Document Processing: Assignment: Part 1

Software Architecture: Project Assignment 2014-2015

The project involves the development of a system for electronic document processing as a service. It consists of three parts: part 1 (3 lab sessions) involves both analysis of the problem domain and requirements elicitation and engineering. Part 2a (2 lab sessions) involves the initial design and inception of a software architecture for the document processing system. Finally, in part 2b (3 lab sessions) the initial architecture is extended and completed.

This document presents the specific assignment for part 1. In this part of the assignment, you are asked to extend and complete (i) our initial analysis of the problem domain in Section 1, (ii) our initial set of functional requirements in Section 2 and (iii) our initial set of non-functional requirements in Section 3. Section 4 provides more details about the requested format of the report.

1 Domain Analysis

Figure 1 presents the initial domain model, describing already some concepts related to the interaction between eDocs and its users.

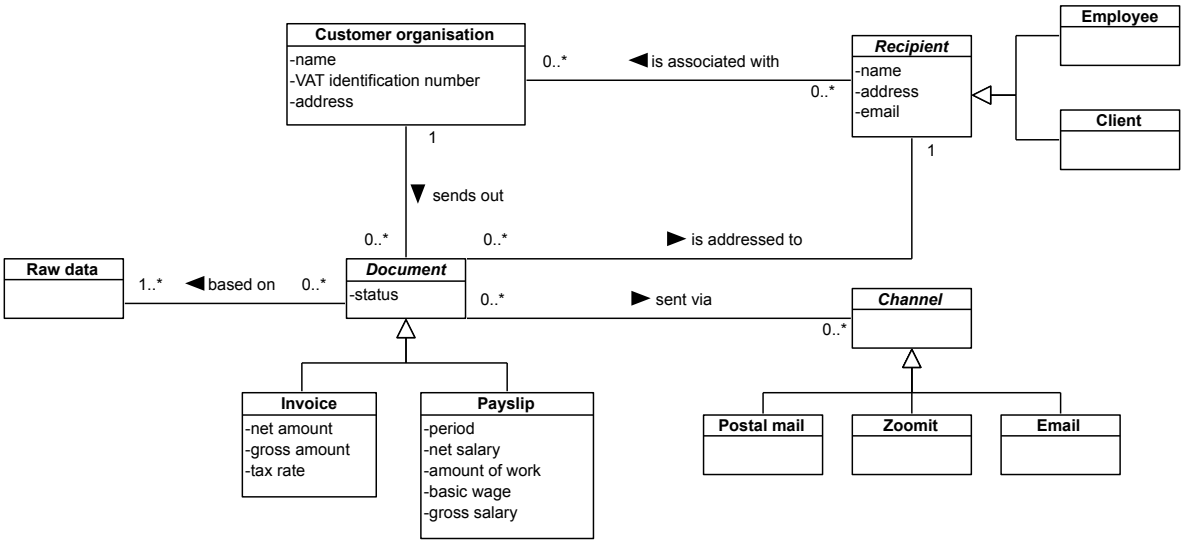

Figure 1: Initial domain model for remote measurement and monitoring.

Initial constraints. Some additional constraints that apply to this domain model are:

- The `basic wage` and a`mount of performed work` on a `payslip` are expressed in function of the same units.
- `Customer organisations` only send out documents related to `Recipient` associated to them.
- `Recipients` only receive `Documents` addressed to them.

Initial glossary.

- **Customer organisation:** an organisation that outsources its document processing for payslips and/or invoices to eDocs.
- **Invoice:** a commercial document issued for the sale of goods and/or services from a seller to a buyer.
- **Payslip:** a document sent out by companies to their employees that describes how their salary for a certain period is calculated.
- **Raw data:** the unformatted information that is represented in a document.
- **Recipient:** a person or organisation that receives a document, e.g. an employee of a customer organisation or a client subscribed to the services of a customer organisation.

Assignment. You are asked to extend and complete the analysis of the problem domain to cover missing elements and aspects related to generating, delivering and storing documents. Specifically, you will extend the domain model (in one or more diagrams), including the most relevant constraints, as well as the glossary.

2 Functional requirements

Sections 2.1 to 2.6 present six initial use cases related to registering to eDocs, authenticating to eDocs and using the personal document store. Figure 2 presents the overview use case diagram.

