République Algérienne Démocratique et Populaire
Ministère de l'Enseignement Supérieur et de la Recherche Scientifique
Université Benyoucef BENKHEDDA-Alger1

Faculty of Sciences
Computer-Science departement

End of Study Project for obtaining the diploma of the
Computer science degree

Option : Informatic Systems

Thème

Conception and realization of an electronic document management system

Framed by:

- M Maredj Azze-addine

Realized by :

-Djalal Aymen

- Djemaoui Ahmed

2023/2024

Remerciements

Je remercie / Nous remercions ...
Rappelez-vous de remercier LaTeX:-)

Dédicaces

Lorem ipsum ...

Noms des étudiants 1 et 2

resume

Contents

List of Figures

List of Tables

Chapter 1

1.1 introduction

In today's world, businesses are constantly bombarded with a large amount of information. From this flood of information, it is important to be able to retain important information, store it, and be able to find it easily and quickly. Electronic Document Management (EDM) tools are solutions to this problem that businesses face.

1.2 What is an EDM

An EDM is a system that helps businesses manage their electronic documents. It can be used to store, organize, and retrieve documents from a central location. EDMs can also be used to automate workflows, track document changes, and ensure compliance with regulations.

1.2.1 Document Lifecycle in an EDM

Aquisition

La capture des documents papier, réalisée via un scanner, et celle des documents électroniques produits par l'entreprise, sont gérées par les utilisateurs à travers leurs logiciels externes, ainsi que les courriers électroniques, pièces jointes, documents provenant d'autres applications telles que les formulaires web et les scans mobiles.

traitement

Cela comprend le choix du format, la compression, l'indexation, la classification et le nommage des documents.

Compression

consiste à réduire l'espace occupé par l'ensemble des données concernées. Il s'agit d'un processus de modification, d'encodage ou de conversion de la structure des bits (représentation binaire des informations) qui peut s'appliquer à tous les types de fichiers

Formats des fichiers :

Après l'acquisition des documents le problème résidera dans le choix du format le plus adéquat. En effet Les formats de sauvegarde des données constituent un problème important dans le stockage des documents.

Indexation

L'indexation est une procédure consistant à associer ou à marquer des documents ou d'autres contenus avec des informations permettant par la suite de rechercher et récupérer ces documents.la (GED) utilise une ou plusieurs types d'indexation.

indexation par mots clés

L'indexation par mots-clés consiste à attribuer des mots-clés à un document pour en décrire le contenu. Ces mots-clés peuvent être tirés d'un langage d'indexation ou d'un thésaurus. L'indexation par titre consiste à utiliser le titre du document comme critère d'indexation.

indexation par métadonnées

Les métadonnées sont des informations structurées qui décrivent un document, telles que le titre, l'auteur, la date de création, le format, etc. L'indexation par métadonnées consiste à utiliser ces informations pour organiser et retrouver des documents. Les métadonnées peuvent être externes à la ressource qu'elles décrivent (comme une notice dans un catalogue) ou internes (comme une balise méta d'une page web).

Utilisation quotidienne

- **Document search:** The search for documents is done by means of a language that can be reduced to its simplest expression. Several techniques exist and are implemented in GED systems, either independently and exclusively from each other, or for some of them in combination. The main search methods are:
- **Boolean search** which is to present the query in the form of a Boolean expression of keywords belonging to the dictionary or the thesaurus, words separated by boolean operators (et, ou, sauf ...).
- **Full text or full text search** which avoids the prior indexing of documents; they are described by themselves, i.e. by the string that constitutes them. Boolean research also applies to this case. However, this technique is often associated with so-called natural-language questioning methods, i.e. based on a linguistic analysis of a user's query and leading to the development of an "internal" query from which the search is carried out.
- **Search by navigation or hypertext technique** this technique presupposes a pre-structuring of documents with the implementation of hypertext links

Control de la diffusion :

There are two techniques for disseminating documents.

