Implications of Age Assurance on Privacy and Data Protection: A Systematic Threat Model
Research Paper

Marta Beltrán(✉) and Luis de Salvador

Agencia Española de Protección de Datos (AEPD), Madrid, Spain
{mbeltran,lsalvadorc}@aepd.es
https://www.aepd.es/en

Abstract. In today's digital world, children are encouraged to develop a significant part of their daily lives by online means. Age assurance solutions have become essential tools to ensure the protection of their fundamental rights. This is reflected in different regulatory frameworks, strategies, codes, and recommendations concerning children's protection and a safer Internet. These solutions estimate or verify users' ages, allowing for the establishment of age restrictions for content, services, and goods. However, the intersection of safety and privacy within this field poses significant challenges, not only for children but for all Internet users.

This research explores the privacy and data protection implications of age assurance, analyzing existing solutions and proposing a comprehensive privacy threat model. The threat model developed in this paper can be widely applied to improve the design and implementation of current solutions, policy and guidelines formulation, and awareness initiatives about age assurance challenges. By balancing adequate age assurance with robust and compliant data protection, we can create a digital environment for everyone that guarantees the protection of fundamental rights.

Keywords: Age assurance · Children protection · Data protection · Privacy risk management · Threat model

1 Introduction

In an increasingly digital world, where children and young users actively engage with online content and services, ensuring their safety becomes paramount. Policymakers have recognised the significance of age assurance mechanisms within this context in different regulatory frameworks [1,3,4] and design codes [2,18]. These solutions, designed to estimate or verify users' age, are devoted to safeguarding children from harmful content, age-inappropriate interactions, exploitative practices, or the purchase of age-restricted goods, to mention some examples [24].

M. Jensen et al. (Eds.): APF 2024, LNCS 14831, pp. 1–22, 2024.
https://doi.org/10.1007/978-3-031-68024-3_1

Age assurance solutions play an essential role in determining the age of users, but they also raise concerns about privacy and data protection [28]. These solutions often require collecting personal data such as birth dates, identification documents, or biometric information, which can pose risks when processing, storing, or sharing this data. There is often a lack of transparency regarding user age data processing, with many providers relying on third-party age assurance services. This can introduce additional privacy threats, as users may not know how their data flows beyond the platform they interact with. Moreover, when combined with other identifiers such as IP addresses or browsing history, age assurance data can potentially identify users, revealing more information about them than intended.

In addition, age assurance solutions should focus on estimating or verifying age. However, over time, some of them are expanding their functionality to collect additional data (e.g. identity, location, interests), generating parallel identity frameworks with new and different purposes different from age assurance. In these cases, age assurance solutions may inadvertently create profiles based on user data. These profiles can be used for targeted advertising, behavioural analysis, or discriminatory practices [13].

Age assurance solutions must manage a careful balance between safeguarding children and respecting the rights of all individuals. To effectively mitigate these risks, strong privacy-enhancing measures are necessary. However, it is important to note that there is no comprehensive list of all potential privacy threats arising from data processing for age assurance purposes.

This research paper contributes significantly to the field by 1) Analysing existing age assurance solutions, bringing to light their common aspects and identifying the most prevalent architectures and their privacy implications. 2) Proposing a comprehensive and systematic threat model to identify privacy threats specific to age assurance systems. 3) Discussing the produced model to offer practical recommendations to enhance privacy while maintaining adequate age assurance.

The rest of this paper is organised as follows: Sect. 2 summarises the related work on age assurance and threat modelling and outlines the motivation of this research. Section 3 describes our research method. Section 4 analyses the most extended architectures used to solve age assurance. Section 5 describes the produced privacy threat model, while Sect. 6 discusses the identified threats and provides initial recommendations. Finally, Sect. 7 summarises our main conclusions.

2 Related Work and Motivation

2.1 On Age Assurance

Age assurance is a term adopted to describe different methods used to determine users' age on online platforms, applications, and services [5]. Age estimation and age verification are two critical components of age assurance, each having a specific function. Age estimation uses probabilistic methods based on biometry

(often voice or facial analysis), capacity testing, behaviour patterns, and language use to guess a user's age. Age verification, conversely, guarantees age by using an existing digital identity or providing a document, such as a passport or other government-issued ID, to confirm an individual's age. A subcategory within verification is composed of the methods inferring the age or age range due to the possession of a particular document or certificate, such as a bank card or digital credential.

Previous work has identified the main challenges concerning age assurance:

- Accuracy: False positives and negatives can occur, leading to incorrect age assurance.
- Lack of Standards: The absence of a unified standard and the adoption of diverse strategies and approaches create a fragmented landscape, presenting challenges for the industry in integrating their platforms, applications and services with different alternatives.
- Cost: Given the fragmentation mentioned above, implementing age assurance systems may be expensive, especially for small and medium-sized providers.
- User Experience: Assurance processes might drive users away from the platform due to a negative experience (usability issues, lack of accessibility, etc.).
- Universality and Inclusiveness: There are limited options for assuring users' age online. Each method has its advantages and disadvantages, but covering all domains' and users' requirements is challenging without causing exclusion.
- Compliance: Providers must ensure compliance with different regulations (children protection, online safety, digital services, audiovisual contents, data protection, etc.) and keep up with changing and heterogeneous legal requirements.
- Security: Users trying to access unauthorized content or services can circumvent or work around age assurance systems. Malicious actors may also be interested in attacking these systems to steal personal data, commit fraud, or carry out some impersonation.
- Privacy and Data Protection: Age assurance systems often require users to provide personal data, which raises privacy concerns.

Research efforts have been focused on exploring ways of overcoming some of these challenges [11,20]. However, most of the works focus on discussing whether it is the most appropriate approach in different use cases [7,8,14,22,29] or on ethical and public policy aspects [9,25,32].

Furthermore, the British Standards Institute (BSI) and the Digital Policy Alliance developed a code of practice for online Age verification service providers. This code, called PAS 1296:2018 [10], applies to providers who must perform age assurance processes to determine whether or not a citizen can access age-restricted goods, content or services.

The International Standards Organisation (ISO) is currently working on developing ISO/IEC 27566 [19]. One of the main drivers for this effort is the growing agreement that a straightforward method to explain the levels of confidence attained by different assurance components would benefit service providers, relying parties, and regulators.

