# An Investigation Into Using Computational Ethnography as an HCI Method for Evaluating eID Systems

Author: Corlynne O'Sullivan[1,2]

[1] Cyprus University of Technology, Limassol, Cyprus
[2] Tallinn University, Tallinn Estonia
https://www.idmaster.eu/

Supervisor: Mati Mõttus[2]

**Abstract.** This thesis investigates how computational ethnography can be used within HCI for evaluating eID systems. The case study applies text analysis techniques and natural language processing to evaluate the Dutch implementation of eID – the DigiD application. The purpose of this is to gather user data in an unobtrusive and automated methodology. The goal of this study is to present a viable methodology for further research in analysing and improving the user experience of using eID technologies.

**Keywords:** Electronic Identification · Computational Ethnography · Content Analysis · Machine Learning · NLP (Natural Language Processing) · Sentiment Analysis · European Regulations · User Experience

Author:  04.01.2021

Corlynne O'Sullivan

Supervisor:  5.01.2021

Mati Mõttus

## Author's Declaration

*I declare that apart from work whose authors are clearly acknowledged, this manuscript is a product of the author's original work, and it has not been published before, neither is not currently being considered for publication elsewhere for any other comparable academic degree.*

*This master thesis document has been supervised Mati Mõttus (Tallinn University, Estonia).*

Author: Corlynne O'Sullivan
Date: 04.01.2021

## Non-exclusive license to reproduce a thesis and make the thesis available to the public

*I, Corlynne O'Sullivan (date of birth: 30.10.1992), grant Tallinn University a permit (a non-exclusive license) to reproduce for free and make public in the repository of Tallinn University Academic Library a piece of work created by me,*

*"An Investigation Into Using Computational Ethnography as an HCI Method for Evaluating eID Systems", supervised by Mati Mõttus.*

*I am aware of the fact that the author also retains the rights mentioned above.*

*I certify that granting the non-exclusive license does not infringe the intellectual property rights of other persons or the rights arising from the Personal Data Protection Act.*

Amsterdam, 04.01.2021

# 1 Introduction

## 1.1 Research Problem Statement and Significance

A political priority of the European Union is the enablement of cross-border mobility of e-Government services, but the user experience and adoption of such services remain largely fragmented. The concepts of electronic identity (eID) are crucial enablers of e-government solutions in Europe however, a lack of sufficient trust is affecting the adoption and interoperability of eID, this is due to a lack of focus on socio-technical aspects of the implementation of eID and eIDAS regulations (M, Alexander et al., 2017).

In a 2007 survey, the European Commission found that a majority of the countries (28 out of 32) already use or have plans to implement an electronic ID scheme. Although there are some countries with signed agreements for mutual recognition, the implementation of eID systems varies across the member states and interoperability across countries at the time was almost nonexistent (Secure Electronic Identity Across Europe, 2010).

Verdegem and Hauttekeete (2010) argue that the low uptake of eGovernment by citizens has been caused by an over government-centric and technologically deterministic way of working, which has resulted in the creation of e-services that neglect the user. The lack of eID socio-technical input has led to the needs of the end-users being neglected, with more focus being placed on the technical and political aspects of eIDAS regulations and implementation. Thus, the adoption and interoperability of eID in Europe have remained fragmented between cross-border, private and public sectors and EU citizens acceptance of such technologies. A publication from The European Union Agency for Network and Information Security (ENISA, 2018) revealed that there is an insignificant understanding of trust and the implications on users - both citizens and enterprises.

When looking at the longevity of eID, notable findings indicate that "young people see risks to personal data and identity as a continuum that spans from the virtual to the real world. This risk greatly hampers the take-up of eID services" (Lusoli & Miltgen, 2009). Issues that are both highly political and socio-technical in nature, arise when this socio-technicality is ignored. A lot of the approaches to eID have treated the social and the technical as different aspects of the eID, rather than as integrated parts (K, Hedström et al., 2016).

The key problem highlighted throughout the aforementioned papers is the lack of consideration for the socio-technical aspects of eID, such as political and cultural, economic and market structures and technology availability. (Bruggers et al., 2014), as well as lack of, or difficulty of user experience evaluation for eID services.

To explore a more specific example, the Dutch eID service – DigiD states on their website that more organizations are moving towards mobile authentication as a means of logging in. Moreover, they claim that the DigiD app is the easiest way to log in securely. The benefit of using the application for authentication is that you no longer have to remember a complicated password. Now users just

use a pin code. Soon to access DigiD, users will always need their phone to log in, this is a measure to better protect personal data.

Reviewing how DigiD is adapted for mobile is significant considering that "mobile applications are gaining in popularity because of the significant advantages of mobile devices, such as portability, location awareness, electronic identity, and an integrated camera" (F. Nayebi, J. Desharnais and A. Abran, 2013).

### 1.2 Research Goals and Motivations

Unfortunately, until recently eID systems have mostly been viewed from a technical perspective alone, which does not involve socio-technical aspects of adoption and use. Without a good user experience, aiming for interoperability between eID systems will be difficult to achieve as adoption will be low due to users' frustration and lack of trust. Governments must therefore ensure that these pan-European eID services are not only available but also usable and accessible. Completing user evaluations on eID systems can provide insights that promote their adoption by understanding what constitutes a good user experience when interacting with eID systems. As of recently, mobile eID systems have been growing in popularity as a means of authentication, therefore evaluating the user experience of eID applications is both important for the ease of access to eID services.

This thesis intends to construct a case study on the eID implementation in the Netherlands – the DigiD application and will demonstrate how computational ethnography methods can be applied to gather and analyse user feedback in an unobtrusive manner, which could provide guidance for improving eID services. The overarching goal is to prove that computational ethnography can be used in human-computer interaction evaluations as an unbiased, unobtrusive means of gauging user experience of eID services, with the resulting insights highlighting key socio-technical painpoints that service providers and regulators could focus on improving.

### 1.3 Research Questions

1. How can collection and analysis of user experience data be applied to address key concerns, or prioritise improvements and feature development for eID applications? In order to answer this question, we need to further investigate:

(i) How can we gather user feedback in an unobtrusive manner, without introducing bias?

(ii) How can computational ethnography methods used in HCI be applied to evaluate this data to gain insights into their experience?

(iii) How can these insights help shift discussions around eID from technical to a sociotechnical outlook that takes into account the end user exeprience, when developing eID applications.

### 1.4 Research Methodology

A content analysis will be carried out on 100 of the DigiD app store reviews, providing insights into key user concerns. Further sentiment analysis will be run on a larger dataset, which according to Corallo et al, (2015) is a capable method for analysing user interaction with public services and identifying focus areas. The objective of this mixed method text analysis is to systematically transform a large amount of text into an organised and concise summary of key results to provide a set of guidelines for obtaining and using data provided by users when reviewing eID applications to help further research into the user experience of eID applications.

### 1.5 Research Procedure

This section describes in detail the research procedure. This study aims to provide a set of recommendations to help further research and implementation of the user experience based on reviews left for applications developed with the eIDAS regulations. This study picked these methodologies to maintain objectivity and avoid bias. Obtaining all data directly from the user's reviews of the DigiD application helps avoid any possible transference that could happen in popular methodologies such as user surveys.

1. **Theoretical understanding** – Gather and analyse relevant knowledge and insight to have a better understanding of the problem and previous approaches concerning eID and the broader context of eGovernment.
2. **Content analysis** – Using deductive methods to discover key themes affecting the user experience when interacting with DigiD by manually analysing latest 100 reviews on the IOS and Google Appstore
3. **Sentiment analysis** – Deploy Natural Language Processing through Google Cloud Natural Language API for Sentiment analysis to detect sentiment (positive, neutral or negative) by scraping and analysing latest 1000 reviews on the IOS and Google Appstore.
4. **Provide a set of guidelines forgathering eID reviews for useability evaluations and compare manualcontent analysis methods with Natural Language Processing methods for analysing text** – The end goal is to evaluate how textanalysis methods such as content andSentiment Analysis can be used to help gain a better understanding of eID systems.

### 1.6 Methodology Justification

1. The first method that will be deployed to evaluate the reviews will be a manual Content Analysis. This method was chosen based on research completed in the literature review, showing that the majority of studies on eID usability used surveys, which are biased and time consuming when finding participants, whereas content analysis takes advantage of the wide range of readily available data that can be mined from the web to gauge user opition

and gain insights. In "The Role of the Profile and the Digital Identity on the Mobile Content" was listed as a previous method for reviewing how users deal and understand their online profiles and digital identities management through mobile devices. "Firstly, national and international surveys focused on users' actions and preferences concerning mobile adoption and content access. Secondly, content analysis about users' performances on apps and platforms, establishing relationships, in many cases, with the users' awareness about the logic of the devices and the media involved." This method can be deployed for organising and categorising key themes from the reviews that will be retrieved from both the Google and Apple Appstore. A content analysis carried out on Chinese eGovernment services found that the population may have been underrepresented in certains parts since data was taken from websites whereby data was gathered by citizens of individual municipalities (Zhou., Xiang, 2004). By evaluating users' reviews on the app store, a wider insight can be assumed since the reviews can come from any municipality of The Netherlands.

