

Identification of persons for data processing in social security

Francisco Delgado Azuara
National Institute of Social Security
Spain

Index

1. Introduction.
2. What is identification?
3. Identification processes
 - 3.1. Identity Registration
 - 3.2. Identity Verification on Providing Access to Subsequent Services
4. Data needed to identify an insured person in electronic processes.
5. Study of registration keys:
 - 5.1 Types of registration keys in EU social security institutions
 - 5.2. National keys
6. Identification keys in Spain
 - 6.1 National Identification Number.
 - 6.2 Social Security Number
 - 6.3 Individual's identifier
7. Electronic identity
 - 7.1. Digital certificates
 - 7.2. Electronic ID's
 - 7.3. Electronic Id in Europe: STORK
8. Conclusions
9. References

1. Introduction

Identification has been an issue as long as people have stepped out of their known environments and started interacting elsewhere. Having to prove ones identity is something typical in situations where the other party involved is not aware of your true identity. In the early nineties, when European Union faced the electronic data exchange, one of the main issues was the management of personal identification. Interoperability is only possible if we definitely know, without doubt, the person we are talking about. In this document we are going to share some lessons learned from the European Union's experience and from that of a more in depth look at the Spanish Social Security System.

After this introduction, chapter 2 is devoted to clarifying the identification concept. The next section (chapter 3) tries to set out questions about the identification process. Some answers to these questions are suggested in chapter 4, which describes the data needed to identify an insured person. Some important recommendations are made in this section and the main one is the focus of one part of this paper (chapter 5) which is the study of registration keys. Concentrating on Social Security problems is chapter 6 which looks at the Spanish experience with identification keys. At the end, Chapter 7 briefly presents the digital identity models of Spain and the European Union. The document finishes with some conclusions (chapter 8) and references.

2. What is identification?

Generally speaking, the function of identification is to map a known entity to an unknown entity so as to make it known. The known entity is called the identifier (or ID) and the unknown entity is what needs identification. A basic requirement for identification is that the ID be unique. IDs may be scoped, that is, they are unique only within a particular scope. IDs may also be built out of a collection of quantities such that they are unique on the collective.

Looking for a more specific definition in an Information Technology (IT) environment we could say that: "Identification is the capability to find, retrieve, report, change, or delete specific data without ambiguity". This applies especially to information stored in databases. In database normalization, it is the central, defining function to the discipline. To illustrate this definition some examples could be useful. For instance, in Wikipedia: "To retrieve the data stored about an individual with the Internal Revenue Service (IRS) of the United States, the IRS needs to be supplied with the individual's Social Security number (SSN). The SSN, then, "identifies" a unique individual to the IRS. No other living person has that SSN, and that individual is assumed to not have more than one SSN. Other data that the individual's SSN identifies with the IRS would be such things as his or her name, birth date, and current employer."

This and some other examples bring questions such as the following:

- A worker asks for their pension, how can this person be best identified and thereafter receive her/his contributions and rights due to them?
- If a particular worker worked in more than one country, schema or any other circumstance that means there is relevant information stored in different data bases. How will every institution be able to find this worker in order to provide the information needed to gather the necessary information in order to decide on the request?

On the other hand, using IT systems implies that the messages exchanged can be automatically treated (this means by a computer system without manual intervention) by the receiver of the message. To allow such treatment, the data in the message must be precisely defined and the computer must be able, with this data to, automatically identify the insured person concerned by the request. This is why the identification processes should be studied carefully.

3. Identification processes

In the framework of electronic message exchange between social security institutions, the identification of insured persons by automatic processes has been a priority for the European Union. As it has said before, only a correct identification scheme of the insured person allows interoperability of Member states IT systems.

In the early nineties, the European Commission promoted the electronic exchange of Social Security information between Member states. Some working groups have been established in related projects on electronic exchanges (SOSENET¹, TESS² and EESSI³). In 2007, the Administrative Commission established an Ad Hoc group⁴ to study the Identification of persons for the purpose of sending and receiving electronic messages between states' IT systems. The main findings have been extracted and adapted for this paper.

The process of identifying a person consists of two key elements: identity registration and identity verification.

