

DELIVERABLE

Project Acronym: **EUCases**

Grant Agreement number: **611760**

Project Title: **European and National Legislation and Case Law
Linked in Open Data Stack**

D2.5 Database schema

Revision: Final

Authors:

Alexander Konstantinov APIS

Project co-funded by the European Commission within FP7-ICT-2013-SME-DCA		
Dissemination Level		
PU	Public	
PP	Restricted to other programme participants (including the Commission Services)	
RE	Restricted to a group specified by the consortium (including the Commission Services)	
CO	Confidential, only for members of the consortium (including the Commission Services)	

List of Abbreviations

DB	Database
EU	European Union
NLP	Natural Language Processing
TSQL	Transact Structured Query Language
SOAP	Simple Object Access Protocol
XML	Extensible Markup Language

Revision History

Rev.	Date	Author	Partner	Description
0.1	12.06.2014	Alexander Konstantinov	APIS	First draft
0.2	20.06.2014	Tenyo Tyankov	APIS	Revision
1.0	30.06.2014	Alexander Konstantinov	APIS	Final draft

Status, abstract, keywords, statement of originality

Date of delivery	Contractual:	30.06.2014	Actual:	30.06.2014
Status	final			

Abstract (for dissemination)	<p>The Front-end Database (Front-end DB) is used by the Front-end Web Service which allows EUCases' applications (such as the ConsumerCases web application and the EULinksChecker add-in tools) and third party applications to query the database and to extract information based on various search criteria. Its main purpose is to supply the Web service with data and to ensure fast response to user requests. That is why all the potentially queryable elements are extracted from the input XML files in Akoma Ntoso format and populated in relational structure. The development of the database schema aims at ensuring building, storing, maintaining and tuning of the Linking Platform Front-end DB so that it can serve the online requests received by the Front-end Web Service. From its content coverage are excluded the language resources and all other data which is not intended to be used by end-user applications. The Front-end DB is optimised for speed data extraction according to the web application needs.</p>
Keywords	Relational database; database requisites, indexes, views and stored procedures

Statement of originality

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

Executive summary

The present deliverable describes in details the database schema of the Front-end database. It includes the structure of the Front-end DB: tables, fields, relationships, views, constraints, indexes, stored procedures, functions and other substantial DB elements. The database's main input is deriving from the Back-end DB where the data is stored in Akoma Ntoso XML format. When transferred to Front-end DB the XML files are processed and their main requisites are populated in the DB structure. This is done in order to ensure the querying process and fast responding to user queries sent by the Front-end web service.

Table of Contents

- 1 Role of the Front-end DB and communication with other databases and modules ... 6**
- 2 Front-end Database Input Data 8**
- 3 Database structure..... 9**
- 4 Relationship diagram 17**
- 5 Database functionality 18**
- Appendix: Database Schema 20**

1 Role of the Front-end DB and communication with other databases and modules

The Front-end Database (Front-end DB) supplies clients' end-user applications with well-structured and semantically enriched data deriving from the EUCases Linking Platform back-end. Its main purpose is to give fast responses to user requests. To this end all potentially queryable elements of the legal documents transferred in Akoma Ntoso XML format from the Back-end DB are extracted and populated in relational structure. The Front-end DB is built on MS SQL 2012.

Within the project framework two end-user applications will make use of the value added data ensured by the Front-end DB – the ConsumerCases web application and the EULinksChecker add-in tools. These will query the database and extract information via the services provided by the Front-end Web Service. The same approach could be used for supporting search queries of third party applications.

In the following paragraphs we briefly describe the communication of the Front-end DB with other databases and modules of the EUCases Linking Platform.

Back-end DB ⇔ Front-end DB

The conversion of data between the Back-end and the Front-end DBs' structures are done mainly using TSQL operations. The two SQL servers are as "linked servers".