Finally, it is worth mentioning two European initiatives. The first, the euCONSENT project [15], working on proposing operation extensions to the eIDAS (electronic IDentification, Authentication and trust Services) infrastructure to deliver pan-European, open-system, secure and certified interoperable age verification and parental consent to access Information Society Services. The second, the Task Force on Age Verification under the Digital Services Act, set up in 2024 to progress towards a harmonised EU approach to age verification [12].

2.2 On Threat Modelling

Threat modelling is a systematic process used to identify, understand and communicate threats and mitigations while protecting something valuable. It is a structured representation of all the information that affects the safety, security or privacy of a wide range of assets, including software and applications, systems, networks or business processes [31].

While traditional threat modelling concerns safety and security threats, privacy threat modelling explicitly addresses risks related to collection, recording, organisation, structuring, storage, adaptation or alteration, retrieval, consultation, use, disclosure by transmission, dissemination or otherwise making available, alignment or combination, restriction, erasure or destruction of personal data.

Privacy threat modelling has been applied in various domains. The LINDDUN methodology is the most widely used in the field of privacy engineering [30]. It can anticipate and prevent data processing that may lead to privacy risks or analyse privacy requirements. However, it has also been used as a mnemonic for a brainstorming-style exercise rather than a means to elicit privacy threats systematically [16,23,26]. Furthermore, since a PIA (Privacy Impact Assessment) or a DPIA (Data Protection Impact Assessment) implies a risk assessment which typically follows a step-by-step process of risk identification and risk mitigation, a LINDDUN-based privacy threat model can also be a systematic and traceable starting point to conduct this kind of exercises [17].

2.3 Motivation

To date, no systematic threat model has been proposed for identifying the privacy and data protection threats that arise when using age assurance systems. This model would provide critical insights into the adequacy of existing solutions in terms of both data protection and regulatory compliance, as well as their alignment with the rights and freedoms of citizens. In the absence of such a model, making recommendations to relevant stakeholders, including providers, third parties, regulators, control authorities, and users, has been challenging. Adopting a systematic threat model for data protection in age assurance systems promises to address these challenges by providing a comprehensive framework for evaluating existing solutions and identifying opportunities for improvement.

3 Method

This research proposes a method with two main phases to address the identified research gap.

1. Architectures analysis: The first phase involves reviewing the age assurance literature and analysing the proposed architectures (see Sect. 4). Searches of standards, scientific papers, commercial products, and patents need to be carried out using the keywords age assurance, age estimation, and age verification. The aim is not to produce a survey of the different solutions and methods but to understand the targeted systems and their common aspects: processes, involved agents, information flows, data formats, etc. These common aspects in the analysed architectures can be used as the input for the second phase.
2. Threat model: This study's second phase involves the threat modelling process, as outlined in Sect. 5. To accomplish this, the LINDDUN methodology has been selected, a well-established and privacy-focused threat modelling approach that has been empirically evaluated. We have adopted the LINDDUN methodology in two critical aspects of our work. Firstly, a Data Flow Diagram (DFD) has been used as a graphical tool to model the system under analysis and guide the threat identification and analysis process. The LINDDUN catalogues and body of knowledge have been used to aid in this process [21]. Secondly, we have analysed threats across various categories, some consistent with the LINDDUN approach. In contrast, others have been modified or added to align with the primary objective of this research, which is data protection, regulatory compliance, and the protection of individuals' rights and freedoms. The LINDDUN methodology is an acronym for Linking, Identifying, Non-repudiation, Detecting, Data disclosure, Unawareness, and Non-compliance. Some categories have been retained, while others have been removed or added as necessary. Specifically, we have eliminated the Non-compliance category, from the regulatory point of view, as data processing that does not comply with regulations cannot be implemented. Additionally, we have added three new categories: Inaccuracy, Exclusion, and Data breaches. These categories better reflect the focus of our research and enable us to more effectively identify and address potential threats to data protection and regulatory compliance. The considered categories, LIINE3DU, are then as follows:
   - Linking: This threat involves associating different data items or user actions to learn more about a data subject or group.
   - Identifying: This threat involves learning the identity of a data subject directly (through leaks, for example) or indirectly (through deduction or inference, for example).
   - Inaccuracy: This refers to using obsolete, wrong, incomplete or low-quality data that may lead to incorrect decisions or actions, potentially causing inconvenience or even harm to the data subject.
   - Non-repudiation: This refers to the ability to attribute a claim to the data subject (something they know, they are, they do, etc.).

- Exclusion: This threat involves unintentionally or deliberately failing to adequately serve a data subject, hindering their participation or involvement in physical or digital life.
- Detecting: This threat involves deducing the existence of data items or user actions through observation.
- Data Breach: This threat involves destruction, loss, alteration, unauthorised disclosure of, or access to, personal data by mistakes, malicious insiders or cyberattacks.
- Data Disclosure: This threat involves excessively collecting, storing, processing, or sharing personal data.
- Unawareness and Unintervenability: This threat involves insufficiently informing, involving, or empowering data subjects in the processing of their personal data.

The last two categories are intrinsically linked to an organisation's risk management processes and to the data processing design, and entail non-compliance with data protection regulatory frameworks. Conversely, the remaining proposed categories may directly impact individual rights and freedoms from a broader perspective.

4 Architectures Analysis

The market for age assurance is changing with the introduction of new products and services. Additionally, regulatory actions for online protection of children and data privacy are shaping the market.

Regarding age assurance, self-declaration is no longer considered a reliable solution in almost all use cases. In light of this, three different architectural approaches have emerged. The first approach involves direct interaction between the user and the provider, who is responsible for ensuring their age without the involvement of specialised providers. The other two approaches are "tokenised", with the provider delegating age assurance to a specialised third-party provider capable of returning the necessary information about the user in a date of birth, an age, an age range or an age threshold fulfilment token.

When a user wants to access age-restricted content or services, the data flow begins with an access request. During this process, participating agents store data based on their roles and level of access. This can include user data, metadata, logs, and the age assurance result.

4.1 Age Assurance Solved by the Content or Service Provider

In this first architecture, the steps followed to solve age assurance are:

1. Age Assurance Request: Once the user attempts to access age-restricted content or services, the system triggers the age assurance process. The system prompts the user to provide information or evidence concerning their age.