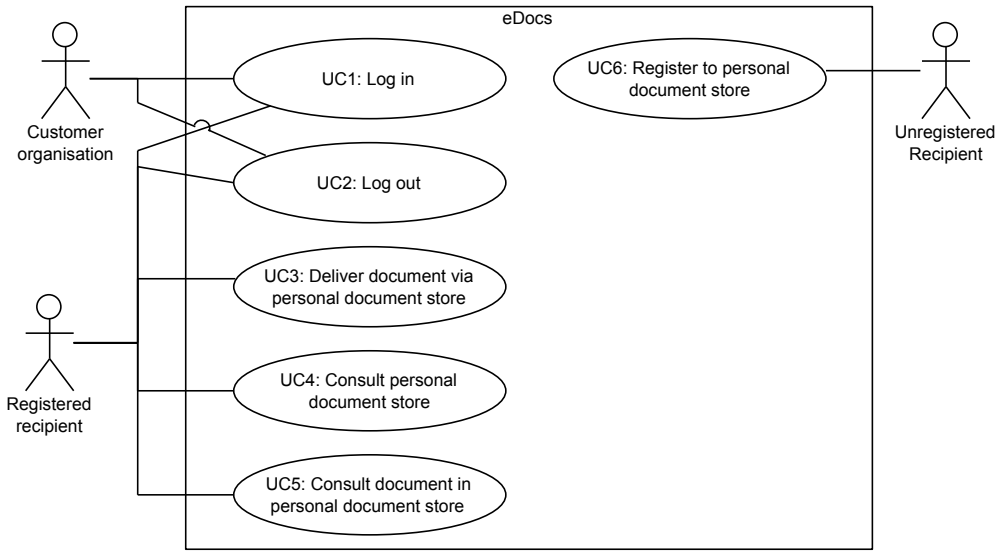

Figure 2: Use case diagram.

Assignment. In this part, you are asked to complete this set of functional requirements[1] to cover at least the functionality illustrated in the scenario presented in Section 6 of the case description document.

2.1 *UC1*: Log in

- **Summary:** A Customer organisation or Registered recipient wants to use the application and logs in by providing authentication credentials.
- **Primary actor**: Customer organisation or Registered recipient
- **Secondary actors:** None
- **Interested parties:**
  - eDocs: wants to protect the data of its customer organisations and users by means of authentication.
  - Customer organisations: want to have sole access to their data.
  - Recipients: want to have sole access to their documents.
- **Preconditions:**
  - The Primary actor is registered.
- **Postconditions:**
  - The Primary actor is logged in and can use the System.
- **Main scenario**:
  1. The Primary actor indicates he wants to log in to the System and provides authentication credentials, e.g. user name and password.

[1] To write the textual use cases, comply to the format used in the initial set of use cases.

2. The System verifies the provided authentication credentials, confirms successful log in to the Primary actor and logs him or her in.

- **Alternative scenarios:**

2a. If the provided credentials are incorrect, the system notifies the Primary actor of this. Restart from step 1.

- **Remarks:**

  - This use case abstracts from the method of authentication that is used (cf. step 2 and 3). For example, authentication can be done using techniques such as user name-password, security tokens, IP-based, using cryptographic keys etc. Among others, the employed method depends on the type of Primary actor.

2.2 *UC2*: Log out

- **Summary:** A logged in Customer organisation or Registered recipient signs out of the System.
- **Primary actor**: Customer organisation or Registered recipient
- **Secondary actors:** None
- **Interested parties:**

  - eDocs: wants to protect the data of its customer organisations and users by means of authentication.
  - Customer organisations: want to have sole access to their data.
  - Recipients: want to have sole access to their documents.

- **Preconditions:**

  - The Primary actor is logged in (cf. *UC1*: Log in).

- **Postconditions:**

  - The Primary actor is logged out.

- **Main scenario:**

  1. The Primary actor indicates he wants to log out of the System.
  2. The System logs the Primary actor out and indicates success to the Primary actor.

- **Alternative scenarios:** None
- **Remarks:**

  - This use case abstracts from the method of authentication that is used (cf. *UC1*: Log in).