Front-end DB ⇐ Front-end Web Service

The Front-end DB is queried by the Front-end Web Service via stored procedures. According to the specific service needs different kinds of stored procedures are used. Here are the main groups:

- Retrieval of a specific document and its metadata (e.g. when service needs to pass a specific document for a full-text display)
- Retrieval of a dataset of documents and their metadata covering specific criteria (e.g. by full-text search)
- Retrieval of data concerning user context (e.g. login)
- Storage of data concerning user context (e.g. history of documents opened).

Front-end DB ⇔ Front-end Web Service

The Front-end DB returns data to the Front-end Web Service after a stored procedure is executed by the latter. Data depends on the specific procedure executed and its parameters passed.

Front-end DB ⇔ Multilingual Access Module

Multilingual Access Module receives a user query as a string and information in which language the query is written and the languages in which the query needs to be translated or conceptual information. Multilingual Access Module is calling either the NLP Toolkit for ontology-based semantic annotation or the Machine Translation Module, depending on the settings of the system.

Front-end DB ⇐ Multilingual Access Module

Multilingual Access Module is producing new versions of the user query in two formats: ontology-based query in terms of classes and/or textual strings in the languages in which the query has to be translated.

Front-end Web Service ⇐ EULinksChecker

The EULinksChecker is an add-in tool. Its purpose is to give access to the Front-end Web Service functionality to the users of the main popular browsers, text editors and viewers. Here is a list of all functions executed by the Front-end Web Service for the needs of the EULinksChecker add-in:

- Add-in tries to logon to the service using specific credentials
- Add-in sends text to the service for citations and term recognition
- Add-in sends text to the service for conversion to legal XML
- Add-in requests a list of document citing a specific document divided into following groups:
 - EU Legislation
 - EU Case Law
 - National Legislation
 - National Case Law
 - All documents
- Add-in requests a list of document classified with a specific legal term divided into following groups:
 - EU Legislation
 - EU Case Law
 - National Legislation
 - National Case Law
 - All documents
- Add-in requests a cited legal text (e.g. provision) being part of a specific legal document
- Add-in requests a definition of a specific term.

Front-end Web Service ⇔ EULinksChecker

The Front-end Web Service serves the above listed requests of the EULinksChecker and returns the appropriate answers via classic web server response (SOAP).

Front-end Web Service ⇔ ConsumerCases DB

ConsumerCases DB main purpose is to serve the online requests which the ConsumerCases application receives. That is why this database will be optimised for speed data extraction according to the web application needs. The conversion of data between the Back-end and the Front-end DBs' structures will be done mainly using TSQL operations. The two SQL servers will be connected via the Front-end Web service using SOAP.

2 Front-end Database Input Data

The input to the Front-end Database is data stored in Akoma Ntoso XML format in the Back-end DB. After being downloaded legal documents are stored in the Back-end DB as raw data. Then several modules (or module groups) are processing the data. These are:

- Legal Text-to-XML Tools
- Linking Tools
- NLP Toolkit
- Tools for Classification and Summarization
- Knowledge Representation Module.

After all the processes are completed the record is marked as ready to enter the Front-end Database. That is how Back-end manager is notified that the record is ready to be stored. The XML document is extracted from the Back-end database and stored in the Front-end Database using a stored procedure **pDocumentSet**. This stored procedure analyses the structure of the passed XML and discovers its type. If it represents a court decision **pCourtActSet** is called, otherwise **pLegalActSet** is called.

Both of the above mentioned procedures do the following:

- Check if the document is already in the database
 - If yes – check if the document is a newer version of the old one
 - If yes document is stored as new version, and the old version is linked to the new
 - If no the existing document is replaced
 - If no – the document is stored in the database.

Both of the procedures **pCourtActSet** and **pLegalActSet** use the built-in XML functions of MS SQL Server 2012 for querying and extracting XML data. After that data is populated in the further explained structure.

3 Database structure

The following table describes in details the structure of the Front-end Database.