2. Although the content analysis can provide useful information on the user experience, the third method Sentiment Analysis was picked to gain a broader understanding of the overall user experience. According to (Corallo et al, 2015) "In the public administrations, SA is a technique capable of facilitating the creation of relationships between public bodies and citizens. SA is able to discover the criticalities of this relationship in order to focus on taking the right actions." In this study, Sentiment Analysis (SA) was to analyze tweets polarity and to enable the government to quantitatively describe the opinion of active users on social networks with respect to the topics of interest to the Public Administration. In our case Sentiment Analysis
3. This also helps in understanding the cross-border functionality of eID systems. A study by evaluating e-Governance and the use of Sentiment Analysis states that "Through E-governance, government receives the non manipulated feedback from the public and results of these can help to improve the service" (R. K. Bakshi,et.al., 2016). This part of the study will be completed on a larger data set than the content analysis which will also help give us a better understanding of the user's experience that would not be possible to do manually without a NLP processing tool such as MagellanText Mining. The article Classification of Customer Reviews based on Sentiment Analysis states that increasing the sample size affects the specific vocabulary used for customer reviews. This is backed by the fact that customer reviews are relatively brief text snippets. As a consequence, the sample size should be large. (Gräbnera, Zankerb, & Fuchs, 2001).
4. A summary of the information obtained from the content and Sentiment Analysis will help conclude a final set of recommendations for evaluating eID applications.

### 1.7 Expected outcomes

The expected outcome of this case study is to provide insights into the user experience of the Dutch eID management system 'DigiD'. This data will be used to draw conclusions on the main pain points for users when interacting with the eID app. Moreover, this study will show how computational ethnography methods, such as text mining, content analysis and sentiment analysis can be used as unobtrusive and unbiased methods for evaluating user experience using existing, public user feedback. This would demonstrate that trusted service providers like DigiD and eIDAS observatory could easily leverage these methods and apply the findings in their joint decision making on usability and user experience.

## 2 Literature Review

### 2.1 The role of computational ethnography in human-computer interaction

Computational ethnography is an emerging family of methods for conducting human-computer interaction (HCI) studies. Computational ethnography often leverages automated and less obtrusive means for collecting in situ data that reflect end-users' true, unaltered behaviours of interacting with a software system or a device in naturalistic settings. This is done by analysing data, or of digital trace data available within a system. (Patel, Kannampallil, & Kaufman, 2015)

In the paper Computational Ethnography: Automated and Unobtrusive Means for Collecting Data In Situ for Human-Computer Interaction Evaluation Studies, it is stated that HCI researchers are becoming more interested in collecting data such as rich behavioural traces. This helps in the evaluation of user interactions with online systems in situ (natural environment) at a scale not previously possible by thorough data analysis. These logs can be used to characterize user interactions with existing systems and compare different designs.

Natural Language Processing (NLP) and Computational Linguistics (CL) methods are not commonly used for aiding design efforts. Computational ethnography is a means by which using large bodies of data to provide insight into the routine of users for use in subsequent design efforts. Notably for this study, narrative text is often a subset of this data, according to Arnold & Fuller NLP/CL methods are well-matched for analyzing bodies of existing user language (Arnold & Fuller, 2018).

In the paper Understanding User Behavior Through Log Data and Analysis, it is noted that HCI researchers are increasingly collecting rich behavioural traces (data) from user interactions with online systems in situ at a scale not previously possible. Moreover, these data logs are then used to help characterize user interactions with existing systems and compare different designs.

Participatory information spaces including social media platforms provide affordances for users both to leave behavioural traces of their informational activities and to find such traces from other users.

### 2.2 Background of eID in Europe

In the last few years, European countries developed a number of authentication schemes based on electronic identification. The main goal of these authentication schemes is to provide access to different services, called eID services. For these services, access should be provided for citizens and organizations by using information already available on eID cards, for example, on the personal ID card which was issued by a government institution (Engelbertz et al., n.d.). Many countries developed their own authentication schemes, however, a common identification and authentication space is one of the goals set in Europe's Digital Agenda. Interoperability of electronic identities (eIDs) across Europe will facilitate mobility and cross-border e-business and therefore contribute to growth.

### 2.3 eID and the lack of acceptance by end-users

A major challenge remains in growing acceptance for such a system by end-users, service providers and national governments alike (Brugger, Fraefel, & Riedl, 2014). The EU, since the mid-nineties, has been developing a significant number of initiatives (roadmaps, agendas, action plans, research projects, etc) in the field of eID. The overall objective guiding such initiatives has been the construction of a European cross-border eID framework, based on interoperability and mutual recognition of national eID resources and management systems. (N. N. Norberto Andrade, 2013) A major challenge for European electronic identity is legality. Notably, the technology needed to enable an interoperable eID across Europe already exists, however, the issue is not just regarding usability but also the lack of a harmonized compliance between member states (N. N. Norberto Andrade, 2013).

In a study that took into account the perceptions of eID in Germany and the UK, concluded that benefits for the public should be clear to promote the adoption of eID systems. A study initiated as part of the EU project FIDIS (Future of Identity in the Information Society) revealed some key aspects that users and policymakers expect from a public identity management/eID systems. The first expectation is that the system should help reduce online fraud and promote public trust. Cost reduction was also cited as a beneficial outcome influencing the need for interoperable identity service. However, a striking divide was noted between the needs of the government and the citizens. "If perceived benefit and public value are important evaluation considerations for all stakeholders in the development of eGovernment and, indeed, a pre–requisite of the citizen–centric vision, then policy–makers must address this emerging gap" (Halperin & Backhouse, 2012).

### 2.4 European adoption and interoperability of eID management system

In the 2018 publication "Institutional Legitimacy and Digital Public Cross- Border Service Delivery Between Denmark/Sweden and Denmark/Germany", recog-

nition is given to a number of initiatives within the EU, whereby member states are facilitating digital cross-border service delivery. The journal states that this is limited and far from seamless in all cases of collaboration. The journal draws on a current example of cross border mobility; between Estonia, Lithuania, Finland and Belgium. In this arrangement, citizens of these countries can access the Estonian Digital public government to business services using their national credentials. There is also a cross-border effort between the Benelux countries - Belgium, Luxembourg, and the Netherlands. Cross border mobility in its current infancy, "is far from the vision of achieving a seamless cross border public digital service delivery"(Halperin & Backhouse, 2012).

Another point highlighted in this publication is that currently, "different EU member states run different IT systems". Most of these IT systems are not interoperable with one another". The paper defines two more major points on why this is a relevant issue:

– "The design, implementation and service delivery processes via the national public digital service in each country differ." – "The design of the IT system and the service processes produced in these systems are enabled and backed by national laws, national norms, and their culturally cognitive way of doing things." (Williams, Falch, & Tadayoni, 2018).

A recent study by Windisch & Müller (2018) concluded that Nordic countries, such as Denmark, Finland, Norway and Sweden, the subject of e-Identity looks comparatively well-organized and adoption and use remain high. Correlating to this point, in 2012 a Danish law was passed making all official communication available in digital form only. The percentage of citizens using eID therefore is close to 100% (Windisch & Müller, 2018).

### 2.5 Electronic identification and the eIDAS regulation

eIDAS stands for 'Electronic Identification and Trust Services for electronic transactions in the Internal Market'. The eIDAS regulation was created by the European Union in 2014, since then it has received much traction from EU member states hoping to create a single digital market for economic benefit and to work towards the European goal for the free movement of people, goods and services. The eIDAS Regulation attempts to provide the European Union with both regulations and a platform for the promotion and adoption of a secure eID across member states. As more member states adopt the eIDAS regulations, it should become possible for a citizen to use their eID to log-in to the online public services of other European countries. However, to drive uptake of eIDAS-based eID across Europe, the experience of doing so must be strong. A study was completed by Colette Cuijpers and Jessica Schroers (2014) to investigate the eIDAS as a guideline for the development of a pan-European eID framework in Future ID concluded that it is difficult to tell whether or not the eIDAS regulation will lead to cross border electronic identification and authentication large scale acceptance and use in Europe considering that the prior eSignatures Directive was dubbed to be 'used by few and ignored by many' Colette Cuijpers & Jessica Schroers, 2014).

This EU regulation provides the basis on which public-sector organisations in EU member states can accept each other's electronic identification schemes. Once a national eID has been approved by the EU, you will also be able to use it in other EU member states. If you have an approved eID, you can use it to access Dutch public services online, such as pension funds National eIDs such as DigiD cannot automatically be used in other EU member states. First, the EU needs to approve the eID. The eIDAS Regulation of the European Union goes beyond electronic identification. An important part of the Regulation deals with electronic trust services, such as digital signatures, electronic seals, time-stamps, electronic delivery, and website authentication. The Regulation ensures that an electronic signature has the same legal value as its paper-based equivalent, allowing transactions such as signing contracts to be finalised online.

### 2.6 How do European citizens encounter eIDAS?

Government organisations in EU member states usually identify people based on their passport. Which country issued the passport is not relevant; it is still a valid form of ID. In the Netherlands, for example, you can use an Italian passport as proof of identity. This is the case for electronic identification (eID), like DigiD. Many countries only accept an eID issued in the citizens country as a valid form of electronic identification. Following the introduction of the eIDAS Regulation, this is changing. Currently, many European countries have a nationally issued eID method in place for citizens, residents and businesses, these include 'DigiD and 'eHerkenning' in the Netherlands, the 'Buergerkarte' in Austria, 'Identity Malta' in Malta and 'SuisseID in Switzerland'. Esposito, S. (2015).