3.1. Identity Registration

Identity Registration can be best described as the process and standards which apply when a service provider (e.g. government) is establishing the identity of an individual for the first time, so they can be recorded within the states Social Security system. This process includes the setting up of a computer record and the issuing of a token or credential, for example a Social Security card or number, which allows the service provider to identify the correct record on future transactions with that individual.

Key elements of the Identity Registration process are:

- Who does this person claim to be? What different attributes can they provide in terms of who they state they are?
- What information the person knows or possesses? What documents or other data items (credentials) can they provide?
- Do the 'credentials' belong to that person to enable their identity to be authenticated?

¹ Social Security Network

² Telematics for Social Security

³ Electronic Exchange Social Security Information

⁴ The author of this article was a member of this ad hoc group and Spanish representative in SOSENET, TESS and EESSI projects

- Are the documents genuine? These credentials will provide authentication of the identity.
- Are the credentials current and valid?

Once these key elements of registration have been addressed and authenticity has been established, this leads to enrolment of the person in the State Social Security System and then creation of the user's tokens, such as a Social Security card or the issue of a Social Security number.

3.2. Identity Verification on Providing Access to subsequent services

Identity Verification is the process that is inherent in each future contact with the citizen in order to establish that they are the true, genuine and rightful owner of the identity they are claiming. This process ensures the government has correctly ascertained the identity of the citizen and, as such, is happy to provide access to ongoing services on this basis.

Key issues to be addressed to ensure that an identity has been correctly verified are as follows:

- Is this person whom they claim to be?
- What documents or credentials can they provide?
- Do they or should they always have to provide documents?
- What are the considerations for how should identity verification be conducted across different channels of communication?
- If documents are not provided, what knowledge based test can they pass so that we can assure that their identity is correct? How else can identity be assured?
- How should an identity verification test be best conducted, e.g. by asking questions based upon what we already know about the individual and by their successful answering of questions, we can be sufficiently assured of their identity?
- Can it be done in different contact channels, e.g. face-to-face, telephone, e-channel, e.g. do they have electronic ID or electronic signatures established with government through e-registration and verification?

Identity registration is a previous process, in some way the "entry door" to the system. Identity verification could be a front end process or a computer process depending on the circumstance. A new Ad Hoc group was established to study the front end identification project but its conclusions are considered out of scope of this document. Quite the opposite, the data needed to identify a person in an electronic process is considered very relevant.

4. Data needed to identify an insured person in electronic processes.

The main conclusion of the Ad Hoc group established in 2007 was that a Minimum Data Set for the Electronic Identification of Persons should be set up and it should cover two scenarios:

- a) where a Personal Identification Number is present and
- b) where a Personal Identification Number is not present.

The respective data set agreed was:

Minimum Data Set Where a Personal Identification Number is present:

Where a Personal Identification Number (e.g. a National Registration Number or Sectoral Reference Number) **is** present the minimum data set is:

- Personal Identification Number;
- Surname/Family Name;
- Forename(s);
- Date of Birth;
- Sex;
- Name at Birth (Surname/Family Name and Forename(s))⁵.

Minimum Data set where a Personal Reference Number is **not** present

Where a Personal Identification Number (e.g. a National Registration Number or Sectoral Reference Number) **is not** present the minimum data set is:

- Surname/Family Name;
- Forename(s);
- Date of Birth;
- Sex;
- Place of Birth;
- Surname/Family Name at Birth;
- Father Surname/Family Name at Birth (if different);
- Mother Surname/Family Name at Birth (if different);
- Forename(s) of Father;
- Forename(s) of Mother;

Moreover, the Ad Hoc group identified a number of key issues where we feel further consideration is required. These issues are encapsulated in the recommendations and comments. The first one is considered the key recommendation and it is especially relevant in the context of this report.

Key Recommendation: It is critical for social security organizations to recognise the value of using Personal Identification Numbers in electronic messages. It is critically important because the use of Personal Identification Numbers significantly increases the likelihood of achieving a match with the right record. Therefore, it is recommended that

⁵ In most European countries, a woman changes her name upon marriage.

either the National Registration Numbers or Sectoral Reference Numbers for the individual from both the sending and receiving states are included in each message.

In other words, it is almost impossible to achieve an acceptable level of success in an electronic identification process without using personal identification numbers.

