were used in the configuration. Even though separation of duties was implemented on system level, it proved difficult to implement similar controls at the personnel level. This was partly due to delays in the delivery of software, which created an unpredictable situation. To alleviate this problem, the EMB identified the most critical keys and secrets and created procedures to ensure that these were safely kept secret and separate. Despite the EMB's best intentions, the actual separation of duties is difficult to verify for an outsider. This would either require long-term observation or very advanced high-security storage equipment.

6 Does the EMB Need Complete Ownership of a Verifiable System?

The Norwegian approach was to assume as much ownership as possible, in order to ensure transparency and public trust. The software vendor was used only for development. On the negative side, assuming ownership means assuming risk. However, the buck will always stop with the EMB, regardless of contractual responsibilities.

It appears to us that end-to-end verifiability may in fact reduce the need for EMB ownership and involvement in the e-voting system. The fact that the processing of votes is independently verifiable means, that the EMB can safely transfer more operational responsibility to external parties, such as the software vendor or data center operator. Some of the challenges encountered by the Norwegian pilot project, such as key management and true separation of duties could have been more manageable with such an approach.

While a verifiable e-voting system may allow the EMB to take a somewhat more relaxed approach to operations, it does not reduce the need for close cooperation with the vendor. Even with small-scale piloting, an Internet voting project demands extensive development of the actual e-voting systems and the legal requirements to conduct such an election. The customer must always assume full responsibility for specification and testing and ensure that the system is, in fact, truly verifiable.

7 Further Research

We would certainly not argue that the Norwegian protocol is perfect. Certain identified threats have not been fully mitigated. For instance, we are not aware of any way to prove that the SMS received by the voter was in fact sent by the authorities. It would be beneficial if the veracity of the SMS could be proven to the voter and the EMB.

Independent researchers have also conducted a series of lab tests trying to exploit the weakest link in the protocol – the voter. In these experiments, test voters were presented with a malicious web site that changed the vote before encryption. Such a web site will never be able to calculate the correct return code, but it could undetectably steal the vote if the voter fails to notice any irregular behaviour. In one of the experiments, the malicious site tricked the voters into both 1) typing in the return code of the chosen vote option and 2) ignoring the fact that they received two text messages – one of them with a "wrong" return code. Disturbingly, none of the test subjects detected the deviation from the protocol [Ol11]. Further research is needed to understand whether or not these results can be applied to actual voting situations. What is certain, however, is that the protocol only requires a very low number of voters to notice irregularities in order for the EMB to detect an attack.

Another hypothetical "attack" is that a group conspires to falsely report wrong return codes. Since it would be impossible for the ministry to know whether reports are truthful or not, this would be a very difficult attack to defend against. One possible defence would be for the EMB to visit every person who reports wrong return codes and physically test their computer. Because the Norwegian EMB is represented by the local government in the municipalities, this would have been feasible but legally and politically unacceptable.

Additionally, the protocol, as it currently exists, makes the rather strong assumption that the vote collector server (VCS) and return code generator (RCG) will not cooperate to violate privacy. On one hand, this is an uncomfortably low number of actors required to guarantee privacy. On the other hand, maintaining even two different operating sites introduced significant unwanted complexity, as described in chapter 5 above. From the EMB's point of view, reducing complexity would be desirable.

8 Concluding Remarks

After reading this paper, the reader might question whether verifiability is worth the time and effort, when trust in the EMB is already high. We contend that the best, and quite possibly only, way to gain trust in the academic community is to implement a verifiable system. Support from the academic community will probably not in itself create trust among the general public. However, a good relationship with the academic community at least reduces the danger of a sudden mistrust of the technical platform.

Furthermore, verifiability is confidence-inspiring for the EMB. While the security measures implemented in the Norwegian e-voting system may appear difficult to live with, the challenge was temporary and most evident during the configuration phase. Once the system was up and the votes were coming in, the benefits became apparent in the very high confidence in the system. Also, piloting a brand new system of some complexity will always be demanding and somewhat chaotic. If piloting electronic voting is continued in Norway, we believe that the process will go more smoothly.

Procuring an E2E verifiable electronic voting system is not a simple task. This is a question of having the right resources available, both in terms of money and personnel. Hence, one should be weary of organisations without sufficient resources piloting electronic voting, as maintaining trust in electoral processes is of great importance to any democracy.

In this paper, we have indicated that with end-to-end verifiability the EMB may be somewhat more relaxed regarding the ownership of the election system and infrastructure. However, this only holds as long as the system is well tested. The Norwegian EMB in no way regrets taking on an active role as customer. The EMB must always assume full responsibility for specification and testing, in addition to ensuring that the final system is, in fact, truly verifiable.

An uncompromising outlook on security can be painful. However, we believe that it's a worthwhile cause. In many countries, the alternative will be distrust from the stakeholders. Verifiability is an important component in such an election, increasing the confidence in the EMB and of the stakeholders during and after the election. However, the intense testing required before the election is one drawback if the necessary resources are unavailable.

Bibliography

[Ev09] The Norwegian E-vote 2011-project, SSA-U Appendix 2B Requirements Table, 2009
http://www.regjeringen.no/nb/dep/krd/prosjekter/e-valg-2011-prosjektet/tekniskdokumentasjon/spesifikasjoner-tilbud-kontrakter.html?id=612121 [February 17th 2012]

[Ev11] The Norwegian E-vote 2011-project, Evaluering av testvalg høst10/vår11, 2011
http://www.regjeringen.no/nb/dep/krd/prosjekter/e-valg-2011-prosjektet/evaluering/evaluering-av-testvalg-host10var11.html?id=653612 [February 17th 2012]

[Gj10] Gjøsteen, K.; Analysis of an internet voting protocol, 2010
http://eprint.iacr.org/2010/380 [February 17th 2012]

[Ka11] Karayumak, F.; Olembo, M.; Kauer, M.; Volkamer, M.: Usability Analysis of Helios - An Open Source Verifiable Remote Electronic Voting System. Presented at EVT/EWOTE'11, 2011
http://static.usenix.org/event/evtwote11/tech/final_files/Karayumak7-27-11.pdf [February 17th 2012]

[NS11] Nore, H.; Stenerud, I.: The good, the bad and the terrible of verifiable electronic voting, VoteID 2011, 2011.

[Ol11] Olsen, K.: Alle ble lurt i falskt e-valg. Published in Teknisk Ukeblad 2011 (31), p. 20-21

[Sc05] Pnyx.core: The Key to Enabling Reliable Electronic Elections. A Description of Scytl's Cryptographic e-Voting Security Software, 2005
http://www.scytl.com/images/upload/home/PNYXCOREWhitePaper.pdf [February 17th 2012]

[SVK11] Spycher, O.; Volkamer, M.; Koenig, R: Transparency and Technical Measures to Establish Trust in Norwegian Internet Voting, VoteID2011, 2011
http://www.regjeringen.no/upload/KRD/Prosjekter/evalg/vedlegg/paper_transparency_and_technical_measures.pdf [February 17th 2012]