2. Data Collection: Depending on whether the system uses age estimation or age verification, the data collected will vary. For Age estimation the system may collect data such as biometry (voice samples or face images), behavior patterns, language use, or other indirect indicators of age. In the case of Age verification, the user is asked to provide a reliable proof of age. This could be in the form of a government-issued ID, a passport, an electronic certificate or credential, credit card information, or another form of age verification accepted by the system.
3. Age Check: The system processes the collected data in-house. For Age estimation the system uses statistical methods, machine learning or artificial intelligence models to estimate the user's age based on the collected data. For Age verification the system checks the provided proof, for example, the date of birth on a government-issued ID, the age on an electronic certificate or credential or the confirmation that a credit card is valid and belongs to the user.
4. Result Communication: The system then communicates the result of the age assurance operation to the user. If the user's age is successfully assured, they are granted access to the age-restricted content or service. If the age assurance process fails, the user is denied access and may be given a reason for the denial.

4.2 Age Assurance Solved by a Third-Party Provider as a Proxy

In this second architecture a new agent is necessary to solve age assurance: a third-party specialised provider. The provider offering contents or services with age-restrictions adopt the role of Relying party, delegating the responsibility on this third-party provider. The steps followed to solve age assurance are:

1. Age Assurance Request: Once the user attempts to access age-restricted content or services, the system triggers the age assurance process. Instead of directly interacting with the user, the system redirects them to a third-party age assurance provider.
2. Data Collection at the Third-Party Provider: The user is then asked to provide the necessary information or evidence for age estimation or verification to this third-party provider.
3. Age Check by the Third-Party Provider: The system processes the collected data. For Age estimation the system uses statistical methods, machine learning or artificial intelligence models to estimate the user's age based on the collected data. For Age verification the system checks the provided proof.
4. Result Communication: The third-party provider then communicates the result of the age assurance process to the original service provider. The service provider then communicates the result to the user. If the user's age is successfully assured, they are granted access to the age-restricted content or service. If the age assurance process fails, the user is denied access and may be given a reason for the denial.

4.3 Age Assurance Solved by a Third-Party Provider as an Autonomous Agent

In this last architecture, a third-party specialised provider is used again. However, it does not act as a proxy between the user and the content or service provider in this case. It only produces the age token the user must present at the content or service provider to ensure their age. This architecture could be understood as a particular case of architecture 1, where the data collected at step 2 is this age token produced by a reliable external agent capable of estimating or verifying the user's age. In this research, we prefer to analyse it as a different architecture to consider its peculiarities in terms of data protection.

The steps followed to solve age assurance in this case are:

1. Age Assurance Request: Once the user attempts to access age-restricted content or services, the system triggers the age assurance process. Instead of directly interacting with the user, the system asks them for a electronic capability assured by a third-party age assurance provider.
2. Data Collection at the Third-Party Provider: The user is then asked to provide the necessary information or evidence for age estimation or verification to this third-party provider.
3. Age Check by the Third-Party Provider: The system processes the collected data. For Age estimation the system uses statistical methods, machine learning or artificial intelligence models to estimate the user's age based on the collected data. For Age verification the system checks the provided proof. The third-party provider generates an electronic capability (signed certificate or unforgeable token), usually only if the user's age is successfully assured. This capability is sent to the user so the user can present it at the original service provider. If the age assurance process fails, the user is usually denied the age token and may be given a reason for the denial.
4. Result Communication: If the user's age is successfully assured, they are granted access to the age-restricted content or service. If the age assurance process fails, the user is denied access and may be given a reason for the denial.

Steps 2 and 3 of this process may or may not be synchronized with access to the content or service provider. This means that the user can request the age token at the moment when they need to assure their age to a content or service provider, or they can do it offline, at a prior time, and store one or more age tokens in a securely protected repository, trusted for future use. In this paper, only the synchronized approach is considered (when the age token is requested when needed) because it is the one that has been found already implemented in real solutions that can be analysed and tested for the proposed threat model.

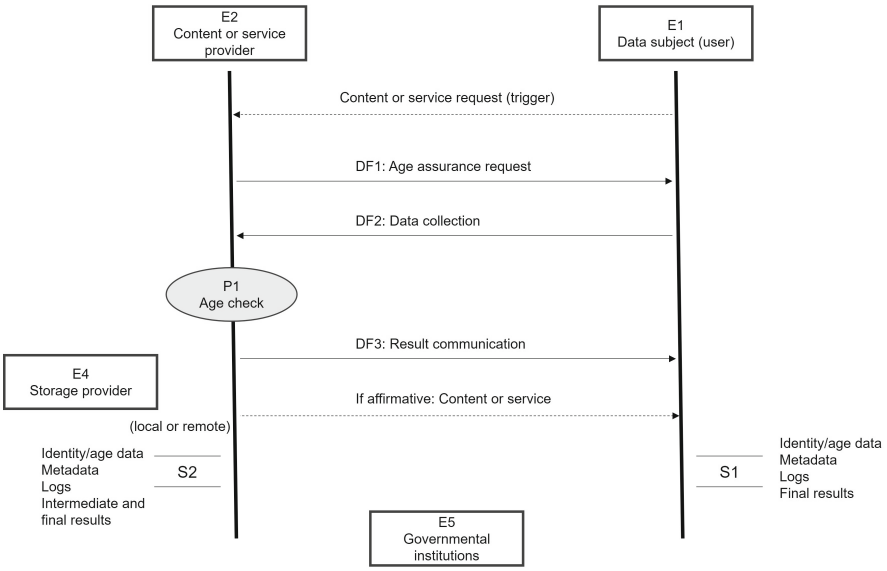

Fig. 1. Architecture 1 (DFD): Age assurance solved by the content or service provider.

5 Threat Model

The Data Flow Diagrams (DFD) produced to conduct the threat modelling process come directly from the description provided in the previous section for the different architectures (Figs. 1, 2 and 3).

Entity E1 is the subject whose age is to be assured. Entity E2 is a content or service provider offering content or services with age restrictions, while entity E3 is the third party specialised in age assurance. E2 and E3 have its own storage, which can be local or remote. For this reason, E4 is a Storage as a Service provider. And E5 are governmental institutions such as judges and courts, security forces and bodies, etc.