Recommendation 1 – There are a variety of National and Sectoral registration schemes in operation across states therefore it is recommended that each message contain the following information about each Personal Identification Number:

- Which state has issued it;
- Whether the number is a National Registration Number or a Sectoral Reference Number;
- Where the number is a Sectoral Number, which sector the number belongs to.

Recommendation 2 – In view of the importance of Personal Identification Numbers in electronic data transfer, it is recommended that states should consider capturing and storing the Personal Identification Numbers of states with whom they exchange electronic messages as this will facilitate the more efficient exchange of data. Processes for updating Personal Identification Numbers will also be required.

Recommendation 3 - To support the use of Registration/Reference numbers, the group felt it would be useful to encourage all member states to provide their citizens with a durable record of the number;

Recommendation 4 - The minimum data set defines how a piece of information should be recorded, it does not describe the factors applied to determine what that item of data is, e.g. where a Date of Birth is not known. These rules vary across states therefore it is recommended that further consideration is given to creating a common framework that accommodates these varying requirements;

Recommendation 5 – A number of issues of detail remain unresolved, e.g. whether there is a legal requirement to record an individual's gender/sex as 'unknown', and the recording of names/places so as to retain their original structure. It is recommended that these issues are investigated further;

Recommendation 6 – The inclusion of 'address' in the Minimum Data Set has been debated however the group decided that 'address' is not one of the data items that forms an identification data set

Recommendation 7- A number of data definition schemes exists, therefore when implementing the Minimum Data Set it is recommended that consideration is given to mapping the Minimum Data Set definitions/values to the prevailing standards in use at that point in time.

5. Study of registration keys

5.1 Types of registration keys in EU social security institutions

It is clear that a numeric key is crucial to identify persons, especially in the framework of the electronic exchange of messages. From a data processing point of view, the most relevant aspect of a PIN is the fact that a numeric key can be used to find specific information. In this way, a distinction has to be made between various registration keys applicable in a country. There are at least five possibilities:

- National Key: this is a key used to identify persons for administrative matters and used in all the Social Security areas.
- Social Security key: this is a key used to identify insured persons only for Social Security.
- Area Key: this is a key used to identify insured persons only for one particular Social Security area.
- Institution key: this is a key used to identify insured persons only at institution level.
- No key: there is no identification key used

One insured person may be identified by more than one registration key depending on the area and the country or by no key at all.

An identification key can change during the life of the insured person. This change may be due to various reasons (case of error, case of change of competent institution ...). Sometimes, a link between the old registration key and the new one is possible, sometimes not.

Establishing the correct combination of keys or PINS for an individual is the goal of the Identity Management process as the inclusion of keys or PINs in electronic messages greatly improves the accuracy and efficiency of the exchange process. For the purpose of interoperability, if an exchange between two different countries, sectors, schemas, etc. is expected, the identification part of every message should contain both the PIN allocated to an individual in the sending State/sector/branch and the PIN allocated to an individual in the receiving State/sector/branch. This will ensure that both parts will be able to exchange any future messages with each other using traceable PINs.

While some countries assign single PINs that are recognised across all branches of Social Security, others assign Sector or Branch-specific PINs. In the latter case, a PIN is only valid within one or a few Sectors/Branches. It is therefore important that States provide either:

- A National PIN (e.g. State-issued ID Number); or
- A National Social Security PIN (e.g. Social Security Number), or
- Sector/Branch PIN (e.g. number issued by a specific Sector/Branch)

It is also recognised that, in exceptional circumstances, an individual may not be able to provide a Social Security-related PIN. Where this is the case, States should attempt to capture any other PIN issued by the state(s) whose legislation the individual has been subject to. These other PINS should be used either in electronic data exchanges and/or in enquiries carried out through the organised network of authorised correspondents.

5.2 National keys

In 2005, Paul Hendriks and Fieke Roozen, from the Dutch Ministry of the Interior and Kingdom Relations, conducted a survey about Citizen Service Numbers in Europe (CSN). The outcome was presented at the e-2005 eChallenges Conference in Ljubljana (<http://www.echallenges.org/2005/>, Session 7b: eGovernment - Services for Citizens). The concept of Citizen Service Numbers is quite similar to that of National keys or National PINs.

This is why we will try to summarise the main findings of this report.