2.3 *UC3*: Deliver document via personal document store

- **Summary:** The System delivers a document to a Registered recipient by adding it to his or her personal document store and notifying him or her of the new document.
- **Primary actor**: Registered recipient
- **Secondary actors:**
- **Interested parties:**

  - eDocs Administrators: want document processing to happen with as little human interaction as possible.
  - Registered recipients: want to receive their documents correctly and as early as possible.
  - Customer Organizations: want their documents to be delivered.
  - Customer administrators: want to have a clear and detailed overview of the status of their document processing jobs.

- **Preconditions:**

  - A document has been generated that should be delivered via e-mail (as indicated by the customer organisation).
  - The addressee of this document is a Registered recipient.

- **Postconditions:**
  - A new document has been added to the personal document store of the Registered recipient.
  - The System has sent a notification to the addressee of the document. This e-mail contains a short description of the document and a link to the document.
  - The System has marked the corresponding job as added to personal document store.
  - If this document is not part of a recurring batch of document processing jobs, the System has billed the customer organisation for it.

- **Main scenario:**
  1. The System adds the document to the personal document store of the Registered recipient.
  2. The System constructs an e-mail. This e-mail is addressed to the name of the Registered recipient and contains a short description of the received document (i.e. the sender of the document, the type of the document and the date at which the document was sent) and a link to the document.
  3. The System sends this e-mail to the e-mail address of the Registered recipient.
  4. The System marks the corresponding document processing job as added to personal document store.
  5. If this document is not part of a recurring batch of document processing jobs, the System adds the cost of delivering the generated document to the bill of the customer organisation.

- **Alternative scenarios:** None
- **Remarks:** None

2.4 *UC4*: Consult personal document store

- **Summary:** A Registered recipient consults an overview of all received documents in his or her personal document store.
- **Primary actor**: Registered recipient
- **Secondary actors:** None
- **Interested parties:**
  - Registered recipients: want to maintain a clear overview of all the documents they have received so far.
  - Customer organisations: want their documents to be delivered correctly.

- **Preconditions:**
  - The Registered recipient is logged in (cf. *UC1*: Log in).

- **Postconditions:**
  - The Registered recipient has received an overview of all documents he or she received and can request a detailed view on each of these documents.

- **Main scenario:**
  1. The Registered recipient indicates that he or she wants to consult his or her personal document store.
  2. The System looks up all documents sent to the Registered recipient.
  3. The System provides an overview of all the documents in the personal document store of the Registered recipient, e.g. as a table or a list, together with the possibility to consult each individual document (cf. *UC5*: Consult document in personal document store) and the possibility to download each such document as a PDF.

- **Alternative scenarios:**

  3a. If the Registered recipient has not received any documents yet, the System indicates this to the Registered recipient. The use case ends.

- **Remarks:** None

2.5 *UC5*: Consult document in personal document store

- **Summary:** A Registered recipient consults a specific received document.
- **Primary actor**: Registered recipient
- **Secondary actors:** None
- **Interested parties:**
  - Registered recipients: want to get a clear overview of each of their documents.
  - Customer organisations: want their documents to be delivered.
  - Customer administrators: want to have a clear and detailed overview of the status of their document processing jobs.
- **Preconditions:**
  - The Registered recipient is logged in (cf. *UC1*: Log in).
- **Postconditions:**
  - The Registered recipient has received a detailed view on the requested specific document.
  - The document has been marked as received.
- **Main scenario:**
  1. The Registered recipient indicates that he or she wants to view a received document, e.g. in the overview of received documents (cf. *UC4*: Consult personal document store) or by following the link in a notification (cf. *UC3*: Deliver document via personal document store).
  2. The System looks up the requested received document and its meta-data, e.g. the sender of the document and the date on which it was received, and presents the document and its meta-data to the Registered recipient.
  3. If the corresponding customer organisation has enabled receipt tracking for its sent documents, the System marks the document as received.
- **Alternative scenarios:** None
- **Remarks:** None