- **mode push :**is the provision of the documents that the user will search in the database. It must then connect to the (GED). It is possible for him to view the documents after a search and selection (if he is permitted to access the document).
- **mode pull :**is the distribution of the document. The system automatically transfers documents (often via e-mail) to previously designated recipients. It is then necessary to manage the mailing lists properly (change of email addresses, departure from the company,...) so that the right people can have the right documents

Editing document:

Document editing is an essential feature that allows users to make changes to existing files. This ability to modify is often accompanied by a version control, which records and organizes the various iterations of a document over time.

Version control ensures traceability of changes made to a document, recording every step of its evolution. This allows users to track changes made, return to previous versions if necessary, and understand who made what changes and when.

With this feature, users can collaborate effectively on a document without fear of losing data or working on an outdated version. Version control also ensures document integrity and facilitates compliance with information management regulations.

In short, editing documents in a GED, coupled with version control, provides a secure and organized way to manage changes to files, improving collaboration, security, and compliance within the enterprise.

Archiving:

In reality, as soon as they are created, the documents must therefore be integrated into a reflection on information management whose general framework is the passage of documents and data through three successive ages:

- **intermediate archives:** This age begins when the document is no longer frequently used (e.g. when the case concerned has ended), but

  The service that produced it may still need it to deal with appeals, possible limitations, or other cases. In the digital context, this often involves restricting access to data in order to ensure its security and confidentiality; These two ages form the Administrative Utility Period (ADP) during which the archives are under the responsibility of the service that produced them.

- **The definitive archives:** Beyond the primary value of the archives that justifies their preservation during their ADP, a scientific evaluation must distinguish: documents that present a permanent historical, scientific, statistical or public interest*, which must be kept indefinitely for the knowledge of our society by our descendants. of those who are deprived of it, and who are then doomed to destruction.

destruction:

The destruction of archives is determined by the archiving policy, since the statutory preservation period is no longer to be respected, because it expires.

1.2.2 Security in an EDM

This software is one of the indispensable dematerialization tools, adaptable to any business and capable of managing any type of files (invoice, mail, contract, payroll...).

Cryptography:

This is the operation of making information indistinguishable without a private key. Encryption can protect files stored in GEDs as well as internal communications to GED applications.

ISO 27001 Certification:

This international standard defines best practices for information protection. An ISO 27001 certification attests to the high level of maturity of an organization's security policy and its GED system.

Access rights management:

This process involves defining and controlling permissions granted to different users to access enterprise resources. In GEDs, this management must be rigorous in order to maintain the confidentiality of sensitive documents.

Compliance with industry regulations and standards:

Adoption of good data security practices is essential to comply with legal and industry requirements such as GDPR, HIPAA, SOX, PCI DSS, etc...

audit and continuous monitoring:

Periodic audits and continuous monitoring are required to verify the effectiveness of the security policy and the implementation of security measures in LDCs.

Employee training and awareness-raising:

Continuously train and raise awareness of data security to reduce the risks associated with human negligence

Chapter 2

Study and Conception Phase

2.1 Choice of UML Language

UML is a unified object modeling language. It was born from the merger of three object-oriented methods: Booch, OMT, and OOSE. UML 2 is a major evolution of the language. Modeling and formalization using a standardized vocabulary, moreover, object-oriented, give the method all its interest. Formalization and modeling indeed facilitate the definition of the problem to be addressed and understanding by all the main stakeholders. Since its inception, it has had use case diagrams that address the upstream phases by representing the modes of use of a system without focusing on its functioning and implementation choices. The advantage of this type of diagram is that it allows formalizing a problem from the initial phases, in a form intelligible to all stakeholders, whether they are computer scientists or not. In design, I will use four types of diagrams:

- Use case diagrams.
- Sequence diagrams.
- Activity diagrams.
- Class diagrams.