Field	Description
Table Languages – contains the languages used in the various fields in the database	
LanguageId	Unique identifier of the entity
LanguageShortName	Short name of the language used in the Akoma Ntoso XML standard
LanguageName	Name of the language
Table Countries – contains the countries used in the various fields in the database	
CountryId	Unique identifier of the entity
CountryShortName	Short name of the country used in the Akoma Ntoso XML standard
CountryName	Name of the country
Table States – contains the names of the states (federal districts) of a specific country	
StateId	Unique identifier of the entity
CountryId	Identifier of the specific country – foreign key
StateName	Name of the state
Table Courts – contains the court names in a specific country	
CourtId	Unique identifier of the entity
CountryId	Identifier of the specific country – foreign key
StateId	Identifier of the specific state – foreign key, optional
CourtShortName	Short name of the court
CourtName	Name of the court
Table Advocates – contains names of advocates	
AdvocateId	Unique identifier of the entity
AdvocateName	Advocate name
Table Dictionary – contains unique group identifiers of a word or phrase translations	
DictionaryId	Unique identifier of the entity
Table DictionaryItems – contains the translations in different languages of a specific group	
DictionaryItemId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
LanguageId	Identifier of the specific language – foreign key
Item	Translation of the word or phrase
Table DecisionTypes – contains types of court decisions	
DecisionTypeId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key

Table ProcedureTypes – contains types of procedures	
ProcedureTypeId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table Sources – contains all the crawler sources	
SourceId	Unique identifier of the entity
SourceName	Name of the source
Url	Url of the source
Table TermsOfUsage – contains different terms of use	
TermsOfUsageId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table AccessRights – contains the different data access rights	
AccessRightId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table PartyTypes – contains the different party types in a court case	
PartyTypeId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table EuroVoc – contains Eurovoc terms	
EuroVocId	Unique identifier of the entity
EuroVocParentId	Unique identifier of the parent – foreign key
OriginalId	The unique id in eurlex site
DictionaryId	Identifier of the dictionary group – foreign key
Table SubjectMatters – contains different subject matter terms	
SubjectMatterId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
OriginalId	The unique id in eurlex site
Table KeyWords – contains different key words	
KeywordId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
OriginalId	The unique id in eurlex site
Table CLDigest – contains Register of EU case-law terms	
CLDigestId	Unique identifier of the entity
CLDigestParentId	Unique identifier of the parent – foreign key
Old	Flag showing that the item is from the old or the new version
DictionaryId	Identifier of the dictionary group – foreign key
Table EUCitations – contains citations to EU acts	
EUCitationId	Unique identifier of the entity
Href	Href of the specific source
CitationText	Text of citation

Table NationalCitations – contains citations to national acts	
NationalCitationId	Unique identifier of the entity
Href	Href of the specific source
CitationText	Text of citation
Table CourtActs – contains metadata and full text of court decisions	
CourtActId	Unique identifier of the entity
EUCasesURI	Unique identifier of entity defined in Akoma Ntoso XML
SourceURL	Url of the source
NationalIdentifier	Identifier from the national source
ECLIIdentifier	ECLI identifier
ReplacedByECLI	Replaced by ECLI Identifier
CountryId	Identifier of the country to which the court belongs – foreign key
StateId	Identifier of the state to which the court belongs – foreign key
LanguageId	Identifier of the language the document is translated in – foreign key
OriginalLanguageId	Identifier of the original language of the document – foreign key
CourtId	Identifier of the court – foreign key
CourtChamber	Name of the court chamber
CourtPanel	Name of the court panel
ReferringNationalCourt	Name of the referring national court
Judges	Names of the judges
JudgeRapporteur	Name of the judge rapporteur
DecisionDate	Date of the decision
LodgedDate	Date the case has been lodged
SourceChangeDate	Date of the amendment in the source
EuCasesChangeDate	Date of amendment in the EUCases database
DecisionTypeId	Type of the decision – foreign key
ProcedureTypeId	Type of the procedure – foreign key
SourceId	Identifier of the source – foreign key
TermsOfUseId	Identifier of the terms of use – foreign key
AccessRightId	Identifier of the access rights – foreign key
Title	Title of the decision
ShortTitle	Short title of the decision
LegalDoctrine	Legal doctrine text
Headnote	Head note text
Summary	Summary text