Notably, regardless of European regulations, there are differences between European countries in how they organize their identification infrastructure. For example, to identify citizens, the Dutch government uses a national identification number called BSN, which can be used exclusively in the public sector. On the other hand, Germany and the United Kingdom have no such number: they use different attributes. Meanwhile, the Estonian government is quite advanced in terms of identification infrastructure and has launched an e-citizenship model which has been adopted nationwide and is considered a potential standard for other EU countries (Dijck & Jacobs, 2019).

### 2.7 Factors that influence the take up of eID

The uptake of eID systems is dependent on socio-technical factors. Bruggers et al. suggest a number of influencing factors over the perceptions of stakeholders; "legal regulators, political and cultural structures, economic and market structures & technical pool". According to Patnaik, Sachidananda & Scholar (2014), trust management has emerged as a key enabler for the collaboration of eServices. Thus, it has gained attention in e-governance services. A study by Karin Axelsson and Ulf Melin (2016) addresses the development of electronic identification (eID) for public e-services and reports from an empirical study of young Swedish university students' attitudes towards eID. The data gathered indicates

two major themes; usability and security in the adoption and use of eID systems. "The perceived level of trust in the eID influenced whether the respondents in our study choose to use the e-service or not. If they cannot remember their password or find it too cumbersome to download a new certificate, they use another communication channel when interacting with the agency" (Axelsson & Melin, 2012).

In the article "Raising Acceptance of Cross-Border eID Federation by Value Alignment" findings derived from STORK 2.0 framework indicate four main causes for users not making use of eID including high price, high complexity and obstacles for acquiring an eID along with a lack of trust in the solution. "These issues must be mainly solved in the national solutions, but are also impacting the interoperability layer: The complexity of the solution must be completely hidden from the end-users while fostering trust at the same time" Furthermore, "Simplicity and user-friendliness of the user interface are one key to grow trust" (Brugger, Fraefel, & Riedl, 2014). Behind E-Government initiatives is a common goal to aim to deliver benefits to improving transparency, efficiency, trust, and citizen participation. However, e-government initiatives face several barriers. One of them is poor usability. (Lyzara, R., et al., 2019). The subject of usability in e-Government is a well-researched topic. None. Having said that, the majority of research done has been on the usability of more simple e-Government websites. The usability of five different e-Government websites has been evaluated in the journal "Usable guidelines for usable websites? An analysis of five e-government heuristics" (Donker-Kuijer, Jong, & Lentz, 2010).

Although much work has been done towards improving these websites the use of eID services have yet to be studied. According to Thomas et al (2012) "Undoubtedly, the usability of e-Government websites is an important topic that needs further exploration. However, techniques to integrate qualified electronic signatures into web-based e-Government applications definitely can not be forgotten" By not addressing usability, it will be difficult to achieve the interoperability and advancement of digitalisation Europe with the goal of creating a single market that can easily facilitate the movement of people and services "Thus, usability evaluations of current e-Government solutions threaten to remain incomplete and to miss relevant aspects that could help Europe develop better interoperability between government portals" (Thomas Zefferer, Vesna Krnjic, 2012).

### 2.8 The role of trust and usability

A study on the perceptions of citizens from the U.K. and Germany on the subject of interoperable electronic identity (eID) systems suggests that the perceived risks are derivative from poor trust in public authorities that are responsible for identity management (Halperin, Ruth & Backhouse, James, 2012). According to a Survey on EU Citizens Trust in ID Systems and Authorities "while there are benefits to be gained from putting trust in such a scheme, such as more efficient travel and convenient health care, if governments of EU Member States do not support a secure and reliable system, users will be faced with a number

of identity-related risks. In particular, risks associated with privacy loss and identity fraud (Halperin & Backhouse, 2012). To back up these findings, It is not only the technical interoperability of electronic IDs which today makes many electronic transactions impossible. The main obstacle to global interoperability of electronic IDs is the lack of trust in digital certificates issued by various Certification Authorities (CAs).

Another notable issue is the "lack of harmonized regulatory framework for developing eID systems. Interoperable electronic identities, at the European level, have been deemed essential for achieving the freedoms of establishment and circulation of goods, capital and services" (Gomes de Andrade, Monteleone, & Martin, 2020). When it comes to the majority of prior research has applied usability of eGovernment websites rather than on eID. However, a survey on the Dutch identity management system DigiD showed that approximately 75 percent of the participants in their survey have DigiD accounts. Those who claimed not owning such an account underscored that they did not find the necessity of having one or that they preferred to transact with government agencies through their physical outlets and not electronically.

According to Geest et.al (2011), none of the participants stated a lack of trust in DigiD for not owning an account. However, the study cites hypotheses for the role of trust and user experience; internet users' trust in DigiD positively impacts their intention to continue using, use, or apply for DigiD . "Internet users' trust in government organizations necessitating the usage of DigiD positively impacts their intention to continue using, use, or apply for DigiD . Internet users' perceptions of the advantages of using DigiD positively impact their intention to continue using, use, or apply for DigiD. The Internet users' perceptions of the problems in using DigiD can negatively impact their intention to continue using, use, apply for DigiD" (Geest, Thea & Beldad, Ardion, 2011).

### 2.9 Description of the Dutch DigiD system

DigiD is the Dutch governmental identity management system; the system is managed by Logius, a subsidiary of the Dutch Ministry of the Interior and Kingdom Relations. DigiD is an electronic identity management system that can be used to access government services and also to a select set of private companies (for example, health insurers and pension funds). In terms of the adoption and use of DigiD the uptake has been notably high. The number of DigiD users was recorded at 3 million in February 2007(Pimenidis et al 2007). DigiD has been optimised for mobile use, as a means of user authentication the application is recommended to safeguard the users data.

– Accessing privacy sensitive information through the DigiD application:

In order to access privacy sensitive information, users need to authorise an 'ID check' to your DigiD app. Certain data is particularly privacy-sensitive, such as information about the users health. To be able to view or change this data online via mobile the user is required to add a once-only ID check of their

passport, driving license or identity card to your app. This will allow you to do even more with your DigiD.

– Accessing DigiD abroad:

DigiD has been optimised for cross-border mobility. Users with Dutch nationality or the nationality of another country in the European Economic Area (EEA) can apply for a DigiD. Users who already have DigiD before moving abroad can continue to use the DigiD application.

### 2.10 The move towards eID mobile authentication

Mobile access to online government services is growing in popularity, public administrations now must try to improve the accessibility of their online services and work towards a better user experience. a trends report on electronic identification carried out by Diollete on behalf of Europa states that "Embracing mobile identity for eGovernment" an overview was provided of various approaches to enable mobile eID, and the key trends witnessed in countries across the EU. The paper provides an overview of the main technical options, moreover, it also provides a set of recommendations for EU administrations on how to improve the mobile experience of eID and cross-border authentication, the recommendations are:

– To encourage service providers in developing mobile applications to access to their services;
– To ensure that service providers websites are responsive for mobile devices
– To develop a mobile-friendly country selector for eIDAS - i.e. the interface by which users select the eID means they wish to use;
– To ensure that the eIDAS nodes' interfaces are mobile friendly;
– To test the overall mobile journey - in order to identify any pain points. (Eichholtzer, 2020)

### 2.11 European Union benefits of eID interoperability

Interoperability within local, national and European public administrations facilitates the achievement of the goals set out by the European Parliament in its resolution of 29 March 2012 on the EU Citizenship Report 2010: Dismantling the obstacles to EU citizens' rights. (Decision (EU) 2015/2240 of the European Parliament). eID offers both societal and economic benefits to European citizens. Firstly, eID provides the means to interact and transact remotely with institutions and also enables the end-user to access a wider range of e-Services without the need to physically visit any premises. Moreover, eID based services will help with both the accuracy and convenience of accessing services, with twenty-four seven availability and accessibility from any location.

Electronic identity undoubtedly can be considered a key driver for the growth of the EU economy and the completion of the Single Digital Market. eID constitutes not only an essential enabler for the deployment of cross-border services

within the EU but also an irreplaceable technological asset for the increase of entrepreneurial activities in Europe. As observed in the Digital Agenda, "electronic identity (eID) technologies and authentication services are essential for transactions on the internet both in the private and public sectors." (Nuno Gomes de Andrade, Monteleone, & Martin, 2013).

### 2.12 Text analysis as a method for gaining both qualitative and quantitative user experience data

Human Computer Interaction has become significantly interested in using machine learning (ML) to improve user experience (UX). However, some design researchers claim that there is a lack of design innovation in envisioning how ML might improve UX (Graham Dove, et.all, 2017). Current studies into the user experience (UX) and/or Human Computer Interaction (HCI) have studied narratives, accounts from users' on their experience when interacting with technology. Such papers have emphasized specific constructs (e.g., hedonics, affect, needs,) and their interrelation, however, few have analyzed the content of the narratives. In the paper a multi-method was applied, this approach consisted of manual (structural analysis of narratives) as well as of automated content analysis methods (machine learning).Their findings proved that technology is positively experienced when it enables users to do things more efficiently or in a new way. In contrast, negative narratives often express anger and frustration due to technological failures. This multi-method approach illustrates the potential of automated (as opposed to manual) content analysis methods for studying text-based experience reports (Alexandre N. et,al. 2013)

### 2.13 The impact of Machine Learning on evaluating user experience data

The development of the web 2.0 has coincided with a plethora of people voicing their opinions online, be it a twitter tweet, article or review. The data left by users can be used by corporations and government entities alike to evaluate the user's opinions, topics and more using natural language processing. In this study, the use of reviews will be used, even though most application stores and websites that provide consumer insights provide a star rating, according to this result can not take into account the actual content being left, therefore, evaluating the user experience based on star rating omits much of the users wrote content, and this data is hard to evaluate without using means of text analysis. Ganu provides the following example:

"40 reviews have a 4 or 5 star rating (out of 5 possible stars). Majority positive reviews, however, praise the ambience of the restaurant, as shown in the following sentences extracted from the reviews:

- "obviously it's not the food or drinks that is the attraction, but the burlesque show"

– "The food was okay, not great, not bad.[...]Our favorite part, though, was the show!"