2.6 *UC6*: Register to personal document store

- **Summary:** An Unregistered recipient registers himself or herself in order to have a personal document store.
- **Primary actor**: Unregistered recipient
- **Secondary actors:** None
- **Interested parties:**
  - eDocs: wants as many Unregistered recipients to register as possible.
  - Unregistered recipients: want to register in the System for the improved functionality and ease of use.
  - Customer organisations: want a reliable way to deliver their documents.
- **Preconditions:**
- **Postconditions:**
  - The Unregistered recipient is now a Registered recipient and can from now on log in to the system to (amongst others) get an overview of all received documents.
  - The System has added all documents sent to the e-mail address of the Unregistered Recipient to his or her personal document store.
  - The Unregistered Recipient is logged in and can immediately use the System as Registered recipient.
- **Main scenario:**
  1. The Unregistered Recipient indicates that he or she wants to register himself or herself.
  2. The Unregistered Recipient specifies his or her details, i.e. his or her first name, last name, e-mail address, postal address.

3. The System verifies whether the Unregistered Recipient has given all necessary details.
4. The System verifies whether no user already exists with the given e-mail address.
5. The System adds all documents sent to the given e-mail address to the personal document store of the Unregistered Recipient.
6. The System indicates success to the Unregistered Recipient and logs him or her in.

- **Alternative scenarios:**

3a. If the Unregistered recipient has not provided all necessary details, the System indicates the missing details to the Unregistered recipient. Continue from step 2.

4a. If another user exists with the given e-mail address, the System indicates this to the Unregistered recipient. Continue from step 2.

- **Remarks:**

  - For brevity, we ignore the verification of the e-mail address of the Unregistered recipient in this assignment.

3 Non-functional requirements

Sections 3.1, 3.2 and 3.3 present an initial quality attribute scenario for respectively Availability, Performance and Modifiability.

Assignment. Document one additional quality attribute scenario for each of these qualities (Availability, Performance and Modifiability) in high level of detail. Make sure that these quality attribute scenarios are sufficiently different from and thus complementary to the initial scenarios presented in Sections 3.1, 3.2 and 3.3. Also make sure you pick sufficiently relevant scenarios.

3.1 Availability: *Personal document store failure*

The internal (sub-)system responsible for storing the documents in personal documents stores fails or crashes.

- **Source:** Internal
- **Stimulus:** The internal (sub-)system responsible for storing generated documents fails or crashes.
- **Artifact:** Internal subsystem
- **Environment:** Normal execution
- **Response:**

  - This does not affect the availability of other types of persistent data, such as (i) the state of ongoing document processing jobs, (ii) generated documents delivered through other channels, (iii) user data, (iv) billing data, etc.
  - This does not lead to loss of documents.
  - Prevention:
    - The storage system should have a guaranteed minimal up-time.
  - Detection:
    - The eDocs Operators are notified of this problem.
    - The System is able to detect this problem and goes into degraded modus:
      - Generated documents to be delivered via the personal document store are temporarily stored elsewhere and are processed when the storage system returns operational.
      - Fail gracefully: the users of the web interface are presented a maximally functional user interface with a clear message that the personal document stores are temporarily unavailable.
  - Resolution:
    - The notified eDocs Operators address the problem, e.g. by replacing failed hardware or by restarting software and, if necessary, reverting the storage system to a previous consistent state.
    - The System automatically brings the repaired storage system up to date by adding the documents that were generated while unavailable.

- **Response measure:**
  - Prevention:
    * Documents that the recipient has not received yet should be available at least 99.9% of the time (measured per month).
    * Documents that have already been received and were generated less than 30 days ago should be available at least 99.5% of the time (measured per month).
    * Documents that have already been received and were generated more than 30 days ago should be available at least 99% of the time (measured per month).
  - Detection:
    * It should be possible to temporarily store at least 3 hours of documents to be delivered via the personal document store.
    * Detection of failed hardware or crashed software happens within 5 seconds.
    * The eDocs operators are notified within 1 minute.
    * In the transition between normal and degraded modus, no documents are lost.
  - Resolution:
    * It should be possible for the eDocs Operators to address the problem within 2.5 hours.

3.2 Performance: *Document look ups*

It should be possible to look up documents via the personal document store in a timely fashion, even in case of a large number of parallel lookups, searches or downloads.