All data flows from DF1 to DF7 in the different architectures are those that have been identified in this research as essential when performing age assurance: Age assurance request, Data collection and Result communication, but also Service request and Service result when a specialised third-party provider is involved. These are the minimum data contents of all these essential flows:

- Age assurance request: DF1 in architecture 1, DF2 in architecture 2 and DF3 in architecture 3, this flow includes at least the age threshold that must be exceeded to gain access to the requested content or service.
- Data collection: DF2 in architecture 2, DF3 in architecture 2 and DF4 in architecture 3, this flow includes personal data. What specific data depends on the method used to perform age assurance, estimation or verification, and the specific approach (biometry, date of birth, etc.).

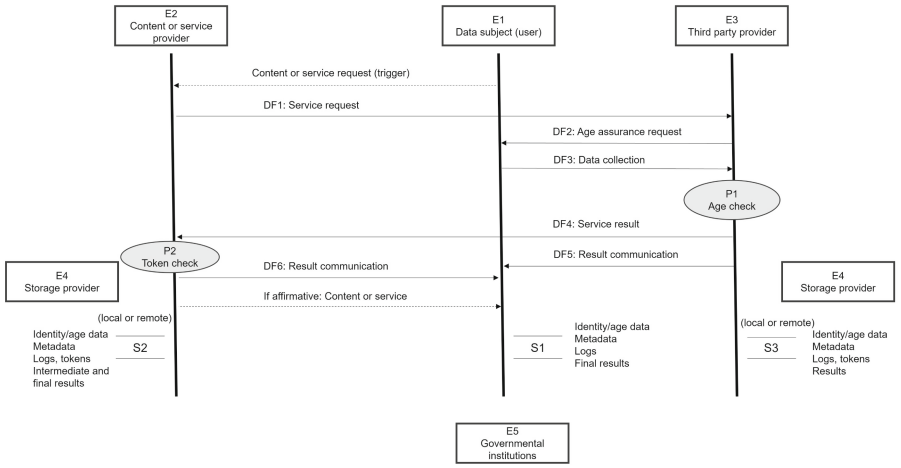

Fig. 2. Architecture 2 (DFD): Age assurance solved by a third-party provider as a proxy.

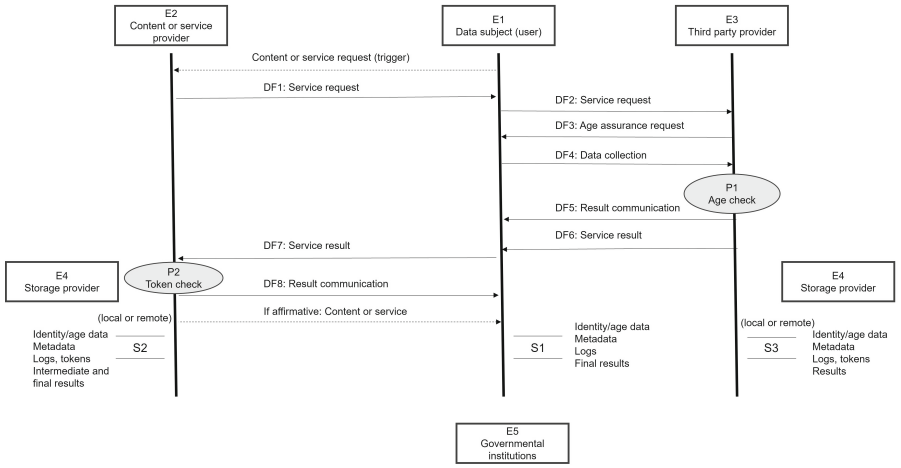

Fig. 3. Architecture 3 (DFD): Age assurance solved by a third-party provider as an autonomous agent (synchronously).

- Result communication: DF3 in architecture 1, DF5 and/or DF6 in architecture 2, and DF5 and/or DF8 in architecture 3. This flow includes a passed (age above threshold)/not passed (age not above threshold, no consent, error) and a motivation for this result.
- Service request: DF1 in architecture 2 and 3, DF2 in architecture 3. This flow includes, at least, information about this specific age assurance process, such as a transaction number or a nonce. It may also include the age threshold.

– Service result: DF4 in architecture 2, DF6 and DF7 in architecture 3. This flow includes a token with information about the user's date of birth, age ("24 years old", for example), the user's age range ("+18", for example) or the fulfilment of the age threshold (passed/not passed). This token is usually signed to allow the verification of its origin and usually includes an expiration time.

Finally, two different processes have been identified. P1 is the Age check, a simple comparison between the user's age and the age threshold. P2 is the Token check, the interpretation and processing of the information received in the token and the verification of its validity (it has not expired, the origin is trusted, etc.). Different agents can execute these processes depending on the specific analysed architecture.

These DFDs have been produced considering the following assumptions:

1. The developed threat model considers that age assurance processes are data processing by themselves.
2. The produced threat model is focused exclusively on the rights and freedoms of the subject undergoing the age assurance process.
3. All elements of the DFD that are the responsibility of the content or service provider and the third parties with which it works are sufficiently protected against malicious insiders and cyberattacks, as well as the communications between them. There is a trusted boundary that groups the essential elements for age assurance.
4. The processes within the DFD are appropriately implemented and perform the function for which they are designed.
5. The elements appearing in the DFD cannot be impersonated except for the subject undergoing identity verification.

In this way, threats related to security are not modelled (a different and complementary modelling process would be needed) but exclusively to privacy and data protection. Under all these assumptions and after analysing 12 different solutions (8 already in production and 4 in a prototype phase), the following main threats have been identified.

5.1 Linking

Threat 1: Unique Identifier. A threat agent is able to learn more than expected about a data subject and their actions by associating this data subject to a unique identifier: this identifier enables the linking of all accesses/requests to the same individual (without identifying this data subject).

– Architecture 1: E2 is able to materialise this threat through the DF2, containing personal data, and through S2, if it stores some of this personal data. E2 is able to link collected data, for example, an email address, a user ID or a face template, with the user's content requests.