The negative reviews complain at length about the price and service. A user not interested in ambience would probably not want to dine at this restaurant. However, a recommendation based on star ratings would label this restaurant as a high-quality restaurant."(A Ganu, G., Elhadad, N., & Marian, A.,2009).

This paper concludes on this point that the user experience can be improved if the review content and it's structure was taken into account. Ganu provides another case to back up this point:

" i.e., if review parts pertaining to different product features (e.g., food, ambience, price, service for a restaurant), as well as the sentiment of the reviewer towards each feature (e.g., positive, negative or neutral) were identified. This information, coupled with the metadata associated with a product". (A Ganu, et.all.,2009).

Machine learning has a wide range of potential analysis techniques that can be deployed. However, Sentiment analysis is cited to be particularly good for obtaining the opinions of user reviews . The sentiment analysis technique set's out to understand the underlying sentiment behind the content within the review, whether a user's opinion is negative, positive or neutral. Since the start of the decade, Sentiment analysis has been a popular method which is in turn fueled by the desire to understand and classify user reviews.. (Batra, S, et.all., 2010).

### 2.14 Final considerations

Taking a holistic approach to the literature reveals the intricacies of the European plans for the implementation of eID and their cross-border functionality and user experience affecting the possibility for interoperability between eID systems. Although we have extensive regulations set out by the european union (eIDAS) there is little research into the user experience. As authentication becomes more reliant on mobile applications to access eGovernment services, interaction design should be taken into account to help a heterogeneous experience for EU citizens. There has been a move towards mobile eID systems as a secure means of two step authentication. The majority of applications used for eID will store reviews on app stores, this provides opportunity for using text analysis methods, such as content or by deploying Natural Language methods such as Sentiment analysis can provide usability insights. Analysing text and using natural language processing techniques is a new practice for User Experience and the area of Human Computer Interaction (HCI). However, with the plethora of users reporting their experiences online this would be a missed opportunity to not avail of the data to obtain opinions from users, also star ratings are unable to take into account the actual content, text analysis and natural language processing methods can be used to gauge the frequency of topics occurring in the text and the sentiment, whether positive, neutral or negative.

# 3 The Study

## 3.1 Purpose of the study

The objective of this mixed methods study is to prove the value in gathering and analysing user experience data of eID applications for the purpose of applying theses insights to improve the user experience. With this goal in mind, this study will attempt to gain a better understanding of the user experience when interacting with the DigiD e-identification application. This will be done by obtaining and analysing reviews from the Google and Apple app store. The first part of this study will be a content analysis, this will provide key themes based on the users reviews and the frequency at which they occur. The second part will be a Sentiment Analysis on a larger data set which will inform us if the overall user experience is negative, neutral or positive alongside the location of the users, which will help also to draw some conclusions on the cross-boarder functionality of the DigiD application.

## 3.2 Content analysis

According to Huang, Z. (2006) for government service using internet media, content analysis presents an effective computational ethnography methodology for evaluating e-government practice. In the case of this study, the goal is to analyse user feedback to gain insights into their experience using the Dutch DigiD e-ID service. To carry out the content analysis, raw user feedback data was required, to gather it in an unobtrusive, unbiased manner random sampling of Google Play Store and Apple App Store reviews for the DigiD mobile app was completed. DigiD is an identity management platform in the Netherlands and follows the eIDAS regulations, the data was obtained from ratings and comments left on the Google and Apple app stores. The motive was to produce a set of key insights into the end users motives for using eID and what issues they run into that affect the usability, adoption and use of eID applications. Furthermore, this approach should highlight the value of gathering user experience data, where most frequently occurring issues could be used to focus the development process of eID applications to improve the user experience, for example if 10% of the users are having issues logging in with a second account, with over 1,000,000 installations on Android alone, this means over 100,000 users may be unable to access critical eServices and additional effort to resolve this issue would improve the experience for a significant amount of people.

**Process** To complete the qualitative content analysis; 100 random reviews were taken from the Apple and Google App store for DigiD and a code book was developed. Codes are defined as ‘ ‘tags or labels for assigning units of meaning to the descriptive or inferential information compiled during a study’ ’ (Miles & Huberman, 1994). The objective of this qualitative content analysis is to systematically transform a large amount of text into an organised and concise summary of key results. Analysis of the raw data transcribed from websites to

form categories or themes is a process of further abstraction of data at each step of the analysis; from the manifest and literal content to latent meanings. The content analysis followed a five step process to obtain the correct information:

1. Identify and collect the needed data – A random sampling technique was used to obtain key insights from stakeholders who have left reviews or comments regarding the use of (i) DigiD application on the Apple and Google Appstore. 50 reviews were obtained from each Appstore.

2. Decide on coding categories – Categories are patterns or themes that are directly expressed in the text or are derived from them through analysis. Once the 100 reviews were obtained, the data was read through and central coding categories were drawn up.

3. Code the content – This was achieved by creating a code book (see Appendix A) by dividing up the text into units and then condensing meaning from these units into themes

4. Ensure validity and reliability of the content – In order to do this, an investigation was be carried out on the levels of abstraction, from manifest to latent content. Further abstraction was taken on content with latent meaning and a meaning was given to the unit

5. Analyze and present the findings – In this case the code book will be thematically analysed. Each code will present a theme that expresses the end users intentions based on the obtained data.

6. Find the Frequency – Record the frequency of a central theme occurring throughout the reviews, the higher the Frequency, the more important the theme.

A hands-on guide to doing content analysis Christen, Erlingsson (2017) provides a glossary of terms as used in this hands-on guide to doing content analysis, which was implemented for this thesis:

- **Condensation** – Synopsis of content, reading through and understanding the general latent and manifest content.
- **Code** – A label that describes what the unit is about.
- **Category** – A group of codes that relate to each other's context or content.what is visible and obvious in the data.
- **Theme** – Is an understanding of the underlying meaning, referred to as latent content, recurring in the thesis. Themes are expressing data on an interpretative (latent) level. Themes address questions such as why, how, in what way? A theme is intended to communicate with the reader on both an intellectual and emotional level. Therefore are descriptive and are made up of verbs, adverbs and adjectives

### 3.3 Sentiment analysis

While content analysis allowed for a deeper insight into key issues highlighted by users, it is a slow and very manual process that could be difficult to apply at scale, on a large dataset. Sentiment analysis is machine-learning-driven, natural language processing method that assists in deciding whether a piece of text (review, tweet, etc.) is positive, negative or neutral. It can be used to identify

the customer or follower's attitude towards an entity through the use of variables such as context, tone, emotion, etc. Stakeholders developing eID applications can use sentiment analysis on reviews left on App Stores to research public opinion and analyze customer satisfaction. The goal of a more automated approach with sentiment analysis is to gather significantly larger amounts of user feedback in an unobtrusive manner and gain higher-level insights into their experience. With these high level insights, how can they be applied to help with decision making when developing eID applications? If the majority of the users feel nagatively about DigiD, how can sentiment analysis help discover key areas of focus?

**Process**

1. A Python script *scraper.py* (see Appendix B) was written for the purpose of retrieving a large amount of reviews from each of the app stores, taking advantage of existing open source libraries - google-play-scraper and app-store-scraper for Google Play Store and Apple App Store respectively. Another open-source library – googletrans, was included for the purpose of translating review text to English.
2. The scraping script is run for each of the platforms individually, retrieving 500 latest reviews from each of the sources, transforming the data into a consistent format, translating the review text and finally storing the results into a CSV file, resulting with *android.csv* and *ios.csv* files.
3. A second, *analyzer.py* (see Appendix C) Python script was created following Google Cloud Sentiment Analysis Tutorial, with adjustments to enable it to read a provided CSV file in the format produced in the scraper script, go over each line of the file and pass it to the Sentiment Analysis API as an individual document. The overall document sentiment for each review was then extracted from the response, taking two values – *score*, which ranges from -1 to +1; -1 being negative and +1 being positive (recorded as *sentiment*); and *magnitude*, which expresses how strongly the given sentiment was expressed. The two values were then appended to each equivalent row in the CSV document and saved as a separate file. This process was completed for both Google Play Store and iOS App Store review data.
4. Both files were then imported into Google Drive and converted to Google Sheets spreadsheet documents. They were then combined into a single spreadsheet for easier analysis across the full dataset.
5. Some of the reviews failed to be translated in the initial scraping script and the native language could not be detected by the sentiment analysis, resulting in 0s for both sentiment score and magnitude. These results were removed, reducing the dataset from 1000 to 911 reviews.
6. A further 44 reviews that were dated before 1st of January 2020 were also removed from the dataset to avoid reviews for much older versions of the app affecting the accuracy of the current sentiment towards DigiD app; resulting with the final set being 867 reviews.
7. Following Google Cloud Natural Language API Basics guide, thresholds were set to interpret and group the sentiment score and magnitude values. The

data was sorted a number of times based on both of these values as well as review score and a number of sample reviews were taken from the dataset for manual content analysis. This helped establish thresholds for clearly positive, likely positive, clearly negative, likely negative, mixed, and neutral sentiment groups (see Table 1).

