# For presentation – about the dataset and the experiments with QA

The CASS dataset - ARRÊTS PUBLIÉS DE LA COUR DE CASSATION :

* Contains multiple type of cases but we will focus just on civile (?)
* published in the Civil Chambers Bulletin since 1960,

Some experiments:

* Preprocessed a part of the **civile type** dataset 🡪 generated the .story files which contains as a first part the **decision of the french court of cassation** and in the second part the **summaries of these decisions made by a lawyer**
* For now we have 6056 files (there are more but it takes a lot of time and being just an experiment for now the number is enough)

Example of story file: the 2 parts are separated by **“@highlight”**

Text

Description automatically generated

For generating synthetic question-answer pairs we used the model **T5 Question Generation and Question Answering** which isa T5 Transformers model fine-tuned in French on 3 tasks: question generation, question answering and answer extraction.

The datasets used for fine tuning are FQuAD and PIAF.

FQuAD - French Question Answering Dataset; contains 25,000+ question and answer pairs

# PIAF - French-language dataset of Questions-Answers

Evaluation results of the model:

Graphical user interface, application

Description automatically generated

Graphical user interface, text, application, email

Description automatically generated

Important to mention ! – we just used it , we didn’t do any training or fine-tuning , nut I think its important to show its results

For QA generation we will use as context text the summarization made by the lawyer because the summarization is more intelligible by people who are not expert in this domain. Also, our model will be used by normal people so their questions will use a normal language and not a legal language.

Ex1 : “ les traites secrets ou contre-lettres, en matière de vente des offices, sont frappes d'une ite d'ordre public “

QA generated :  
Diagram

Description automatically generated with medium confidence

Ex2 : “une décision constitutive de divorce se trouve exécutée quant à ses effets d'état par la rupture du lien conjugal réalisée, aussitôt qu'elle est passée en force de chose jugée et indépendamment de toute transcription”

A picture containing diagram

Description automatically generated