- Architecture 2: E2 is able to materialise this threat through the DF4, containing the token data, if it contains unique identifiers. Moreover, through S2 if it stores some of this token data. E2 is able to link token data, such as an email address, a user ID, or a face template, with the user's content requests. E3 can also materialise this threat through the DF3, which contains personal data, and S3 if it stores some of this personal data. E3 is able to link collected data, for example, an email address, a user ID or a face template, with the service requests coming from the content or service provider (E2).
- Architecture 3: E2 is able to materialise this threat through the DF7, containing the token data, if it contains unique identifiers. Furthermore, through S2 if it stores some of this token data. E2 is able to link token data, such as an email address, a user ID, or a face template, with the user's content requests. E3 is able to materialise this threat through the DF4, which contains personal data, and S3, if it stores some of this personal data. E3 is able to link collected data, for example, an email address, a user ID or a face template, with the service requests coming from the user (E1).

It is worth noting that the privileged position of the third-party provider in architectures 2 and 3 allows it to trace one user's interactions with different content and service providers. This means that all online activities related to adult content or sites could be associated with this specific data subject.

Any entity E4 could also materialise this threat if any storage elements are remote and consumed as a service from an external provider.

Threat 2: Linkable Data Through Combination, Profiling or Inference. A threat agent is able to learn more than expected about a data subject and their actions by associating this data subject to all their accesses/requests (without a unique identifier and without identifying this data subject). Access requests and service requests usually contain attributes that, when combined, are unique to a data subject. The more data is collected (volume and variety) and the more detailed it is, the easier it is to find unique patterns for linking.

- Architecture 1: E2 is able to materialise this threat through the DF2 and S2, which may contain a variety of metadata and logs. For example, E2 may link different requests to the same user tracing their IP address, their location, combining data such as their date of birth and their postal code, performing browser fingerprinting during their interactions, analysing timing patterns or analysing behaviour or language.
- Architecture 2: E2 is able to materialise this threat through the DF4, containing the token data, depending on its content (and through S2). E2 is able to link token data, such as IP address, location, date of birth, and postal code, with the user's content requests. E3 can also materialise this threat through the DF3, which contains personal data, and through S3. E3 is able to link the above-mentioned "quasi-identifiers" or users' profiles collected or built during user interactions, with service requests coming from the content or service provider (E2).

– Architecture 3: E2 is able to materialise this threat through the DF7, containing the token data and through S2. E3 is able to materialise this threat through the DF4, containing personal data, and through S3. E3 is able to link "quasi-identifiers" or user's profiles collected or built during user's interactions with the service requests coming from the user (E1).

5.2 Identifying

Threat 3: Identified Data. A threat agent is able to collect and process the data subject identity within the age assurance process.

– Architecture 1: E2 is able to materialise this threat through the DF2 and S2, which may contain a variety of metadata and logs. For example, E2 requires a registration that includes a full name. Alternatively, when the user ID within the platform is based on a combination of first and last names.
– Architecture 2: E2 is able to materialise this threat through the DF4, containing the token data, depending on its content (and through S2). For example, name and surname. E3 can also materialise this threat through the DF3, which contains personal data, and through S3.
– Architecture 3: E2 is able to materialise this threat through the DF7, containing the token data and through S2. E3 is able to materialise this threat through the DF4, which contains personal data, and S3.

Threat 4: Identifiable Data. A threat agent is able to learn or infer the data subject's identity within the age assurance process. For example, reversing a pseudonym, using a reverse image search tool to identify the data subject from a face image, etc. Different attributes or combinations of attributes are unique to a specific data subject and enable their identification. The scenarios are the same as in threat 3, but the threat agent does not collect/process the identity directly but instead has to perform some operation to derive it. There is a particular case when an entity E5 (usually, it has to be governmental to have the capacity to materialize this threat) is the only one that can identify the data subject, requesting the necessary data from entities 2 and 3 to do so. Each of these entities, separately, cannot identify the user with the data they handle. However, E5, collecting the data processed by both (for example, with a court order), can infer the data subject identity.

5.3 Inaccuracy

Threat 5: Error. Age assurance is performed by processing obsolete, wrong, incomplete or low-quality data. The threat materialises unintentionally. For example, the identity document with which the age verification is carried out has expired or the data subject has turned years old and their data still needs to be updated. In the case of age estimation, the data used to make the age prediction is not of sufficient quality. In a cold start, for example, with a new

user about whom the provider barely has any information or with an image of a face taken in bad lighting. There are also errors inherent to estimation methods since, by their very nature, they cannot guarantee 100% accuracy (methods performance, potential bias or discrimination, etc.). This threat is not only about direct data collection, it can also be materialised with expired or deficient tokens in architectures 2 and 3. The elements of the DFD involved in this threat are the Data collection flows (DF2, DF3, and DF4 in architectural designs 1, 2, and 3, respectively) and the processes P1 and P2, Age check and Token check.

Threat 6: Circumvention and Evasion. Age assurance is performed processing obsolete, wrong, incomplete or low-quality data, intentionally because a user wants to pretend they have a different age from their real one. User's binding is essential in verification methods, for example, since one user could use another's document, certificate or credential if proper checks are not performed. Alternatively, because the E2 or E3 entities have an interest (commercial, directly or indirectly) in ensuring ages differ from their real ones. For example, to allow more users to access the platform, to monetize a more significant number of positive results, etc. The elements of the DFD involved in this threat are again the Data collection flows (DF2, DF3, and DF4 in architectures 1, 2, and 3, respectively) and the processes P1 and P2.

5.4 Non-repudiation

Threat 6: Attributable Data Evidence. A threat agent can use data to prevent the data subject from denying an access/request. Threats 1, 2, 3 and 4 may magnify the impact of this one because a group of different actions can be attributed to the same data subject (by materialising Linking) even knowing who this data subject is (by materialising Identifying).

- Architecture 1: E2 is able to materialise this threat through S2, which may contain a variety of data, metadata and logs.
- Architectures 2 and 3: E2 is able to materialise this threat through S2 and E3 through S3. Digitally signed tokens are an additional source of attributable data evidence that may prevent the data subject or E3 from denying their involvement in a transaction later.

There is a particular case when entity E5 is the only one that can attribute an action to the data subject, requesting the necessary data from entities 2 and 3 to do so. Even requesting the contents of S1 in the user's device, for example, the Result communication logs that could be combined with browser history to substantiate claims about their online activities.