8. Based on the above thresholds, the number of reviews falling into each category was calculated in the spreadsheet. About 10% of the dataset did not fit into any of the aforementioned categories and these reviews were counted in a separate "Unclear" category. Another similar calculation was made, aggregating the number of reviews based on the user rating, i.e. how many users rated 1-star, 2-stars, etc.
9. Bar charts (see Figure 1 and Figure 2) for both sentiment analysis and user rating were created using the calculated data.

| Sentiment | Min. Score | Max. Score | Min. Magnitude | Max. Magnitude |
|---|---|---|---|---|
| **Clearly Positive** | >=0.5 | <=1 | >0.2 | Infinity |
| **Likely Positive** | >0.1 | <0.5 | >0.2 | Infinity |
| **Mixed** | >=-0.1 | <=0.1 | >=1.5 | Infinity |
| **Neutral** | >=-0.1 | <=0.1 | >=0 | <2 |
| **Likely Negative** | >-0.3 | <-0.1 | >0 | Infinity |
| **Clearly Negative** | >=-1 | <=-0.3 | >=0.7 | Infinity |

**Table 1.** Thresholds defined for classifying sentiment analysis results

## 4 Results and Discussion

### 4.1 Content analysis

The content analysis method provided key themes throughout the 100 reviews that were analysed. These themes were aggregated by order of frequency, highlighting which were the most common topics, as shown in Table 2. These themes, or results from a similar content analysis could be used by the developers of eID applications and regulations such as DigiD and eIDAS to find areas of focus. A significant advantage of content analysis is that it can use already publicly available user feedback, expressed in free form and not constrained by specific survey questions which could result in biased answers or missing key issues if the correct questions are not asked. One of the most recurring themes is a lack of clear instructions, with the theme showing up in 17% of analysed reviews. With over 1 million DigiD installations from Google Play Store alone, this means over 170,000 users could be experiencing issues due to lack of instructions. DigiD could use this data to focus their efforts on improving this specific aspect of the application, and as a result improve the experience for a significant portion of their users. This shows the value of using content analysis to analyse user feedback, as key issues can be identified and prioritised.

### 4.2 Sentiment analysis

After carrying out sentiment analysis of over 800 of the most recent reviews, it was revealed that the majority of the reviews had a negative sentiment, with 54.4% classified as clearly negative and 5.7% classified as likely negative. This correlated with the star ratings associated with these reviews with 55.9% 1-star reviews and 8.7% 2-star reviews, with the correlation visually apparent in the bar charts (see Figure 1 and Figure 2) produced. This contrasts with the current public ratings of 4.2 on Apple App Store and 4.1 on Google Play Store, which aggregate reviews over the lifetime of the application, indicating that the negative user opinion is a more recent development. Sentiment analysis proved

| **Code** | **Theme** | **Frequency** |
|---|---|---|
| Clear instructions | A lack of clear instructions was one of the areas of impact on the user's accessibility and ease of use. | 17.0% |
| Ease of authentication | User's struggled with the method of authentication, QR code scanning. | 13.0% |
| Cross device ability | Users are having difficulty understanding and/or using the cross device functionalities for authentication. | 10.0% |
| Cross border ability | Users can access the DigiD app and use it in another EU country | 9.0% |
| Opinions on DigiD | Users opinion on the developments and working of DigiD app. | 8.0% |
| Issues after update | Issues occurring after updates negatively affect the users opinion on DigiD and the Government. | 7.0% |
| Trust in government | Users have high expectations from government developed applications and are affected by pre-formed opinions on the government. | 7.0% |
| Somaesthetics | Ease of embodied interaction and how the user thinks they should interact with the application. | 7.0% |
| Perceived need | Users see a need for the DigiD application. However, users should not feel obliged. | 5.0% |
| Perceived risk | If the perceived risk is low, users will use DigiD as a more secure way to access government services. Those who perceived the risk as high also lacked trust in the government. | 4.0 |
| Assign multiple users | Multiple users, ease of authentication and identification of each user's data is required. | 3.0% |
| Users age | Older users found the application less accessible and complicated to use. | 2.0% |

**Table 2.** Key findings from the content analysis

to be a great method for a quick, unbiased analysis of a large amount of data, in this case it was 1000 most recent reviews, however it could be applied to specific cohort, platform or other more focused area to extract more specific results. While positive results might not provide as much value, an overall negative sentiment indicates the presence of issues with DigiD application, informing DigiD developers that despite the current overall rating in the app stores showing relatively high, there majority of recent reviews and users' sentiment is negative, highlighting the need to resolve issues users may be experiencing. Although sentiment analysis carried out in this study did not identify *which* issues should be prioritised, content analysis could be used to achieve this goal, or additional natural language processing could be performed to extract the most recurring issues.

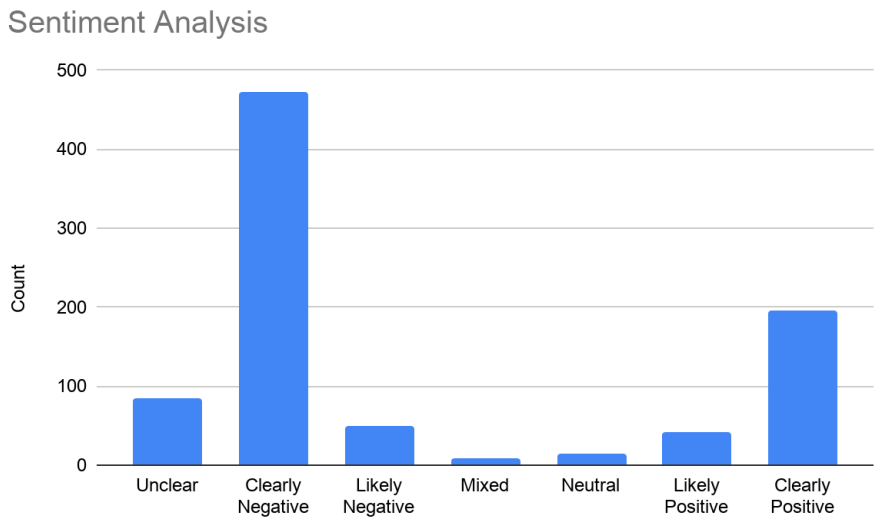


**Fig. 1.** Results of Sentiment Analysis grouped by sentiment.

### 4.3 Discussion

This thesis set out to investigate what is the value in gathering user experience data from eID applications, what methods can be applied to gather and analyse this data and how can we use these insights to shift discussions around eID from technical to a sociotechnical outlook when developing applications.

This case study explored how both manual and automated computational ethnography methods could be applied to evaluate user feedback for the Dutch eID application - DigiD, how this data could be evaluated using content analysis

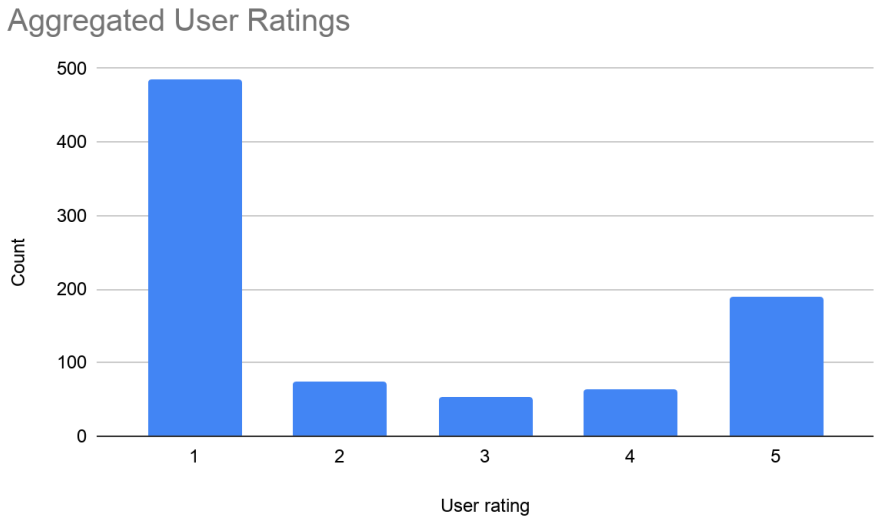


**Fig. 2.** User ratings out of 5 aggregated by count

and sentiment analysis, and how the derived insights could be applied to fixing or improving the DigiD app.

Delving deeper into the results of both methods, content analysis provided a better idea of specific areas of focus required, for example one of the most recurring themes was a lack of clear instructions, we can therefore conclude that an important user experience improvement would be to update and make the instructions for using the application clearer. On the other hand, sentiment analysis did not provide such easily identifiable issues, however it showed that there is an overall negative sentiment towards the app and in particular this seems to be a more recent development as the resulting ratings from 2020 were below the overall app rating. This could possibly be correlated with a theme cropping up in the content analysis where over 7% of the users reported having issues after updating the app. A possible recommendation could be made to improve the quality assurance practices for the DigiD app, which would result in an improved user experience.

