H. DOCUMENT ANALYSIS

Concept of document analysis

In document analysis an algorithm scans the content of (many) documents to extract data from them and to process it further.

For example we can use this technology for **classification tasks**:
Are contracts a financial lease or an operating lease according to IFRS 16? This can be a quite complicated decision that needs many pieces of information.

But we can illustrate the principle of document analysis in a much simpler case. Here we need to **extract only limited data items**. The process to use this data afterwards for accounting purposes is also fairly simple:

The XFashion PLC has extra capacity from warehouses it does not need any more. Therefore XFashion PLC is offering to 3rd parties that they can use storage capacity and pay for that service.

XFashion PLC has many contracts about providing warehouse storage space. They are also available in electronic form as pdf files. The data in these contracts are relevant for accounting in two ways:

- First, each warehouse is a separate Cash Generating Unit (IAS 36.65 ff.). Cash inflows per building might be so low that an impairment is necessary.

- Second, XFashion's customers paid a security deposit to XFashion. The sum of all payments per customer should add up to the liability in the books.

Because there are many contracts, the content of these contracts shall be extracted automatically using text mining software.

Case Study Rental Agreements

There is already some software code in Rapidminer.

It reads contract files that a stored in a certain directory location. For each file (i.e. contract) the code tries to find data items that we are interested in. The data items are described with search patterns using "regular expressions". The code already includes search patterns for the contract terms WAREHOUSE, SIZE, ENDDATE.

- Define new search patterns using "regular expressions" to extract additional contract data for LEASEHOLDER, PAYMENT, DEPOSIT.

- Export the results as an Excel file (where to store?).

- Analyze the data and discuss results:

- Compare over all contracts the **sum of the deposits per lessor** with the liabilities per lessor shown in the books. You can use a pivot table in Excel to add numbers.

For comparisons use the data in **Check of liabilities for deposits (Task).xlsx**

- Use the **sum of the rents per building** (warehouse) to compare them with the cash flow projections for the building. You can use a pivot table in Excel to add numbers.
  To calculate the "value in use" (IAS 36.30 ff.) of cash generating units assume that XFashion's warehouses have no operating costs (i.e. cash flow = rental payments), current rental cash flow for the year is expected to continue as a perpetuity, but it will grow at 2% p.a. The discount rate is 6%. Assume for simplification also that all payments will made at the end of each calendar year

For comparisons use the data in **Check of value in use for warehouses (Task).xlsx**