5.5 Exclusion

Threat 7: Limited Access. A data subject is marginalised, unable to complete age assurance processes, which hinders their participation in different online

activities and restricts their autonomy because they have restrictions to use the required elements or technologies. For example, a user does not have a smartphone with a high-quality camera or an NFC reader, or is not in possession of the document or certificate requested to verify their age.

Threat 8: Limited Literacy or Confidence. In this case the data subject is marginalised because they have a low-proficiency level, lack of knowledge or lack of confidence in the provided methods. For example, the user does not have the capability to create an account in a specific platform, or does not want to use a governmental solution.

5.6 Detecting

Threat 9: Observed Communications. The access/request of a data subject or an attribute of this subject can be deduced through the observation of communication within the DFD. For example, Service request data flows at E3 (architectures 2 and 3) imply that users are trying to access age-restricted content or services. If threats 1, 2, 3 or 4 have been materialised, this user is also a specific one. Observing Service result flows (architectures 2 and 3) and Result in communications, in general, can also enable detection, even inferring if a user is a child or their age range.

Threat 10: System Responses. The access/request of a data subject or an attribute of this subject can be deduced by examining system responses. For example, E2 or E3 can infer if a user is a child or their age range with the results obtained from P1 and P2. Logs, error messages, metadata, etc., stored in S1, S2 and S3 also allow detection.

5.7 Data Breach

Threat 11: Breach of Stored Data. Even when proper cybersecurity risk management procedures have been implemented, data breaches affecting information that resides in databases, files or other storage systems may be inevitable. They can affect S1, S2, and S3 and all the personal data stored in these repositories (also in logs, error messages, metadata, etc.). Threats 1, 2, 3 and 4 may magnify the impact of this one because the external threat agent materialising this threat can attribute the leaked data to a specific data subject.

Threat 12: Breach of Data in Transit. In this case the breach affects data in motion being transferred between two DFD elements over networks. The elements of the DFD involved in this threat are the Data collection flows (DF2, DF3, and DF4 in architectures 1, 2, and 3, respectively) but also the rest of the flows containing personal data, for example, Service result flows containing tokens in architectures 2 and 3. Threats 1, 2, 3 and 4 may magnify the impact of this one because the external threat agent materialising this threat can attribute the leaked data to a specific data subject.

Threat 13: Breach of Data Being Processed. The data breach affects data being actively manipulated by applications, scripts or services. Sometimes, it is during processing that the data has the most value, because it is usually necessary to decrypt it in order to work with it. The elements of the DFD involved in this threat are P1 ans P2 when they process personal data.

5.8 Data Disclosure

Threat 14: Unnecessary Data Types or Volume. The data minimisation principle is not fulfilled, and personal data is not adequate, relevant, or limited to what is necessary concerning the age assurance process. E2 and E3 can materialise this threat through the Data collection flows (DF2, DF3, and DF4 in architectures 1, 2, and 3, respectively). For example, a provider may ask the user for more personal data (name, surname, address, email, postal code, etc.) than functionally required by the solution, which only requires data concerning age. In fact, in most cases it is not necessary to know the date of birth, age or age range, simply if the user's age exceeds a certain threshold. It must be taken into account that the data is sensitive, not only because it could allow minors to be located but also other people from vulnerable groups such as the old adults. Another example is providers that use more than one age assurance method for the same user/process, when with only one it is possible to check whether the user can access the requested content or service. In other cases, personal data is collected more frequently than functionally needed, for example, when the requested content is for all audiences instead of age-restricted. Furthermore, when personal data is collected of more data subjects than functionally needed, for example, children's data when only the adults' age should be ensured. Special mention deserves the cases in which special categories of data are processed without taking into account that their processing is prohibited unless any of the assumptions included in the regulation occur. This is the case, for example, of age estimation when it is based on biometric data enabling the identification of a natural person.

Threat 15: Unnecessary Processing. The purpose limitation principle is not fulfilled, and the collected personal data are processed in a manner incompatible with the initially specified, explicit and legitimate purposes concerning age assurance. E2 and E3 can materialise this threat through P1, P2, or other new processes different from those required to conduct age assurance. There are many cases where personal data may be further treated, analysed, and enriched in an unnecessary way to achieve the solution's functionality. For example, providers relying on age estimation may use users' face templates to analyse facial expressions or to recognise emotions, users' behaviour datasets to profile them or analyse their habits, etc. Retaining users' personal data long after they have been through the age assurance process may also be problematic (in S2 and S3), not fulfilling the storage limitation principle. Some providers, for example, use the data to test different methods or to train machine learning or artificial intelligence models (research&development purposes).

Threat 16: Unnecessary Exposure. In this case, personal data are made accessible to more parties than functionally necessary. E2 and E3 can materialise this threat through P1, P2, S2, and S3 or create new data flows different from those required to perform the age assurance process. For example, we have found providers propagating data to third-party providers, offering them storage services (E4) but also AI services (for age estimation), liveness tests in real-time (when the veracity of a physical document needs to be verified), additional background checks (banks or telecom operators are sometimes included in age assurance flows because they know their customers and can help to verify some of their attributes), etc. This threat also materialises when E2 or E3 share data with advertisers or use tracking and analytics services, to mention only some additional examples. Alternatively, when they publish research datasets.

5.9 Unawareness and Unintervenability

Threat 17: Lack of Information. In this case the lawfulness, fairness and transparency principle is not fulfilled and the data subject is not aware about the collection, processing, storage, or disclosure of their personal data. This threat can be materialised by E2 and E3. For example, long and incomprehensible privacy notices leading the data subject to not be aware of the collected data and metadata, the processing purposes, the third parties with whom their data will be shared, retention periods, etc. We have also found that the default settings of different solutions may not adhere to the highest standards of privacy and may employ dark patterns that can influence users to inadvertently share more data or consent to terms that they do not fully comprehend.

Threat 18: Lack of Control. In this case the data subject is not capable of exercising their data rights (the right of access, the right to rectification, the right to erasure, the right to restrict processing, the right to data portability, the right to object and the right not to be subject to a decision based solely on automated processing). This threat can be materialised by E2 and E3.

6 Discussion

Age-restricted content and services providers and third-party specialised providers must assess the potential impacts for privacy and data protection associated with their age assurance solution and their design and implementation decisions. This process should fulfil all the fundamental rights protection requirements. We hope that the model proposed in the previous section will help them.