The sentiment analysis also complements the content analysis where the theme does not always make it clear whether it was recurring due to positive or negative user feedback, for example cross-border ability showed up in 9% of the analysed reviews, however without looking at the individual reviews it is not clear if they were negative or positive. With sentiment analysis producing a majority negative result, it can be deduced that this theme is discussed in a negative manner in the majority of the reviews. This insight allows us to suggest to the DigiD developers to further investigate and resolve cross-border functionality issues to improve the user experience.

Overall, this case study has shown that it is indeed possible to apply computational ethnography to gather and evaluate user experience data and gain insights which can then be applied to making improvements to eID applications. However, it's important to choose the right tools or the right combination of tools for a specific scenario, for example using a deeper, per-theme sentiment analysis and only selecting negative results to find key user pain-points, which then indicates what issues should be fixed first.

### 4.4 Limitations of the study

The first study was a manual content analysis, which did not use any computation, data processing or machine learning, so it could be argued that reviews were subjected to increased chances of human error, particularly when relational analysis is used to attain a higher level of interpretation. Content analysis, while useful for identifying key recurring themes in user feedback, is often devoid of theoretical base, or attempts too liberally to draw meaningful inferences about the relationships and impacts implied in a study, as well as taking up a significant amount of time to manually analyse and collate the data.

The second method applied was sentiment analysis, which deployed machine learning to gauge users' opinions, there was no human interference with the data, aside from removing invalid data that could not be processed, so therefore it is not subjected to human-biased interpretation of the review content, resulting in conclusive data that can not be argued. The only manual interpretation required is setting thresholds of what is considered a negative or a positive sentiment based on its score and magnitude.

A possible future study could also use natural language processing instead of manual content analysis and identify key themes/topics by creating a topic model, which is a type of a statistical model for discovering abstract topics by analysing semantic structures in a text body. Furthermore, the sentiment analysis could be carried out not just for each review as a whole, but for each recognised topic or "token" and collated to get users' sentiment towards a specific theme, for example, gauging users' sentiment to "eID" or "Dutch government" entities.

### 4.5 Conclusions

This study tried to prove the value of gathering user experience data for eID applications - how that data can be collected in an unobtrusive, unbiased way, how can computational ethnography and modern data analysis methods be applied to analyse this feedback, and finally how can these findings bring value - how can they be applied to improve eID applications and regulations to not only focus on the technical, but also consider the sociotechnical aspects, in particular ensuring a consistent, positive user experience.

Both content analysis and sentiment analysis methods used in this study used existing, public, and readily available user feedback from Google Play Store and Apple App Store for the DigiD app. This shows that it is possible to gather

user feedback in an unobtrusive manner as users themselves have volunteered to provide it without being prompted. Furthermore, this method allows the users to freely express their views and does not introduce bias like, for example, a survey would by providing restrictive or constrained questions.

The content analysis method was applied to analyse 100 most recent, manually gathered reviews from the aforementioned app stores and provided key themes recurring in these reviews, with majority highlighting issues, rather than positive features. The most recurring issue was found to be the lack of clear instructions, with the theme showing up in 17% of analysed reviews. With over 1 million DigiD users on Android alone, over 170,000 users could be experiencing this issue. Findings such as this can provide value to DigiD or other eID providers, by using this data to prioritise key issues, and as a result improve the experience for a significant portion of their users.

The sentiment analysis method analysed over 800 most recent app store reviews, indicating a majority negative sentiment. While on it's own this may not seem like a valuable result, this indicates that despite relatively high overall ratings of 4.2 on Apple App Store and 4.1 on Google Play store, the most *recent* user sentiment has been negative. These findings can provide value by indicating a presence of significant issues that could be prioritised to improve the user experience. While the specific issues themselves are not immediately available following sentiment analysis, this approach can be combined with content analysis, for example performing it on the reviews with the most negative sentiment, or additional natural language processing methods could be used to extract the most common issues.

It has become common practice for people these days to take to the internet, be it Twitter, Instagram, Facebook or the app store, to share their experiences and leave reviews. With slow-to-respond government organisations, however mundane it might seem, people feel compelled to share their experience. This content should not be dismissed, considering the difficulty of testing e-Government services, as user reviews provide an easily available, reusable and unbiased source of user feedback. This data could provide significant value to eID developers to find specific areas to focus on and issues to resolve, to improve the user experience for significant portions of their user base, to ensure wider usability by resolving critical issues which prevent users from using their services, and in turn drive adoption of eID services reducing the overhead on in-person or on-paper bureaucratic processes. Furthermore, these insights could be used by governing bodies to improve their regulations, for example results highlighting an overall subpar user-experience could help shift the discussion about eIDAS regulation from entirely technical to sociotechnical, adding guidelines and requirements not just on how the authentication methods are implemented, but how the resulting mechanisms are presented to the user and how they are able to interact with it, resulting in a better user experience in all eID applications across the European Union.

## References


1. Alexandre N. Tuch, Rune Trusell, and Kasper Hornbæk. 2013. Analyzing users' narratives to understand experience with interactive products. In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems (CHI '13). Association for Computing Machinery, New York, NY, USA, 2079–2088. https://doi.org/10.1145/2470654.2481285
2. Arnold, T., & Fuller, H. (2018, July 21). In Search of the User's Language: Natural Language Processing, Computational Ethnography, and Error-Tolerant Interface Design. https://doi.org/10.1007/978-3-319-94947-5$_4$.
3. Axelsson, K., & Melin, U. (2012). Citizens' Attitudes towards Electronic Identification in a Public E-Service Context – An Essential Perspective in the eID Development Process. Retrieved, from https://www.researchgate.net/publication/266864117_Citizens'_Attitudes_towards_Electronic_Identification_in_a_Public_E-Service_Context_-_An_Essential_Perspective_in_the_eID_Development_Process
4. Björneborn, Lennart. (2011). Behavioural traces and indirect user-to-user mediation in the participatory library
5. Brugger, J., Fraefel, M., & Riedl, R. (2014). Raising Acceptance of Cross-Border eID Federation by Value Alignment. E-Government Institute, Berne University of Applied Sciences, Berne, Switzerland, 261–269. Retrieved, from http://www.ejeg.com/issue/download.html?idArticle=352
6. Corallo, A., Fortunato, L., Matera, M., Alessi, M., Camillò, A., Chetta, V., . . . Storelli, D. (2015). Sentiment Analysis for Government: An Optimized Approach. Machine Learning and Data Mining in Pattern Recognition Lecture Notes in Computer Science, 98-112. https://doi.org/10.1007/978-3-319-21024-7_7
7. Cuijpers, C., & Schroers, J. (2014). eIDAS as guideline for the development of a pan European eID framework in FutureID. EIDAS as Guideline for the Development of a Pan European EID Framework in FutureID, 1–10. https://core.ac.uk/display/34614563
8. De Andrade, N. N. (2013). Electronic Identity for Europe: Moving from Problems to Solutions. Journal of International Commercial Law and Technology, 8(2), 104-108.
9. December 1999 on a Community framework for electronic signatures THE EUROPEAN PARLIAMENT AND THE COUNCIL OF THE EUROPEAN UNION
10. Decision (EU) 2015/2240 of the European Parliament and of the Council of 25 November 2015 establishing a programme on interoperability solutions and common frameworks for European public administrations, businesses and citizens (ISA2 programme) as a means for modernising the public sector (Text with EEA relevance)
11. Digital Government Factsheets - 2018. (2018). Retrieved from https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/digital-government-factsheets-2018
12. Dijck, J. V., & Jacobs, B. (2019). Electronic identity services as sociotechnical and political-economic constructs. New Media & Society, 22(5), 896-914. https://doi.org/10.1177/1461444819872537
13. Donker-Kuijer, M. W., Jong, M. D., & Lentz, L. (2010). Usable guidelines for usable websites? An analysis of five e-government heuristics. Government Information Quarterly, 27(3), 254-263. https://doi.org/10.1016/j.giq.2010.02.00
14. Dumais, Susan & Jeffries, Robin & Russell, Daniel & Tang, Diane & Teevan, Jaime. (2014). Understanding User Behavior Through Log Data and Analysis. https://doi.org/10.1007/978-1-4939-0378-8$_14$.