Regarding the characteristics of CSN, some countries explicitly choose a meaningless number, whereas other numbers contain information regarding the holder. A meaningless number refers to a selection of numbers, mostly in combination with a control mechanism such as a checksum or control number or control letter. Meaningful numbers contain specific information, such as gender, date of birth, region of birth, a few letters relating to the name at birth, as well as serial numbers referring to the sequence of birth. Also, a serial number is used to distinguish people having the same meaningful information within the number.

The stored type of information, as well as the quantity of information, varies greatly depending on the country. For example, some countries minimise the amount of information kept centrally, whereas others store large amounts of information. In some countries, deceased people are removed from the central database and stored in another database, where others keep all living as well as deceased citizens within the same database.

All countries have legislation concerning those who may use the CSN and access the corresponding registries. However, laws concerning access differ from country to country. Not all countries have an independent supervisory board that assures that rules regarding the CSN and its use are properly obeyed.

Regarding the usage of the CSN, in nearly all countries, public organizations have access to the CSN and related information, with some exceptions. The information is often accessible via a private network or by using a structured data exchange. Access via Internet is supported only in half of the countries. Public organizations generally do not need permission from their citizens to use their number, nor are all government organizations obliged to use the CSN.

The results of the survey do not show large differences in the sectors in which the CSN is used. However, the range of access for public organizations is diverse. The tax and social security sectors are often found to be the basis for any information exchange. These sectors are followed immediately by that used for the registration of births, deaths and marriages. Immigration offices have access as well while justice, health care,

education and municipalities are last on the list, but still half of the countries mention information exchange in these sectors.

In comparison to public organizations, there is a less interaction regarding information exchange between private organizations and the primary register of the CSN. Private organizations, when granted access, mostly do not need permission from their citizens regarding the use of their number, nor are private organizations obliged to use the CSN. Access to the register and CSN is in all cases explicitly regulated by national legislation.

In most countries, citizens have access to their information contained in the registry associated with the CSN. In half of the countries this service is provided via Internet.

Not only are there differences in the access provided, but also in the various “groups” who receive the number. When, how and if the number is provided to a citizen, immigrant, student or temporary worker differs. For instance, in some countries immigrants cannot obtain the number, in others it is provided as soon as certain conditions are fulfilled. Temporary workers and students are also handled differently between countries.

The number is often used to provide e-services to the citizen such as municipal services, first day registration of new employees, unemployment registration, medical declarations with public health insurance organizations, and financial transactions with financial institutions.

CSN are often used for information exchange amongst public organizations. In a number of countries it is also used for exchange with the private sector. Personal data can be combined in different ways amongst organizations. Three possibilities exist on how to couple identifying information.

1. CNS itself: the same number is used in different sectors for data exchange.
2. Referring index: If not all sectors use the same number; a referring index can be used. For each CSN, such an index contains references on the key information of an individual.
3. Encryption of numbers: Sector specific numbers are derived from the CSN using an encrypting algorithm.

Based on the collection of resources, the conclusion was that the basic assumptions of each of the CSN in the twelve countries studied were diverse. The characteristics of the CSN, the registry and registry owners diverge greatly. Variety is found on nearly all aspects, e.g. who is allowed to receive a number, what type of information is registered, and to whom and how access-rights are provided.

6. Identification keys in Spain

6.1 National Identification Number.

The Spanish national identification document is named “documento nacional de identidad (DNI)”. It’s a plastic card with the following details: names and surnames, birth date, address, parents’ name, gender, address, birth place and photo. The DNI is issued by the Police Department and its expiration date depends on the person’s age. To get a DNI it is necessary to have Spanish nationality.

Every DNI has a unique number, in the format 00000000-A (where 0 is a digit and A is a checksum letter). Foreign nationals are issued with a similar identity card, with a number in the format X-00000000-A (again, 0 is a digit, A is a checksum letter, X is a letter, generally X, and lately also Y), called an NIE Number (Número de Identificación de Extranjeros, Foreigner's Identity Number). The numbers are used as identification for almost all purposes. This is required for all transactions related to the tax authority.

Since 2006, this document has been issued in an electronic version that will be explained later.

6.2 Social Security Number

In Spain, Social Security has a specific card and number. It is necessary to make the distinction between Social Security number and Affiliation Number. The Social Security number identifies citizens in their relations with the Social Security system and may be requested by any citizen who does not have one, even if he or she does not work.. Whereas a Social Security number becomes an Affiliation Number at the time when a labour activity determining inclusion within the Social Security System begins.