2.2 Conception

2.2.1 Use case diagrams

-Use case diagrams are UML diagrams used to provide an overview of the functional behavior of a software system.

here is the functionalities that our system provide

Label	Description
Register	Fill a form and wait for your activation mail on your Email
Login	Authentication with your Email and password
Upload	Upload documents into the server
Consult list of documents	Display list of documents
Download	Download document from the server to your local machine
Search	Search for documents in the list of the document that you uploaded and received
Diffuse	Diffuse the document that you uploaded to other users with setting their permission
Validate document	Only responsible users can validate documents in the server so they can be used in the organization
Admin Functionalities	
Account activation	Only the admin can activate an account of a user through his interface
Edit user	Edit user roles and permissions
Delete user	Deleting a user while keeping a record of the documents associated with him

- 1-General Use Case Diagram

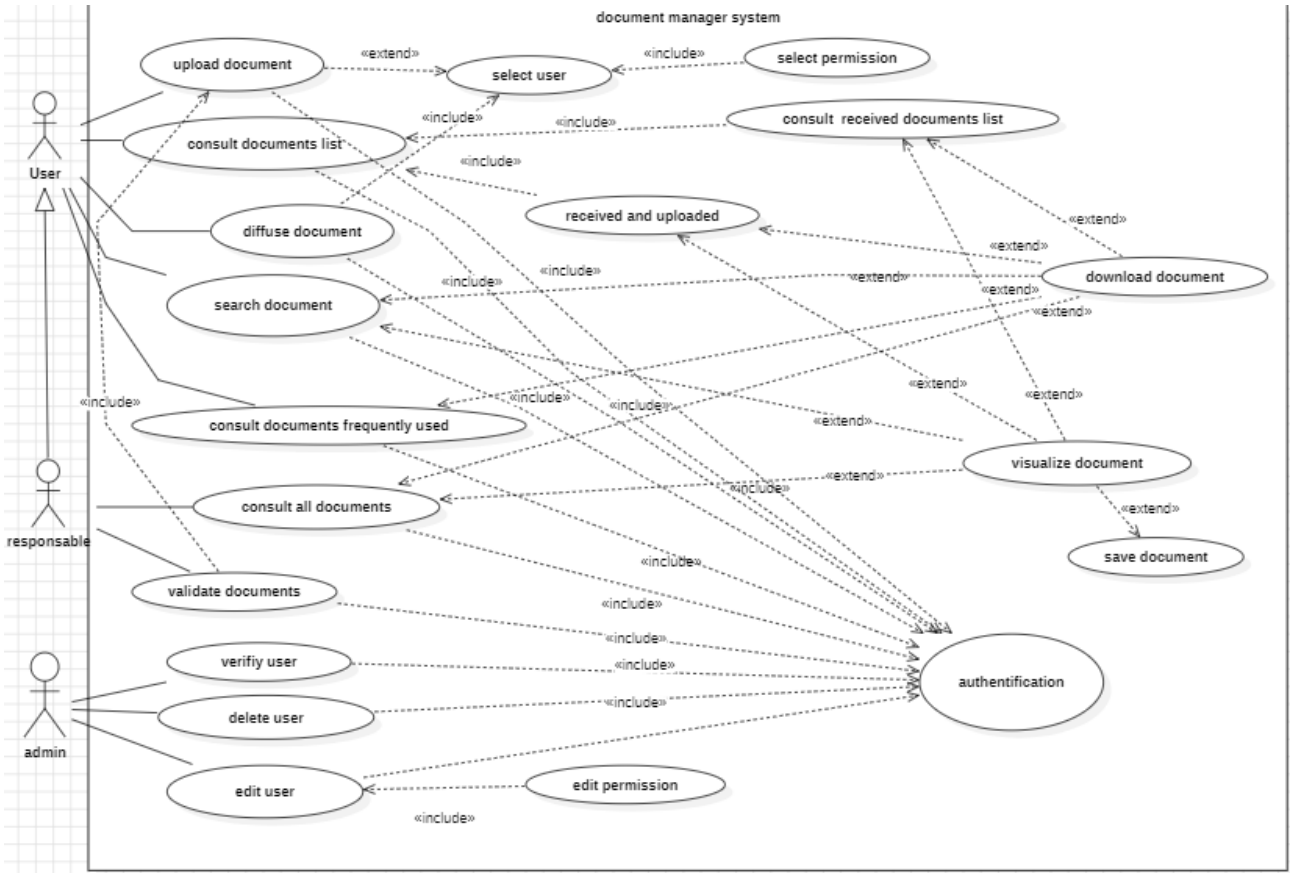

Figure 2.1: Use case diagram representing all functionality that the system provides