15. Eichholtzer, M. (2020). Embracing mobile identity for eGovernment. Retrieved from https://ec.europa.eu/cefdigital/wiki/display/CEFDIGITAL/2020/05/06/Embracing+mobile+identity+for+eGovernment
16. eIDAS: Overview on the implementation and uptake of Trust Services. (2018, January 24). Retrieved from https://www.enisa.europa.eu/publications/eidas-overview-on-the-implementation-and-uptake-of-trust-services
17. eIDAS. (2019, November 19). Retrieved from https://www.digitaleoverheid.nl/dossiers/eidas/
18. Engelbertz, N., Erinola, N., Herring, D., Somorovsky, J., Mladenov, V., & Schwenk, J. (nd). Security Analysis of eIDAS – The Cross-Country Authentication Scheme in Europe. Chair for Network and Data Security, Horst Görtz Institute for IT-Security Ruhr University Bochum.
19. Esposito, S. (2015). Secure idenTity crOss boRder linKed-eLearning and Academic Qualifications Pilot at IST.
20. European Commission, ISA: European Interoperability Framework (EIF) for European public services, Annex 2. pp. 2-21 (2010).
21. European Commission, ISA: European Interoperability Framework (EIF). p.5-37 (2017) European Commission, ISA: The ISA programme: Overcoming eBarriers to European public services. p2 (2011).
22. European Union, and European Parliament. "Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on Electronic Identification and Trust Services for Electronic Transactions in the Internal Market and Repealing Directive 1999/93/EC, CELEX1." Publications Office of the European Union, Publications Office of the European Union, 23 July 2014, https://op.europa.eu/en/publication-detail/-/publication/23b61856-2e82-11e4-8c3c-01aa75ed71a1/language-en
23. F. Nayebi, J. Desharnais and A. Abran, "An Expert-Based Framework for Evaluating iOS Application Usability," 2013 Joint Conference of the 23rd International Workshop on Software Measurement and the 8th International Conference on Software Process and Product Measurement, Ankara, 2013, pp. 147-155, https://doi.org/10.1109/IWSM-Mensura.2013.30.
24. Foundation, P. by D. (n.d.). Choose IRMA. Put a digital passport on your own mobile. Retrieved from https://irma.app/
25. Ganu, G., Elhadad, N., & Marian, A. (2009). Beyond the Stars: Improving Rating Predictions using Review Text Content. WebDB.
26. Geest, Thea & Beldad, Ardion. (2011). Citizens' trust in Dutch e-government and DigiD.
27. Gomes de Andrade N., Monteleone, S., & Martin, A. (2020). Electronic Identity in Europe. Legal Challenges and Future Perspectives, 11-57. https://doi.org/10.5771/9783748900092
28. google-play-scraper. (n.d.). Retrieved November 10, 2020, from https://pypi.org/project/google-play-scraper/
29. googletrans. (n.d.). Retrieved November 22, 2020, from https://pypi.org/project/googletrans/
30. Graham Dove, Kim Halskov, Jodi Forlizzi, and John Zimmerman. 2017. UX Design Innovation: Challenges for Working with Machine Learning as a Design Material. In Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17). Association for Computing Machinery, New York, NY, USA, 278–288. https://doi.org/10.1145/3025453.3025739
31. Höök, K. (2018). Designing with the body: Somaesthetic interaction design. Cambridge, MA: The MIT Press.

32. Halperin, R., & Backhouse, J. (2012). Risk, trust and eID: Exploring public perceptions of digital identity systems. First Monday, 17(4). https://doi.org/10.5210/fm.v17i4.3867
33. Hansteen, Kjell, Jon, Alvik, & Tor. (1970, January 1). Nordic digital identification (eID) : Survey and recommendations for cross border cooperation. Retrieved from https://core.ac.uk/display/36357362
34. Harbach, M., Fahl, S., Rieger, M., & Smith, M. (2013). On the Acceptance of Privacy-Preserving Authentication Technology: The Curious Case of National Identity Cards. Privacy Enhancing Technologies Lecture Notes in Computer Science, 245–264. https://doi.org/10.1007/978-3-642-39077-7 13
35. Hedström, Karin & Karlsson, Fredrik & Söderström, Fredrik. (2016). Challenges of introducing a professional eID card within health care. Transforming Government: People, Process and Policy. 10. 26-46. 10.1108/TG-10-2013-0044. Huang, Z. (2006). E-GOVERNMENT PRACTICES AT LOCAL LEVELS: AN ANALYSIS OF U.S ... Retrieved from https://iacis.org/iis/2006/Huang.pdf
36. Human-centred data science provides an opportunity for both qualitative and quantitative traditions. Digital ethnography has gained widespread adoption as qualitative researchers adapt traditional ethnographic methods to online spaces. (Dumais, Sesan,.et. all)
37. Lee, Y., Kozar, K.A. and Larsen, K.R.T. (2003) "The Technology Acceptance Model: Past, Present, and Future", (Online) http://cais.isworld.org
38. Lusoli, W., & Miltgen, C. (2009). Young People and Emerging Digital Services An Exploratory Survey on Motivations, . Perceptions and Acceptance of Risks, 1–83. Retrieved from http://ftp.jrc.es/EURdoc/JRC50089.pdf
39. M. Alexander, Thomas Zefferer, Florian Reimair, Cagatay Karabat, and Elif Ustundag Soykan, Leveraging the adoption of electronic identities and electronic- signature solutions in Europe, In Proceedings of the Symposium on Applied Computing, pp. 69-71 (ACM, 2017).
40. MacQueen, McLellan, Kay, & Milstein, 1998 Miles & Huberman, 1984
41. NIFO - National Interoperability Framework Observatory. (n.d.). Retrieved from https://joinup.ec.europa.eu/collection/nifo-national-interoperability-framework-observatory/digital-government-factsheets-2018.
42. Official Journal of the European Communities 19. 1. 2000 DIRECTIVE 1999/93/EC OF THE EUROPEAN PARLIAMENT AND OF THE COUNCIL of 13
43. Patel, V. L., Kannampallil, T. G., & Kaufman, D. R. (2015). Cognitive informatics for biomedicine: Human-computer interaction in healthcare. Cham: Springer.
44. Patnaik, Sachidananda & Scholar, Research. (2014). Trust Management in e-Governance Services. https://doi.org/10.13140/2.1.3345.6961.
45. Pimenidis, Elias & Savvas, Ioannis. (2007). E-identification technologies for e-government interoperability in the EU. International Journal of Electronic Security and Digital Forensics. 1. https://doi.org/10.1504/IJESDF.2007.016866.
46. Public Participation Readiness Toward E-Gov 2.0: Lessons from Two Countries. ICEGOV2019: Proceedings of the 12th International Conference on Theory and Practice of Electronic Governance. 240-243. https://doi.org/10.1145/3326365.3326397. Community participation in e-Government development initiatives..
47. R. K. Bakshi, N. Kaur, R. Kaur and G. Kaur, "Opinion mining and sentiment analysis," 2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom), New Delhi, 2016, pp. 452-455.

48. Regulation (EU) No 910/2014 of the European Parliament and of the Council of 23 July 2014 on electronic identification and trust services for electronic transactions in the internal market and repealing Directive 1999/93/EC. (n.d.). Retrieved from http://www.legislation.gov.uk/eur/2014/910/introduction/adopted
49. Roelofs, F. (2019). Analysis and comparison of identification and authentication systems under the eIDAS regulation.
50. Sentiment Analysis Tutorial  |  Cloud Natural Language API. (n.d.). Retrieved November 22, 2020, from https://cloud.google.com/natural-language/docs/sentiment-tutorial
51. Siddharth Gulati, Sonia Sousa & David Lamas (2019) Design, development and evaluation of a human-computer trust scale, Behaviour & Information Technology, 38:10, 1004-1015. https://doi.org/10.1080/0144929X.2019.1656779 20.
52. Siv E. R. Skard, Herbjørn Nysveen, & Helge Thorbjørnsen. (2016). Influences of Perceived E-Service Risks: The Moderating Role of Perceived Benefits. E-Service Journal, 10(1), 1-23. https://doi.org/10.2979/eservicej.10.1.01
53. Thomas Zefferer, Vesna Krnjic: "Usability Evaluation of Electronic Signature Based E-Government Solutions". In Proceedings of the IADIS International Conference WWW/INTERNET 2012, pp. 227-234, IADIS, 2012.
54. van Dijck, J., & Jacobs, B. (2020). Electronic identity services as sociotechnical and political-economic constructs. New Media & Society, 22(5), 896–914. https://doi.org/10.1177/1461444819872537 22. European Commission. Secure Electronic Identity Across Europe, European Commission – Information Society and Media DG, November (2010)
55. Williams, I., Falch, M., & Tadayoni, R. (2018). Institutional Legitimacy and Digital Public Cross-Border Service Delivery Between Denmark/Sweden and Denmark/Germany. Retrieved April 1, 2020, from https://www.riverpublishers.com/journal_read_html_article.php?j=NBJICT/2018/1/8
56. Yang, B., Liu, Y., Liang, Y., & Tang, M. (2018, December 28). Exploiting user experience from online customer reviews for product design. Retrieved October 05, 2020, from https://www.sciencedirect.com/science/article/abs/pii/S0268401218305437
57. Zheng, Kai & Hanauer, David & Weibel, Nadir & Agha, Zia. (2015). Computational Ethnography: Automated and Unobtrusive Means for Collecting Data In Situ for Human-Computer Interaction Evaluation Studies. https://doi.org/10.1007/978-3-319-17272-9$_6$.
58. Zhou, Xiang. (2004). E-Government in China: A Content Analysis of National and Provincial Web Sites. J. Computer-Mediated Communication. 9. 10.1111/j.1083-6101.2004.tb00297.x.