The application for a Social Security Number or Affiliation number must be submitted at the Social Security Treasury or the Administration (local office) that corresponds to the citizen’s home address or the company’s business address, as applicable.

The characteristics of the affiliation are the following:

- It is required for people included in the System with regard to contributory type rights and responsibilities.
- It is unique and covers all the System's Schemes.
- It covers the entire life of the person included in the System.
- It is exclusive.

Affiliation is indicated on a Social Security card which also indicates the forename, surnames and the National Identification Number.

Regarding its characteristics, Social Security Number has three parts:

aa/bbbbbbbb/cc

First part (a) indicates the registration province⁶. For example, 28 means Madrid province and 08 Barcelona province.

Central part (b) is the citizen number in every province.

Final part (c) is deduced from (a) and (b) for control purposes through a formula that consists basically by taking the remainder after dividing by 97.

Social Security number plays a crucial role in IT processes because it's the primary key for accessing affiliation and contribution databases. However, that's not enough to ensure a perfect working of the system. Experience saw that something more was needed.

6.3 Individual's identifier

Data processing has a long tradition in Spanish Social Security. Historically, DNI was used as primary key in pension databases. The Affiliation number was used for both registration and contribution purposes. On the other hand, the affiliation number is supposed to be unique but it could happen for different reasons that one person might have more than one affiliation number. This is why a new key was created in the nineties, when a plan was developed by Social Security IT. This new identifier is called IPF: Individual's identifier⁷, it is only for internal use in data processing and it is formed by 15 characters with the following distribution:

1. Type of document: DNI, Passport, some other special numbers.
2. Document identifier: ten alphanumeric positions. Its composition depends on the type of document.
3. Duplicity: a **two-figure** sequential number to indicate possible duplicities
4. Splitting: a **two-figure** number to indicate that more than one person depend on the same titular, in the special case of minors or mentally handicapped persons.

The function of this identifier is to associate all possible keys to a single one based in the DNI number or its equivalent. On the other hand, it is an alternative access key in case of the loss of the Social Security number or for verification purpose. In the registration process, the citizen is asked for the DNI number; it is introduced into the database and associated to the Social Security number. In this way, almost any possibility of duplication is avoided. But the problem remains from any legacy data, mainly when the databases didn't exist. If this happens, the older number is considered as principal and the other one is associated as secondary. All of them are referenced by the same IPF.

⁶ Spanish territory is organised in 52 provinces. Social Security institutions have a territorial directorate in every province.

⁷ IPF stands for "Identificador Persona Física". In Spanish, "persona física" means individual.

Moreover, IPF is always used in exchanges with other institutions or Government departments. The main reason is that IPF have DNI encapsulated and DNI is generally known in any relationship between citizen and Public Administration.

7. Electronic identity

7.1. Digital certificates

As a complement to this paper about identification for improving data processing, it seems appropriate to quote what our main commitment is nowadays, e-government. In this framework, since the Internet has become part of our life, a new form of identification has become an issue: the digital identification. With people interacting through digital devices, the need for trustworthy identification is of paramount importance. The communicating parties want to be sure with whom they are communicating.

Many digital identification means have been developed, like biometrics, PINs, etc., but one of them is used more and more because of the level of security it provides, the digital certificates (also known as PKI certificates).

Digital certificates are electronic files that are used to uniquely identify people and resources over networks such as the Internet. Digital certificates also enable secure, confidential communication between two parties.

As in the real world (e.g. passport), digital certificates are issued by a trusted third-party (like the passport office) called a Certification Authority, who validates the certificate holder's identity and signs the certificate so that it cannot be forged or tampered with.

A certificate typically includes a variety of information pertaining to its owner and to the Certification Authority that issued it, such as:

- The name of the holder and other identification information required to uniquely identify the holder;
- The holder's public key, that can be used to encrypt sensitive information for the certificate holder;
- The name of the Certification Authority that issued the certificate;
- A serial number;
- The validity period (or lifetime) of the certificate (a start and an end date).

Digital certificates have two basic functions. The first is to certify that the person named is a reliable source, in other words, who they claim to be. The second function is to provide protection for the data exchanged from the visitor and the website from tampering or even theft, such as credit card information.