PressRelease	Press release text
PublicationReference	Reference to publication
PublicationDate	Date of publication
InternationalLaw	International law involved reference
AdditionalInformation	Additional information text
FullText	Full text of the decision
Table CourtAct2Advocates – contains relationship between court decision and advocates	
CourtAct2AdvocateId	Unique identifier of the entity
CourtActId	Unique identifier of the court law document – foreign key
AdvocateId	Unique identifier of advocate - foreign key
AdvocateType	Type of role
Table CourtAct2Parties – contains relationship between court decision and parties	
CourtAct2PartyId	Unique identifier of the entity
CourtActId	Unique identifier of the court law document – foreign key
PartyTypeId	Unique identifier of party type - foreign key
PartyName	Name of party
Table CourtAct2EuroVoc – contains relationship between court decision and Eurovoc terms	
CourtAct2EuroVocId	Unique identifier of the entity
CourtActId	Unique identifier of the court law document – foreign key
EurovocId	Unique identifier of eurovoc term - foreign key
Table CourtAct2SubjectMatters – contains relationship between court decision and subject matter terms	
CourtAct2SubjectMatterId	Unique identifier of the entity
CourtActId	Unique identifier of the court document – foreign key
SubjectMatterId	Unique identifier of subject matter term - foreign key
Table CourtAct2KeyWords – contains relationship between court decision and keywords	
CourtAct2KeyWordId	Unique identifier of the entity
CourtActId	Unique identifier of the court law document – foreign key
KeyWordId	Unique identifier of - foreign key of key word
Table CourtAct2CLDigest – contains relationship between court decision and items from the Register of EU Case Law	
CourtAct2CLDigestId	Unique identifier of the entity
CourtActId	Unique identifier of the court law document – foreign key
CLDigestId	Unique identifier of - foreign key
Table CourtAct2EUCitations – contains relationship between court decision and EU Citations	
CourtAct2EUCitationId	Unique identifier of the entity

CourtActId	Unique identifier of the court document – foreign key
EUCitationId	Unique identifier of citation - foreign key
Core	Flag if the citation is core for the current court decision
Table CourtAct2NationalCitations – contains relationship between court decision and National Citations	
CourtAct2NationalCitationId	Unique identifier of the entity
CourtActId	Unique identifier of the court document – foreign key
NationalCitationId	Unique identifier of national citation - foreign key
Core	Flag if the citation is core for the current court decision
Table CourtAct2CaseLawCitations – contains relationship between court decision and other court decisions	
CourtAct2CaseLawCitationId	Unique identifier of the entity
CourtActId	Unique identifier of the court document – foreign key
CitedCourtId	Unique identifier of cited court decision - foreign key
Table AliasTypes – contains types of aliases	
AliasTypeld	Unique identifier of the entity
AliastType	Name of alias type
Table Aliases– contains legal instrument aliases	
AliasId	Unique identifier of the entity
AliasTypeld	Type of alias – foreign key
Alias	Name of alias
Table LegalActStatuses – contains statuses of legal instruments	
LegalActStatusId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table DirectoryCodes – contains directory codes from the Register of EU legislation on EUR-Lex site	
DirectoryCodeId	Unique identifier of the entity
DirectoryCodeParentId	Unique identifier of the parent – foreign key
OriginalId	Original unique identifier from the EUR-Lex site
DictionaryId	Identifier of the dictionary group – foreign key
Table NationalClassificationTypes – contains national classification groups	
NationalClassificationTypeld	Unique identifier of the entity
TypeName	Name of the group (type)
Table NationalClassifications – contains national classification structures	
NationalClassificationId	Unique identifier of the entity
NationalClassificationParentId	Unique identifier of the parent – foreign key
NationalClassificationTypeld	Unique identifier of the classification group – foreign key
OriginalId	Original unique identifier from the source