Age assurance does not imply the sharing of personal data, such as the date of birth or the age, but the access to the needed information to protect children (access with the meaning of data use, by specific technical, legal or organisational requirements, without necessarily implying the transmission or downloading of data).

In summary, estimation-based approaches often raise specific concerns about 1) inaccuracy caused by errors, 2) data disclosure caused by unnecessary data types or volume (even involving special categories of data), unnecessary processing and unnecessary exposure or 3) unawareness and unintervenability caused by lack of information and control. Verification-based approaches often face challenges related to 1) identification due to the use of identified or identifiable data, 2) inaccuracy caused by issues with document authenticity and user binding or 3) exclusion when documents or certificates are not available.

On the other hand, specific aspects such as the use of unique identifiers, the processing of date of birth or specific age, the need to involve children in the assurance processes, the use of trusted third-party services or the excess of information in the tokens exchanged, to mention just a few examples, make it easier to materialise some of the identified threats. The threat posed by the detection of children through the age assurance solution is always one of the most critical.

Given the eighteen identified threats, some recommendations can be produced for the providers, which are not intended to be a complete or exhaustive set but can serve as an initial element of analysis:

- Linking: Avoid authentication and unique identifiers when possible because it is remarkably effortless to establish links for data subjects who have been authenticated or receive this type of identifier since all requests made within the same session will be connected. Minimise data shared in data flows; each entity should receive only what it needs to do its part of the job. For example, Service requests should not include any information about the content or service provider, and tokens included in Service result data flows should not include more information than passed/not passed (the date of birth or age are not necessary in many use cases).
- Identifying: Avoid processing credentials, attributes, tokens or claims containing identity information; it is not necessary to perform age assurance processes. Anonymous and pseudonymous solutions must be carefully assessed to guarantee robust implementations.
- Inaccuracy: Rely on non-probabilistic methods for age assurance (on age verification methods) whenever necessary, at least in extreme or boundary cases (in doubt). Exhaustively validate the functionality of the processes P1 and P2 and their way of checking age and tokens' content. Eliminate any incentive for fraud, for example, concerning the business model of all the entities involved in age assurance.
- Non-repudiation: Minimise the storage of attributable data and erase this data as soon as possible (remember metadata and logs).
- Exclusion: Provide very different alternatives and advertise them enough. Design easy-to-understand, use and manage solutions. Provide users with the necessary assistance and support.
- Detecting: Design an approach not involving children in age assurance processes; for example, it is the adult who must prove their age to access age-restricted content. Rely on robust encryption so external threat agents may intercept the data flows but not read their content.

- Data Breach: Prepare proactively to fail, to have a data breach. Minimise the collection, processing, and storage of personal data and erase this data as soon as possible (remember metadata and logs). Rely on robust encryption so external threat agents may steal the data but not read their content. Constantly re-assess the proportionate controls using state-of-the-art technologies.
- Data Disclosure and Unawareness: Remember that threats in these two categories must not be able to materialise; they imply regulatory non-compliance. Prevent them from occurring through correct risk management from the point of view of data protection in the organisation: do not collect more data than strictly necessary, ensure that data are not retained longer than strictly necessary, prevent not informed purposes and misuses, provide users with the required information, let them exercise all their rights, do not place them in a situation of urgency, weakness or lack of power in which they must accept any processing, etc.

It is important to note that some threats identified in current solutions cannot be avoided with the available architectures. Therefore, researchers, innovators, and privacy practitioners should explore promising but not yet mature market directions. The first one should be to perform the age assurance process on the user's local devices, making the necessary adaptations to the applications that allow access to content and services or to the operating systems on different devices [6]. The second is to explore the potential of decentralized and user-empowering solutions, such as European digital identity wallets, to solve this problem [12] with the asynchronous approach of the Architecture 3. Furthermore, the third, probably in combination with the previous, is to rely on privacy-enhancing techniques, such as selective disclosure or zero-knowledge proof, to guarantee user unlinkability or anonymity [27].

Concerning regulators and standardization bodies, they should focus on making global and collaborative advances since local efforts in today's digital world will have limited impacts. Standardizing terminology and nomenclature would help in this direction. Regulators should also make clear that the deployment of age assurance solutions must always be aligned with the highest privacy and data protection principles. They should also cooperate with the industry and NGOs (for example, parents or educators associations or professional colleges) to establish global certification schemes, codes of conduct, and regular auditing models which ensure the highest levels of privacy and data protection in the deployed solutions.

Given the potential implications, demonstrated in this research, that all these decisions have, data protection authorities should advise on all these processes, helping to achieve compliance and ensuring enforcement.

7 Conclusions

Ensuring children's fundamental rights is paramount, and age assurance solutions play a pivotal role in protecting them from online inappropriate content.

Deploying these solutions following privacy-conscious practices and data protection regulations is crucial. This paper aims to contribute to the ongoing discourse on this subject and foster a safer digital environment for our younger generations without limiting the fundamental rights of all Internet users.

This paper provides a comprehensive analysis of the current age assurance solutions and their potential implications on privacy and data protection. Our innovative threat model, LIINE3DU, identifies specific threats associated with these solutions. The insights gained from our research have significant implications for solution design, policy formulation, and awareness initiatives. Service providers can utilize our threat model to develop more privacy-respecting and compliant age assurance solutions. Policymakers can leverage our findings to shape regulations that promote safety while safeguarding privacy and data protection. Researchers can focus on advancing in the most essential and promising directions. Furthermore, our study provides society with a nuanced understanding of the implications and challenges involved in age assurance, emphasizing the need for privacy-focused solutions.

This paper shows the results of using our threat modelling method for the first time. If new iterations were required because new solutions and architectures must be added to complete the threat model, this might modify the produced threat list. However, the users of this method would work as illustrated in the previous sections, following the same steps in the same way.

We are now working on analysing, in relation to the first categories of threats (LIINEDD, excluding the third D and the U, from a different nature), what risks to the rights and freedoms of users each threat specifically entails and how the probabilities and impacts associated with these risks could be mitigated. This would be essential to adopt a risk-based approach, so common in the data protection field to decide about the implementation of the appropriate technical and organisational measures to integrate the necessary safeguards into the processing in order to meet the requirements of regulation, to ensure a proper level of security, etc.