- **The diagrams of each actor below :(Supposedly all actor has been already authentified)**
- 2-User diagram

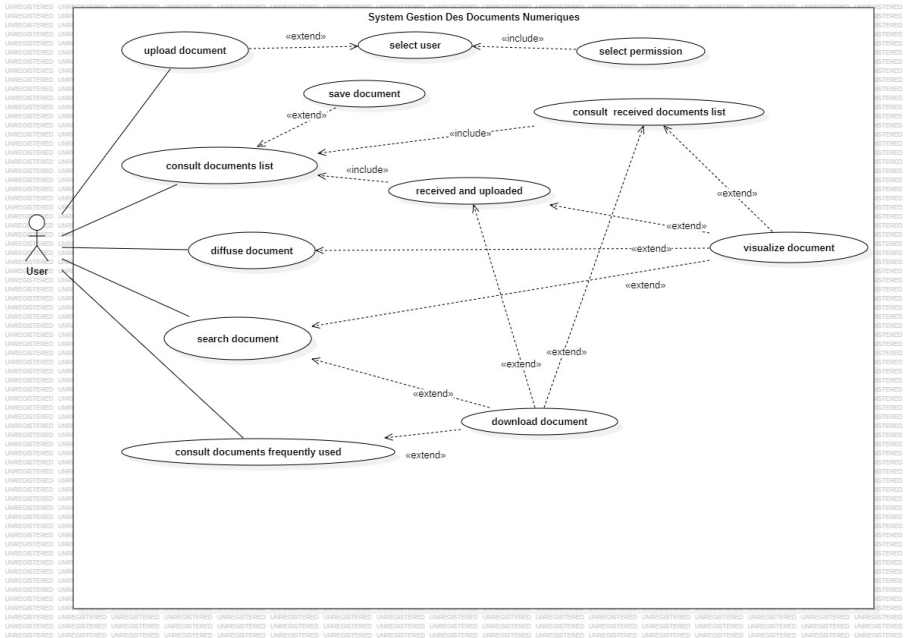

Figure 2.2: use case diagram of the actor user

- 3-Reponsible diagram

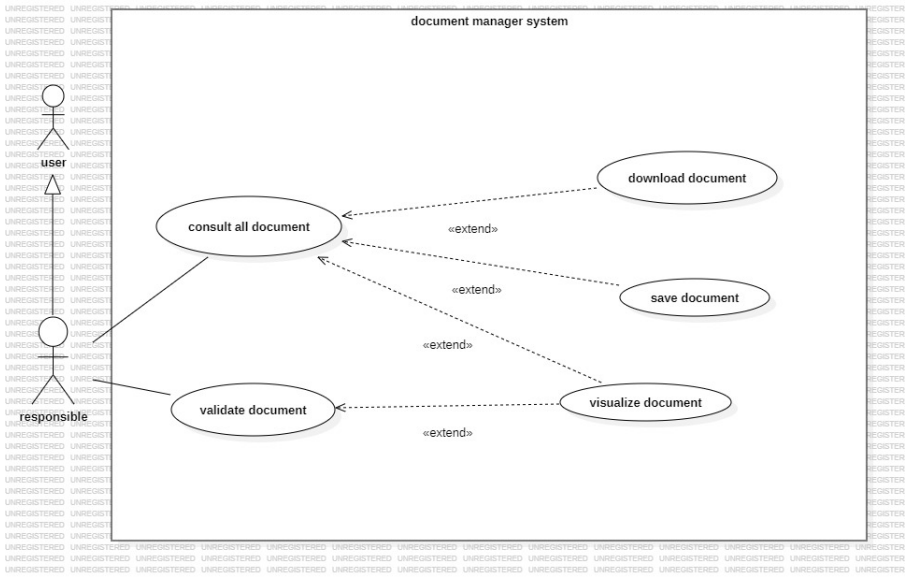

Figure 2.3: Use case diagrame of the actor responsible

- 4-Admin diagram

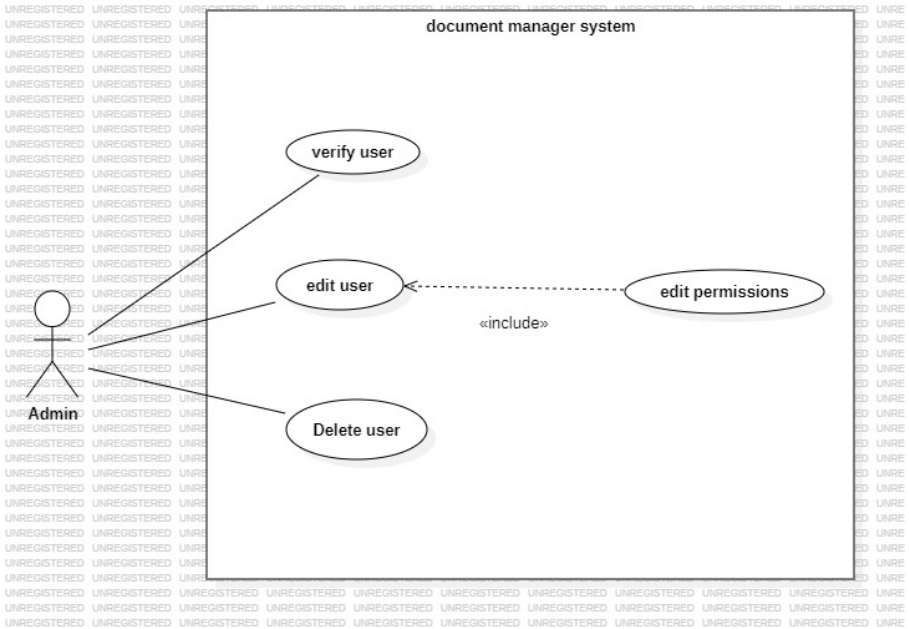

Figure 2.4: Use case diagram of the actor admin

2.2.2 Sequence diagrams

A sequence diagram is a type of interaction diagram in UML (Unified Modeling Language) used primarily to depict the interaction between objects or components within a system in terms of the messages exchanged over time. It illustrates the flow of messages and interactions among the various elements of a system in a sequential manner, typically showing the order in which interactions occur.

some sequnce diagram for the essential functionalies

- **Sequence representing ⟨⟨Upload and validate documents⟩⟩**

Description	Upload and validate document
Actors	User, responsible
Pre-condition	User opens the upload document window
Goal	Upload document on the server and verify