The most common standard for digital certificates nowadays is the so called X.509 certificate.

In Spain Social Security acts as the Registration Authority in two cases:

- For company representatives to update and view information, strictly on the RED System⁸. An internal certificate is produced in this case and called a SILCON Certificate.
- For citizens to have access to all the personal services offered by Social Security, A Class 2 CA certificate from the National Mint and Stamp Factory is used.

Some other certificates from different issuers allow access to the services of the Virtual Office⁹ but with specific conditions in any case. Electronic Id, explained in next section, is also valid.

7.2. Electronic Id in Spain: The electronic National Identity Document (eDNI)

As it has been said before, in Spain, the DNI (Documento Nacional de Identidad in Spanish, meaning National Identity Document) has been the mechanism that has provided every citizen with a legal identity for more than 50 years. It is used in every administrative transaction as well as in many commercial transactions. A DNI's number (a unique number which identifies the holder) is stored in a very high number of databases, for private or public use, as the main identifier of persons.

Law 59/2003 on electronic signatures makes it possible to evolve the former plastic DNI into a new electronic DNI, more appropriate for the Internet and any electronic transaction. The aim of this new eDNI is multiple:

- To prove the citizen's identity in the real world, as well as in the virtual world, allowing the digital signature of any kind on electronic documents.
- To guarantee interoperability with European projects on electronic id.
- To foster confidence on electronic transactions and to drive the development of new services

eDNI contains in an electronic chip the following:

- X509/v3 certificates and related private keys, for authentication and signature
- Personal data about the holder (Name, Surname, Gender, Nationality, Birth date, Birth place, Address)
- Data about the Document (DNI's number, Validity period, issuance date)

⁸ The RED System is a service offered by the TGSS to companies, groups of companies and professionals. Its mission is to permit the interchange of information and documents between the TGSS and users over the INTERNET

(http://www.seg-social.es/Internet_6/SistemaRed/Informacion/InformacionGeneral/QueeselSistemaRED/index.htm)

⁹ http://www.seg-social.es/Internet_6/OficinaVirtual/Certificadosdigital47735/OtrasAutoridadesdeC51674/index.htm

- Use of eID for cross-border electronic delivery for citizens and businesses,
- And for testing the electronic process of address change for EU citizens that move to other Member States.

The project will develop, test and validate common specifications for national eID systems to work together. These will be made freely available. All industries that would like to develop services for eID in the future will have equal and free access to the common specifications.

The project will result in the smooth cross border operation of several key public services. The solution will be scalable to all EU Member States. It will be technologically transparent, robust, with measurable benefits and will be implemented in such a way that it will be sustainable beyond the life of the pilot.

8. Conclusions

Identification is a very important issue for Social Security. There can be no doubt when identifying an individual at the front desk, during the registration process or at any other moment when access to services is required.

In the framework of interoperability, identification becomes a condition sine qua non for successful data or messages exchanging.

According to European working groups on this subject, it is critical to recognise the value of using Personal Identification Numbers in electronic messages. It is critically important because the use of Personal Identification Numbers significantly increases the likelihood of achieving a match with the right record.

Spanish Social Security organizations' experiences confirm this viewpoint and not only from the interoperability aspect but also in order to guarantee good data management and prevent possible mistakes.

Internet is changing communication ways, habits and culture. Almost everything in the future will pass through the Internet and Social Security, its processes and its relations with the citizens will not be an exception. Maybe nowadays Internet can be considered as an alternative channel, but Social Security institutions should be ready to identify an individual through a digital path.

9. References

- SOSENET Project, Deliverable 14, Part 1: Identification of insured persons in European Social Security
- TESS Committee studies 1994-96;
- Final Report of Ad Hoc Group on the Electronic Identification of Individuals
- Final Report of Ad Hoc Group on Identity Management
- Paul Hendriks and Fieke Roozen: Inventory of Citizen Service Numbers in Europe, e-2005 eChallenges Conference in Ljubljana (<http://www.echallenges.org/2005/>, Session 7b: eGovernment - Services for Citizens)
- www.seg-social.es
- <http://www.cert.fnmt.es/>
- <http://www.dnielectronico.es/>
- <https://www.eid-stork.eu/>