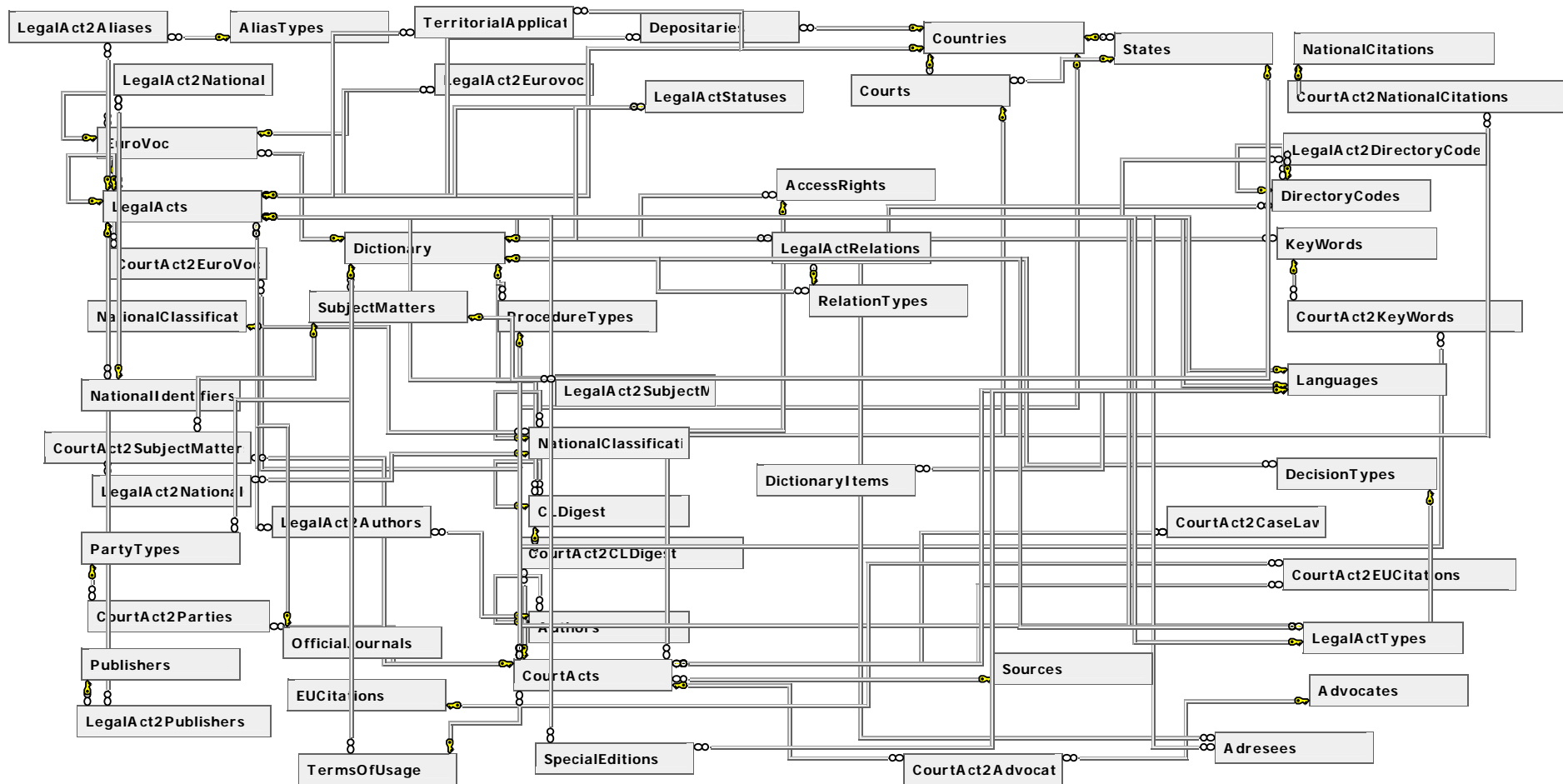
Item	The name of the item
Table OfficialJournals – contains names of official journals	
OfficialJournalId	Unique identifier of the entity
OfficialJournalName	Official Journal Name
OfficialJournalShortName	Official Journal Short Name
Table LegalActTypes – contains types of legal instruments	
LegalActTypeId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table LegalActs – contains the metadata and fulltext of legal instruments	
LegalActId	Unique identifier of the entity
VersionId	Identifier of the version
LeadLegalActId	Identifier of the last version – foreign key
EUCasesURI	Unique identifier of entity defined in Akoma Ntoso XML
SourceURI	Url of the source
FullTitle	Full title text
ShortTitle	Short title text
Alias	Alias text
OfficialJournalId	Identifier of official journal – foreign key
OfficialJournalNumber	Number of official journal
OfficialJournalSeries	Series of official journal
OfficialJournalPages	Pages from official journal where current act is published
OfficialJournalSpecialEditions	Special editions of official journal in which current act is published
ELIIdentifier	ELI Identifier
LanguageId	Identifier of the language the document is translated in – foreign key
OriginalLanguageId	Identifier of the original language of the document – foreign key
CountryId	Identifier of the country which published the current document – foreign key
DocumentDate	Date of the document
PublicationDate	Date of publication of the document
EffectDate	Date of entering into force of the document
EndOfValidityDate	End of validity date of the document
NotificationDate	Date of notification of the document
SignatureDate	Date of signing of the document
TranspositionDate	Date of transposition of the document
LegalActStatusId	Unique identifier of legal status – foreign key
Description	Description text