## Appendix A

Code book developed to systematically analyse the 100 DigiD app reviews obtained from Google Play Store and Apple App Store,

1. **Code:** Clear instructions
   - **Theme:** A lack of clear instructions was one of impact on the user's accessibility and ease of use.
   - **Frequency:** 17
   - **Categories:**
     - The user can find and follow instructions on how to use the DigiD app.
     - The user does not understand the content on the application.
   - **Example reviews:**
     - "If the clear instructions are followed, the app is easy to use"
     - "Too much fuss, unclear instructions"
2. **Code:** Ease of authentication
   - **Theme:** User's struggled with the method of authentication, QR code scanning.
   - **Frequency:** 13
   - **Categories:**
     - The user can use the method of authentication.
     - QR code scanning issues
     - Required two devices.
   - **Example reviews:**
     - "This app makes it easier to authenticate yourself on websites such as the one for my Gemeente"
     - "The QR code is not scanned while other apps on my phone have no problem scanning QR codes"
3. **Code:** Interoperability services
   - **Theme:** DigiD needs to be interoperable with other eGovernment services such as those needed for accessing tax, pension and healthcare information.
   - **Frequency:** 11
   - **Categories:**
     - The ability to access e-government services such as pension, healthcare and taxes.
   - **Example reviews:**
     - "Digid app does not recognize QR code from tax authorities"
     - "Very useful if you need any governmental service in the Netherlands"
4. **Code:** Cross device functionality
   - **Theme:** Users are having difficulty understanding and/or using the cross device functionalities for authentication.
   - **Frequency:** 10
   - **Categories:**
     - Users are having troubles using multiple devices to log in.

    - Experiencing issues with using two devices
    - Confused about using two devices
  - **Example reviews:**
    - "Truly abysmal. It was not designed to run on two devices at the same time."
    - "Terrible integration between the app and the website. Your session on the websites times out every 5 seconds when attempting to log in"
5. **Code:** Cross border ability
  - **Theme:** Users can access the eID app and use it in another EU country
  - **Frequency:** 9
  - **Categories:**
    - Users who were residing abroad accessing the app.
    - Users within the EU having difficulties accessing government services in the Netherlands.
  - **Example reviews:**
    - "I am a Dutch subject living in Spain. I have digid. Unable to activate the app from Spain. In the end I had to travel to the Netherlands to activate it"
    - "NFC does not work, crashes right away. I cannot login with sms-confirmation since i live abroad."
6. **Code:** Opinions on DigiD
  - **Theme:** Users opinion on the developments and working of DigiD app.
  - **Frequency:** 8
  - **Categories:**
    - Users opinions on the app
    - Users mentioning DigiD
    - Users reviews on how the applications development
  - **Example reviews:**
    - "Digid is one of the most annoying tools used by the Dutch government and health insurance companies.",
    - "remains an unbelievable blunder to only come up with a fix after months if an iOS update makes the app unusable. If it were about a thirteen-in-a-dozen photo filter app: but for such a serious app this is not acceptable."
7. **Code:** Trust in Government
  - **Theme:** Users have high expectations from government developed applications and biases from pre-formed opinions on the government.
  - **Frequency:** 7
  - **Categories:**
    - The users opinion of government processes will affect their opinion of the application.
    - Users expect better applications compared to non-government related applications.
  - **Example reviews:**

    - "Typical Dutch governmental IT failure, does not work to scan a QR Code on my Huawei P30 Pro. Just keeps focusing incorrectly."
    - "Good to see 2FA authentication at Dutch government - good effort."
8. **Code:** Somaesthetic
    - **Theme:** Ease of embodied interaction and how the user thinks they should interact with the application.
    - **Frequency:** 7
    - **Categories:**
        - User's are interested in the ways in which they can identify themselves, some suggested using fingerprint and face identification.
        - The way in which the user understands how to use these interactions are important, from how they should hold their phone, to the distance and position.
    - **Example reviews:**
        - "Tried for minutes at different distances and positions and even movements."
        - "Too many steps in the login process. And fingerprint support is missing."
9. **Code:** Perceived need
    - **Theme:** Users see a need for the DigiD application. However, users should not feel obliged.
    - **Frequency:** 5
    - **Categories:**
        - The users would like the application to improve or feel obliged to use.
        - The user feels obliged to use the app.
    - **Example reviews:**
        - "It's a pity that such an important tool is not working as it should be expected."
        - "Terrible app. It feels more and more forced onto Dutch people"
10. **Code:** Issues after updates
    - **Theme:** Issues occurring after updates negatively affect the users opinion on DigiD and the Government.
    - **Frequency:** 4
    - **Categories:**
        - Issues with IOS update.
        - Issues with Android update.
        - Users questioning application testing.
        - Users frustration towards DigiD due to issues with the application updating.
        - Users are frustrated at the Government due to update issues.
    - **Example reviews:**
        - "App not working. I have the latest IOS update but the app does not open."
        - "Does not work after update"
11. **Code:** Perceived risk

- **Theme:** If the perceived risk is low users will use DigiD as a more secure way to access government services. Those who perceived the risk as high also lacked trust in the government.
- **Frequency:** 4
- **Categories:**
  - Does the user trust DigiD with their personal data?
  - Users link their trust in the government with the perceived risks associated with the application.
  - Some users believe the application to be a secure way to access government services.
- **Example reviews:**
  - "A security risk at best, handing out the key to your identity and government critical information is just the top of the iceberg of trouble."
  - "Fast and secured login for my personal docs"

12. **Code:** Assign multiple users
    - **Theme:** Multiple users, ease of authentication and identification of each users data is required.
    - **Frequency:** 3
    - **Categories:**
      - Users who are caregivers may need to authenticate on behalf of another user.
      - Those who file taxes together such as couples require ways to add multiple users.
    - **Example reviews:**
      - "Why can only my husband use the app and not I? I did not have to enter any data when downloading. How come it is automatically in his name?"
      - "A lot of trouble logging in / fails! Only my husband's data, not mine"
13. **Code:** Users age
    - **Theme:** Older users found the application less accessible and complicated to use.
    - **Frequency:** 2
    - **Categories:**
      - Older user are less able to use the application.
      - Older Users can not independently complete government related tasks including accessing the pension website.
    - **Example reviews:**
      - "Couldn't it be easier and easier with the app? people over 40 years old don't understand all of your complicated app"
      - "This simply means that I am shut out of my pension website."

## Appendix B

Python script for scraping app store reviews from Google Play Store and Apple App Store.

```
from google_play_scraper import Sort, reviews # pip install google-play-scraper
from app_store_scraper import AppStore # pip3 install app-store-scraper
from googletrans import Translator # pip3 install googletrans
import sys
import ssl
import csv

def main():
    if len(sys.argv) != 2:
        print("Please pass preferred platform: ios|android")
        exit()

    data = []
    if sys.argv[1] == "ios":
        data = get_ios_reviews(500)
    elif sys.argv[1] == "android":
        data = get_android_reviews(500)
    else:
        print("Platform must be either 'android' or 'ios'")
        exit()

    translateAll(data)

    keys = data[0].keys()
    with open(sys.argv[1] + '.csv', 'w', newline='')  as output_file:
        dict_writer = csv.DictWriter(output_file, keys)
        dict_writer.writeheader()
        dict_writer.writerows(data)

def translateAll(data):
    translator = Translator()
    forTranslation = []
    for review in data:
        forTranslation.append(review["content"])
    translations = translator.translate(forTranslation, src='nl', dest='en')
    for i in range(len(data)):
        data[i]["translated"] = translations[i].text

def get_ios_reviews(count):
    digid = AppStore(country="nl", app_name="", app_id="1208460960")
```

```
    digid.review(how_many=count)

    processed = []
    for review in digid.reviews:
        processed.append({
            "date": review["date"].isoformat(),
            "score": review["rating"],
            "content": review["review"]
        })
    return processed

def get_android_reviews(count):
    ssl._create_default_https_context = ssl.
    _create_unverified_context # ssl fix
    result, _ = reviews(
        'nl.rijksoverheid.digid.pub',
        lang='en',
        country='nl',
        sort=Sort.NEWEST,
        count=count,
    )

    processed = []
    for review in result:
        processed.append({
            "date": review["at"].isoformat(),
            "score": review["score"],
            "content": review["content"]
        })
    return processed

if __name__ == "__main__":
    main()
```

## Appendix C

Python script for passing the CSV review results through Google Cloud Sentiment Analysis API.

```
# Adapted from https://cloud.google.com/natural-language/docs/sentiment-tutorial
from google.cloud import language_v1 # pip3 install --upgrade google-cloud-language
import argparse
import csv

def main():
    parser = argparse.ArgumentParser(
        description=__doc__, formatter_class=argparse.RawDescriptionHelpFormatter
    )
    parser.add_argument(
        "review_filename",
        help="The filename of the review file you'd like to analyze.",
    )
    args = parser.parse_args()

    reviews = parseCSV(args.review_filename)
    analyze(reviews)

    keys = reviews[0].keys()
    filename = args.review_filename.replace(".csv", "") + "_analysed.csv"
    with open(filename, 'w', newline='')  as output_file:
        dict_writer = csv.DictWriter(output_file, keys)
        dict_writer.writeheader()
        dict_writer.writerows(reviews)

def parseCSV(filename):
    with open(filename) as f:
        data = [{k: v for k, v in row.items()} for row in csv.DictReader(f, skipinitialspace=True)]
    return data

def analyze(reviews):
    client = language_v1.LanguageServiceClient()

    for review in reviews:
        document = language_v1.Document(content=review["translated"], type_=language_v1.Document.Type.PLAIN_TEXT)
        annotations = client.analyze_sentiment(request={'document': document})
```

```
37          review["sentiment"] = annotations.document_sentiment.
    score
38          review["magnitude"] = annotations.document_sentiment.
    magnitude
39
40 if __name__ == "__main__":
41     main()
```