Figure below shows the sequence diagram ⟨⟨upload and validate document⟩⟩

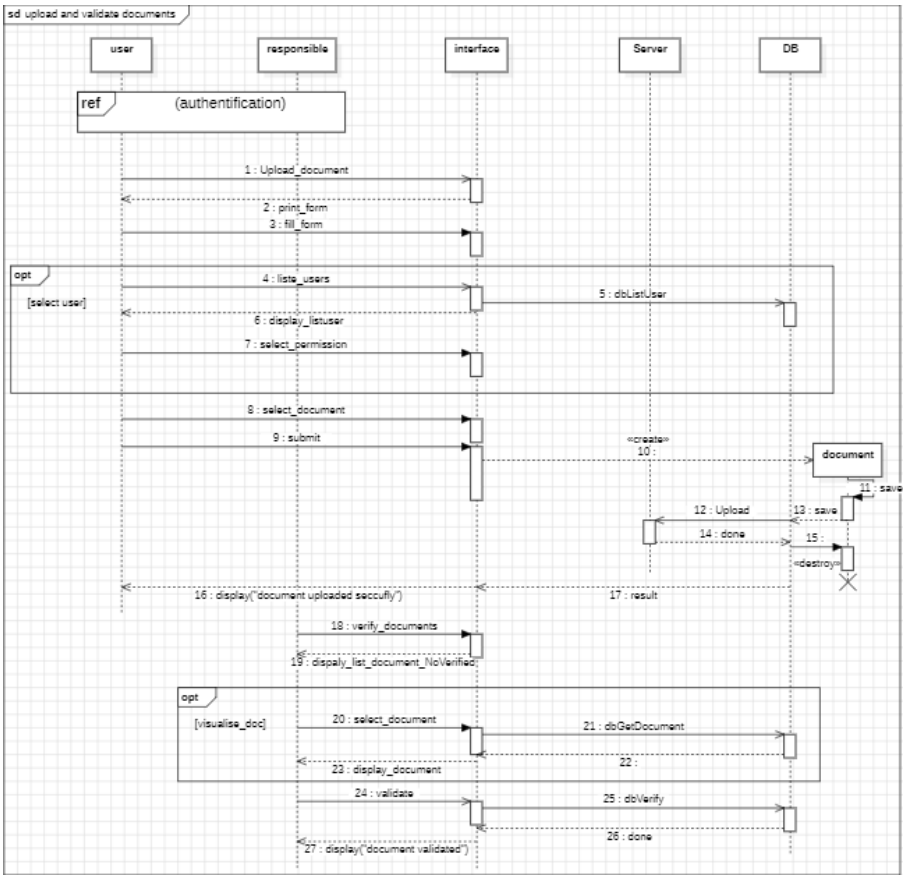

Figure 2.5: sequence diagram upload and validate document

- **Sequence representing ⟨⟨Consult documents list⟩⟩**

Description	Consult documents list
Actor	User
Pre-condition	User opens the Documents list window
Goal	– Consult document that the user uploads and receives – The possibility to visualize or download if the user has the permission

Figure below represents sequence diagram ⟨⟨Consult document list⟩⟩

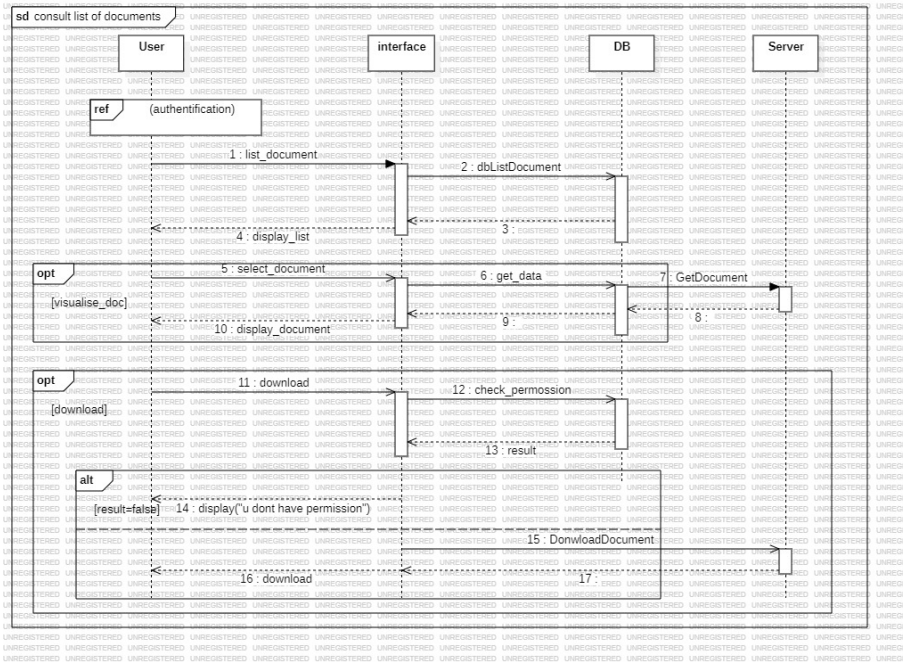

Figure 2.6: Sequence diagram consult documents list

- **Sequence representing ⟨⟨Search document⟩⟩**

Description	Search document
Actor	User
Pre-condition	User opens Search window
Goal	Find the document you need immediately just by entering its metadata