- **Source:** Registered User or Unregistered User
- **Stimulus:**
  - A Registered User consults his or her personal document store.
  - A Registered User consults a specific document in his or her personal document store.
  - A Registered User requests to download the PDF of a certain document in his or her personal document store.
  - A Registered User executes a search query on the documents in his or her personal document store.
  - An Unregistered User requests to download the PDF of a certain document to which he or she received the unique link.
- **Artifact:** The (sub-)system(s) responsible for handling requests of end-users regarding document look ups and downloads.
- **Environment:** Normal modus
- **Response:**
  - The System is able to respond to document look ups and downloads in a timely fashion while the rate of arriving requests is lower than a certain value.
  - If the System receives requests for document look ups and downloads at a higher rate than this value, the System throttles the excessive requests so that this amount of documents continues to be handled in a timely fashion.
  - A large number of document look ups and downloads should not affect the performance of other functionality of the System, such as (i) the personal document store of registered users, (ii) the status overview for Customer Administrators, (iii) delivering the raw data for new document processing jobs, etc.
- **Response measure:**
  - While the rate of arriving requests for document look ups and downloads is lower than 200 requests / second, the System behaves as follows:
    * Of the requests for consulting the personal document store, the System should be able to handle at least 95% within 300ms and 99% within 600ms (measured server-side).
    * Of the requests for searching the personal document store, the System should be able to handle at least 95% within 10s and 99% within 30s (measured server-side).
    * Of the requests for a specific document, the System should be able to handle at least 95% within 500ms and 99% within 1s (measured server-side).

* Of the requests for the PDF of a document generated less than 30 days before, the System should be able to handle at least 95% within 1s and 99% within 2s (measured server-side).
* Of the requests for the PDF of a document generated more than 30 days before, the System should be able to handle at least 95% within 2sec and 99% within 5s (measured server-side).

3.3 Modifiability: *Multiple print & postal services*

eDocs currently works with a single print & postal service to print and deliver its documents by postal mail. As a result, eDocs is highly dependent on the price setting of this one business partner. And because of the amounts of documents processed by eDocs every month, this price setting has a large influence on the profits of eDocs. Therefore, in order to alleviate this situation, eDocs wants to be able to easily switch print & postal service in the future.

- **Source:** The finance department
- **Stimulus:** Wishes to be able to easily switch print & postal service.
- **Artifact:** This modification affects the final steps of the document processing flow, i.e. forwarding the generated digital documents to a print & postal service.
- **Environment:** At design time or at run-time.
- **Response:**
  - Incorporating a new print & postal service in the System does not require changes to any other part of the System.
- **Response measure:**
  - Incorporating a new print & postal service in the System takes less than 1 man month to implement, test and deploy.
  - Switching to another print & postal service can be done without stopping or delaying on-going or future document processing jobs.

4 Format of the report

Integrate the documents and models that are generated in the analysis phase in a coherent document. Adhere to the following structure:

1. Domain Analysis
   (a) **Conceptual model(s) of the problem domain** (UML[2]),
   (b) Relevant additional **domain constraints** (informal text suffices, but OCL is allowed),
   (c) A **glossary**, which presents an overview of the different concepts and terms, together with a short description (only for those concepts that need further clarification)
2. Functional requirements
   (a) Use case overview diagram (UML[3]),
   (b) Detailed textual use case scenarios (roughly around 10 new high-detail textual use cases).
3. Non-functional requirements
   - In total three new quality attribute scenarios[4], one for Availability, one for Performance and one for Modifiability.

[2]For more information, please refer to the Larman book "Applying UML and Patterns, 3rd ed.", chapter 9 (Domain Models).
[3]For more information about modeling use case diagrams, refer to the Larman book "Applying UML and Patterns, 3rd ed.", Chapter 6 (Use Cases).
[4]For more information about the format in which quality attribute scenarios are documented, refer to the book of Bass, Clements and Kazman "Software Architecture in Practice, 3rd ed.", Part 2 (Quality Attributes).

There is no need to include the initial domain model diagram, the initial six use cases (but do include them in the use case diagram), and the initial quality attribute scenarios that have been presented as part of the assignment.

Hand in the report in the dedicated project post boxes (main floor of the department of computer science, in the student printer room A00.03) and submit the digital version (including all sources) via Toledo. The deadline for this is **Wednesday 11th of March at noon**.

Good luck,

The SA team.