LegalActTypeId	Unique identifier of legal instrument type – foreign key
LegalActSubTypeId	Unique identifier of legal instrument subtype – foreign key
AdditionalInformation	Additional information text
ProcedureNumber	Number of the procedure
EuropeanParliamentLegislativeObservatory	European Parliament Legislative Observatory
IndexCM	Index CM
Table NationalIdentifiers – contains national identifiers of legal instruments	
NationalIdentifierId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
NationalIdentifier	National identifier value
Table TerritorialApplications – contains territorial application reference	
TerritorialApplicationId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
CountryId	Unique identifier of the country – foreign key
Table LegalAct2Eurovoc – contains relationship between legal instruments and Eurovoc terms	
LegalAct2EurovocId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
EuroVocId	Unique identifier of the Eurovoc term – foreign key
Table LegalAct2NationalClassifications – contains relationship between legal instruments and National Classificators	
LegalAct2EurovocId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
NationalClassificationId	Unique identifier of the National Classificator – foreign key
Table LegalAct2Abbreviations– contains relationship between legal instruments and abbreviations	
LegalAct2EurovocId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
AbbreviationId	Unique identifier of the abbreviation – foreign key
Table LegalAct2SubjectMatters – contains relationship between legal instruments and subject matter terms	
LegalAct2SubjectMatterId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
SubjectMatterId	Unique identifier of the subject matter – foreign key
Table LegalAct2DirectoryCodes – contains relationship between legal instruments and directory codes	
LegalAct2DirectoryCodeId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
DirectoryCodeId	Unique identifier of the directory code – foreign key
Table Authors – contains names of the authors of the legal instruments	