Acknowledgement. This study was funded by the Agencia Española de Protección de Datos (AEPD), the Spanish Data Protection Authority. The authors want to thank the entire AEPD team and specifically the staff of the Innovation and Technology Division for their involvement in this project and all their generous comments and contributions.

Disclosure of Interests. The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

1. Directive 2018/1808 of the European Parliament and of the Council of 14 November 2018 amending Directive 2010/13/EU on the coordination of certain provisions laid down by law, regulation or administrative action in member states concerning the provision of audiovisual media services (Audiovisual Media Services Directive) in view of changing market realities (2018). https://eur-lex.europa.eu/eli/dir/2018/1808/oj
2. Age Appropriate Design Code (2022). https://californiaaadc.com/
3. Regulation 2022/2065 of the European Parliament and of the Council of 19 October 2022 on a single market for digital services and amending Directive 2000/31/EC (Digital Services Act) (2022). https://eur-lex.europa.eu/eli/reg/2022/2065/oj
4. Online Safety Act (2023). https://www.legislation.gov.uk/ukpga/2023/50/enacted
5. 5RightsFoundation. But how do they know it is a child? Age assurance in the digital world (2021). https://5rightsfoundation.com/uploads/But_How_Do_They_Know_It_is_a_Child.pdf
6. AEPD. Protection of minors on the Internet- Technical note with the description of the proofs of concept (2023). https://www.aepd.es/guides/technical-note-proof-of-concept-age-verification-systems.pdf
7. Bertrand, A., Diaz, M.C., Hair, E.C., Schillo, B.A.: Easy access: identification verification and shipping methods used by online vape shops. Tobacco Control (2024)
8. Blake, P.: Age verification for online porn: more harm than good? Porn Stud. **6**(2), 228–237 (2019)
9. Brennen, S., Perault, M.: Keeping kids safe online: how should policymakers approach age verification? The Center for Growth and Opportunity (2023)
10. BSI. PAS 1296:2018 - Online age checking. Provision and use of online age check services. Code of Practice (2018). https://knowledge.bsigroup.com/products/online-age-checking-provision-and-use-of-online-age-check-services-code-of-practice?version=standard
11. CNIL. Online age verification: balancing privacy and the protection of minors (2022). https://www.cnil.fr/en/online-age-verification-balancing-privacy-and-protection-minors
12. EC. Second meeting of the Task Force on age verification (2024). https://digital-strategy.ec.europa.eu/en/news/second-meeting-task-force-age-verification
13. EDRi. Online age verification and children's rights (2023). https://edri.org/our-work/policy-paper-age-verification-cant-childproof-the-internet/
14. Egan, K.L., Villani, S., Soule, E.K.: Absence of age verification for online purchases of cannabidiol and delta-8: implications for youth access. J. Adolesc. Health **73**(1), 195–197 (2023)
15. euCONSENT. euCONSENT project (2024). https://euconsent.eu/home-euconsent-project/
16. de Farias, J.C.L.A., Carniel, A., de Melo Bezerra, J., Hirata, C.M.: Approach based on STPA extended with STRIDE and LINDDUN, and blockchain to develop a mission-critical e-voting system. J. Inf. Secur. Appl. **81**, 103715 (2024)
17. Georgiadis, G., Poels, G.: Towards a privacy impact assessment methodology to support the requirements of the general data protection regulation in a big data analytics context: a systematic literature review. Comput. Law Secur. Rev. **44**, 105640 (2022)

18. ICO. Age appropriate design: a code of practice for online services (2024). https://ico.org.uk/for-organisations/uk-gdpr-guidance-and-resources/childrens-information/childrens-code-guidance-and-resources/age-appropriate-design-a-code-of-practice-for-online-services/
19. ISO. ISO/IEC WD 27566-1 Information security, cybersecurity and privacy protection -Age assurance systems- Framework (2023). https://www.iso.org/standard/88143.html
20. Jarvie, C., Renaud, K.: Are you over 18? A snapshot of current age verification mechanisms. In: Dewald Roode Workshop (2021)
21. LINDDUN. Privacy threat modeling (2024). https://linddun.org/
22. Nash, V., O'Connell, R., Zevenbergen, B., Mishkin, A.: Effective age verification techniques: lessons to be learnt from the online gambling industry. Available at SSRN 2658038 (2012)
23. Nweke, L.O., Abomhara, M., Yayilgan, S.Y., Comparin, D., Heurtier, O., Bunney, C.: A LINDDUN-based privacy threat modelling for national identification systems. In: Proceedings of the IEEE Nigeria 4th International Conference on Disruptive Technologies for Sustainable Development, pp. 1–8 (2022)
24. OECD. Children in the digital environment: revised topology of risks (2021). https://doi.org/10.1787/9b8f222e-en
25. Pasquale, L., Zippo, P., Curley, C., O'Neill, B., Mongiello, M.: Digital age of consent and age verification: can they protect children? IEEE Softw. **39**(3), 50–57 (2020)
26. Robles-González, A., Parra-Arnau, J., Forné, J.: A LINDDUN-based framework for privacy threat analysis on identification and authentication processes. Comput. Secur. **94**, 101755 (2020)
27. Ronis, J.: Don't trust when you can verify: a primer on zero-knowledge proofs (2024). https://www.wilsoncenter.org/article/dont-trust-when-you-can-verify-primer-zero-knowledge-proofs
28. Sas, M., Mühlberg, J.T.: A risk-based evaluation of available and upcoming age assurance technologies from a fundamental rights perspective (2024). https://www.greens-efa.eu/en/article/study/trustworthy-age-assurance
29. Williams, R.S., Phillips-Weiner, K.J., Vincus, A.A.: Age verification and online sales of little cigars and cigarillos to minors. Tobacco Regulat. Sci. **6**(2), 152 (2020)
30. Wuyts, K., Sion, L., Joosen, W.: LINDDUN go: a lightweight approach to privacy threat modeling. In: Proceedings of the IEEE European Symposium on Security and Privacy, pp. 302–309 (2020)
31. Xiong, W., Lagerström, R.: Threat modeling-a systematic literature review. Comput. Secur. **84**, 53–69 (2019)
32. Yar, M.: Protecting children from internet pornography? A critical assessment of statutory age verification and its enforcement in the UK. Policing: Int. J. **43**(1), 183–197 (2020)