

Subscribe to DeepL Pro to edit this document.  
Visit [www.DeepL.com/pro](https://www.deepl.com/pro?cta=edit-document) for more information.

HealthID - Information Security

Introduction:

Having personal information on the blockchain is not safe for several reasons. The blockchain is designed to be immutable, which means that once data is recorded on it, it cannot be altered or deleted. This is a fundamental feature for guaranteeing the integrity of records. However, if personal information is recorded irreversibly, any error or breach of privacy cannot be corrected.

Instead of storing personal information directly on the blockchain, it is safer to use it to manage access permissions and control the sharing of personal data stored securely on external systems, following strict privacy regulations. This helps to protect individuals' privacy while reaping the benefits of the technology.

In our project we are dealing with the user's medical information, so we need to treat this information and guarantee the user's privacy, and only then insert their medical records into the database.

1. identifying sensitive data:

In the application, the user will upload their medical examination. Once uploaded, the file needs to be converted into text and OCR (Optical Character Recognition) will be used to analyze and search for sensitive information (only non-sensitive data or pseudonymized information should be stored on the blockchain). Personal information that is not essential for the purpose of the transaction on the blockchain (releasing access to medical data) will be removed and stored centrally using anonymization techniques.

2. Physical separation of data:

One way of separating the data is to create different types of records or data structures, for example: **metadata record,** this record will contain control information and metadata about the file, such as unique identifiers, creation dates, document types and any other non-sensitive information that is useful and **content record,** this record will contain the actual content of the documents, but with the sensitive data anonymized. This means that personal information, such as names and identification numbers, is replaced by pseudonyms.

3. Cross-referencing:

To ensure that the metadata record and the content record are related, you can use a unique identifier (such as a hash code or ID) that links these records. This way, you can retrieve sensitive data when necessary, but only if the appropriate authorization is granted.

4. Access control:

We will implement an access control system through smart contracts to ensure that only authorized parties can access sensitive data. This allows you to set detailed permissions on who can see or use this data.

5. Auditing and traceability:

We will record all transactions and activities related to data access and use, including authorizations granted and revoked. This provides traceability and a complete history of interactions with sensitive data.

6. Regulatory Compliance:

All anonymization, pseudonymization and data storage practices must comply with data privacy regulations, such as the GDPR in the European Union.

7. User education:

Users need to be educated about how their data is protected and how they can control access to this information, which can be done periodically.