AuthorId	Unique identifier of the entity
AuthorParentId	Unique identifier of the parent – foreign key
AuthorName	Author name
Table LegalAct2Authors – contains relationship between legal instruments and authors	
LegalAct2AuthorId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
AuthorId	Unique identifier of author – foreign key
IsSub	Flag showing if author has subauthority role
Table Publishers – contains names of publishers of legal instruments	
PublisherId	Unique identifier of the entity
PublisherName	Publisher title
Table LegalAct2Publishers – contains relationship between legal instruments and Publishers	
LegalAct2PublisherId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
PublisherId	Unique identifier of the publisher – foreign key
Table Depositories – contains relationship between legal instruments and country – depository	
DepositoryId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
CountryId	Unique identifier of the country depository – foreign key
Table Addressees – contains relationship between legal instruments and addressees	
AddresseeId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
DictionaryId	Identifier of the dictionary group – foreign key
Table RelationTypes – contains relation types	
RelationTypeId	Unique identifier of the entity
DictionaryId	Identifier of the dictionary group – foreign key
Table LegalActRelations – contains legal instrument relations	
LegalActRelationId	Unique identifier of the entity
LegalActId1	Unique identifier of the legal instrument referring – foreign key
Href	Href of the legal instrument referred
RelationTypeId	Unique identifier of the relation type – foreign key
Table SpecialEditions – contains special editions of OJ	
SpecialEditionId	Unique identifier of the entity
LegalActId	Unique identifier of the legal instrument – foreign key
LanguageActId	Unique identifier of the legal instrument – foreign key
SpecialEdition	Special edition description

4 Relationship diagram

The following diagram displays the relations between the tables defined. Each relation represents a foreign key definition which is accompanied by index.



5 Database functionality

The following list describes the functionality in the Front-end database which is developed in TSQL.

1. Function *f_ak_split_string*

```
function f_ak_split_string (@str nvarchar (max), @delim nvarchar (max))
returns
  @SlittedStrings table
(
  id int identity (1,1) primary key,
  line nvarchar (max)
)
```

Description	Splits string by given delimiter
Parameters	@delim - delimiter @str - string to be splitted
Return value	Returns a table of splitted substrings ordered by id

2. Procedure *p_ak_drop_all_foreign_keys*

```
PROCEDURE p_ak_drop_all_foreign_keys @tableName nvarchar (max)
```

Description	Drops all foreign keys of a specific table
Parameters	@tableName the name of the table

3. Procedure *p_ak_create_fk_indeces*

```
PROCEDURE p_ak_create_fk_indeces @tableName nvarchar (max)
```

Description	Creates indexes on all foreign keys of a specific table
Parameters	@tableName the name of the table

4. Function *f_ak_trim*

```
function f_ak_trim (@str nvarchar (max))
returns nvarchar(max)
```

Description	Trims the given string from all white spaces
Parameters	@str - given string
Return value	Trimmed string

5. Procedure *pDocumentSet*

```
procedure pDocumentSet
  @atXML XML
```

Description	Writes in database all the data of
--------------------	------------------------------------

	a given AKOMA NTOSO XML, deciding if it is a new or ammended document or a new version of a document, and if it is a legal instrument or a court decision
Parameters	@atXML AKOMA NTOSO XML

6. Procedure *pCourtActSet*

`procedure pCourtActSet`
 @atXML [XML](#)

Description	Writes in database all the data of a given AKOMA NTOSO XML for court decision, deciding if it is a new or ammended document or a new version of a document
Parameters	@atXML AKOMA NTOSO XML

7. Procedure *pLegalActSet*

`procedure pLegalActSet`
 @atXML [XML](#)

Description	Writes in database all the data of a given legal instrument AKOMA NTOSO XML, deciding if it is a new or ammended document or a new version of a document
Parameters	@atXML AKOMA NTOSO XML

Appendix: Database Schema

The database schema of the EUCases Front-end DB is provided with the following sql-files:

- 1. Create Frontend DB.sql
- 2. Tools Frontend DB.sql
- 3. Tables Frontend DB.sql
- 4. Indexes Frontend DB.sql
- 5. Stored Procedures Frontend DB.sql

and is published on EUCases web site as zip archive:

http://eucases.eu/fileadmin/EUCases/documents/Database_Schema_Frontend_DB.zip