Figure below represents sequence diagram ⟨⟨Search document⟩⟩

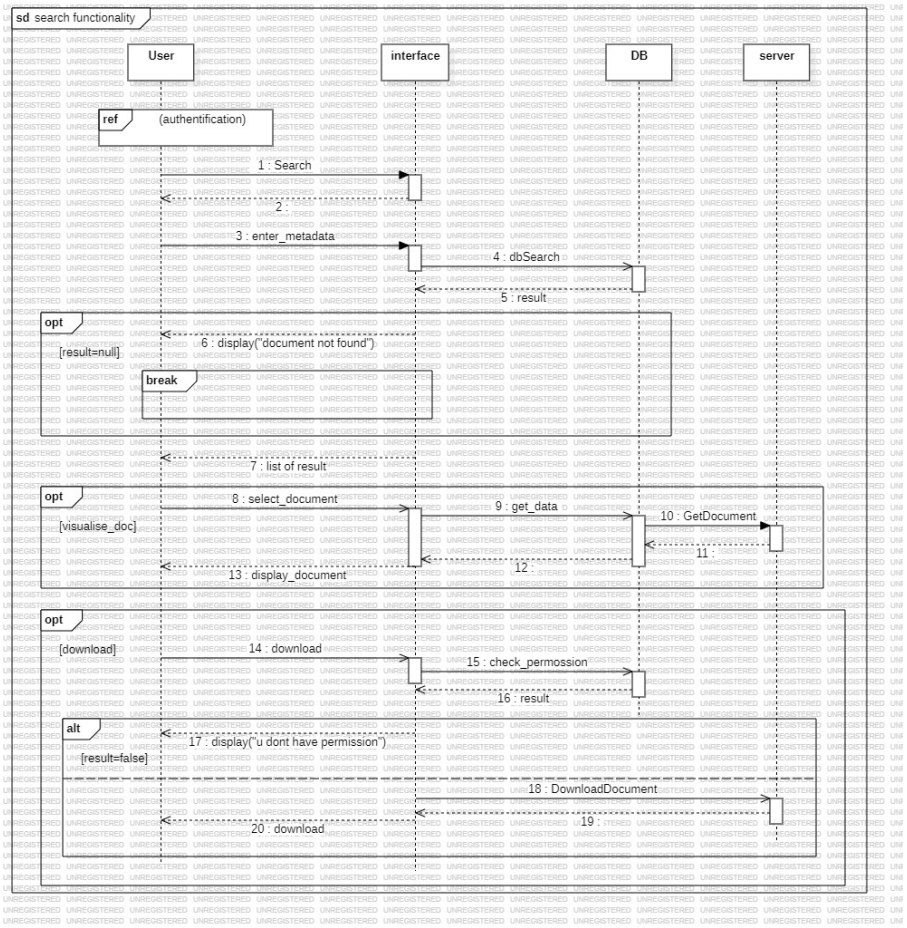

Figure 2.7: Sequence diagram search document

Bibliography

[1] Autour de la ged. `https://www.ged.fr/`. Visité: 2024-02-25.

[2] La ged: gestion électronique des documents. `https://www.archivage-numerique.fr/ged`. Visité: 2024-02-25.

[3] Géraldine DESBETES, Lucile LEROY, and Anne-Gaëlle LIEBERT. La gestion electronique des documents. *Université de Lille*, 3:44, 2008.

[4] Núria Forcada Matheu. *Life cycle document management system for construction*. Universitat Politècnica de Catalunya, 2005.

[5] Comité interministériel aux Archives de France. Référentiel général de gestion des archives. page 60, 2013.

[6] Annie Le Saux. La gestion électronique de documents. 1992.

[7] Dr M. P. Satija. Knowledge organization and processing: Classification. 2018.