

Identifiers, Metadata and Document Structures: Essential Ingredients for Inter-European Case Law Search

European Legal Access Conference, Paris, 10-12 december 2008

Marc van Opijnen
sr. adviser legal informatics
Council for the Judiciary, the Netherlands
mvanopijnen@rechtspraak.nl

Abstract:

After establishing the need for cross-border dissemination of national (EC) law, this paper reviews already existing initiatives like JuriFast, Dec.Nat, Caselex and Jure.

For future development three basic challenges are defined and elaborated.

The lack of a unique and persistent case law identifier gives rise to many problems when citing and searching case law. The paper describes the way this issue is tackled in the Netherlands, by using a national identifier together with a publicly available reference-index. Inspired by the advantages of this system, a proposal is made for a European Case Law Identifier.

To facilitate the inter-European search for, and exchange of case law the paper also discusses the need for a harmonized set of metadata, and the necessity to develop a European interchange format for court decisions.

1 Introduction

Knowledge of Community law is of the utmost importance for the functioning of the European legal order and the common market. This applies equally for all players in the legal arena: citizens, businesses, lawyers, prosecutors and judges. The specific role of the judiciary was recently stressed by the resolution of the European Parliament *on the role of the national judge in the European judicial system* of 9 July 2008.¹

Because of the principles of the primacy and the direct effect of Community law, not only the jurisprudence of the European Court of Justice is an important source, but also the case law of national courts. E.g.: the national judge has to be aware of issues raised by other Member States' courts in the preliminary reference procedure (under Article 234 EC Treaty) in order to prevent raising the same issue twice. But knowledge of case law of other Member States is also important to decide whether a specific question is an '*acte clair*'.² One of many

¹ <http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P6-TA-2008-0352+0+DOC+XML+V0//EN>

² The Acte clair theory means that a legal question does not have to be submitted to the European Court of Justice if over the interpretation or validity of Community law reasonably no doubts can exist.

examples can be found in the Dutch case LJN BD2447³ where the Advocate-general argued, quoting decisions of the French *Cour de Cassation*, the British House of Lords and the German *Oberlandesgerichte Stuttgart* and *Köln* that no preliminary proceedings were necessary for the interpretation of a specific question related to the European Arrest Warrant.

Although the European Parliament correctly noted “(T)hat complete and up-to-date information on Community law is not available in a systematic and proper manner to many national judges (...)”⁴, the availability of this information can be considered to be equally important to citizens and other legal professionals.

Apart from the databases mentioned by the European Parliament⁵:

- a) “National databases on national court rulings concerning Community law”;
- b) The database on “Conventions and Regulation on jurisdiction and enforcement of judgments in civil and commercial matters” (*infra*, section 2.7);
- c) “Databases containing pending references for preliminary rulings from all Member States (... and...) judgments of referring courts applying a preliminary ruling”

several projects show that legal practitioners also have a great interest in judgments from other Member States regarding Community law, other than those on preliminary proceedings or enforcement cases.

Not only do these initiatives not seem to be known by legal practitioners⁶, they also reveal some essential issues which should have to be resolved in the short term in order to enable a well-featured inter-European access to national case law in the long run.

These issues can be summarized as follows:

- 1) **Identification**
how can one cite a case in such a way that both a colleague from abroad, and an automated system can find the same case easily in a completely different database;
- 2) **Metadata**
how can metadata be added to a judgment in such a way that cross-border search is facilitated;
- 3) **Document structure**
Although judgments are comparably structured in various jurisdictions, these structures are not machine-readable. Because of this enhanced search and display features are seriously hampered.

Apart from these three ‘must haves’ there are also some other problems to be tackled. They are also of the utmost importance, but they might be regarded as ‘should haves’.

- 4) **Multilinguality**
Most judgments can not be read or fully understood by lawyers or citizens from other Member States. Therefore, translations or automated translation services should be available;
- 5) **Techniques to cope with overabundance**
The more case law there is available, the harder it will be for the user to find what he is looking for. With taxonomies, rating systems and user input the overabundance might be tackled.

³ <http://www.ljn.nl/BD2447>

⁴ l.c. nr .8

⁵ l.c. nr 10-11

⁶ Report on the role of the national judge in the European judicial system (2007/2027(INI)), Annex: questionnaire sent to national judges, 7-c-iv.
<http://www.europarl.europa.eu/sides/getDoc.do?type=REPORT&reference=A6-2008-0224&language=EN&mode=XML#title2>

This paper focusses on identifiers, metadata and document structure.

First though, section 2 gives an overview of existing initiatives for cross-border access of national (EU) case law. Section 3 zooms in on the necessity for a unified approach on case law identifiers and a reference index. The system as developed in the Netherlands will serve as an illustration for future work. Section 4 discusses the urge for some agreement on metadata and section 5 focusses on the need for an interchange format for judicial decisions. The paper ends with a conclusion (section 6) and an overview of cited literature (referenced in the text in an abbreviated form).

2 Current Websites

2.1 Overview

In this section some websites are discussed which offer cross-border search facilities for national case law, most of them with a focus on Community law. Figure 1 shows the number of cases per country in the reviewed databases. Figure 2 contains an overview of metadata and search criteria. In the sections 2.3 – 2.9 these websites are briefly described.

The metasearch website of the Network of the presidents of the supremes courts of the EU is described in section 2.2, but – due its very special character – not included in the tables hereunder.

Country	Dec.Nat	JuriFast	Caselex	FRA	JURE	IARLJ	Codices
Austria	632	44	118	7	188	0	109
Belgium	1692	42	0	8	81	3	211
Bulgaria	0	0	28	8	0	0	47
Cyprus	1	1	1	6	0	0	30
Czech republic	42	29	25	5	0	0	199
Denmark	254	1	10	8	0	0	38
Estonia	8	13	4	3	0	0	74
Finland	147	16	9	18	8	0	9
France	3108	29	300	6	251	8	177
Germany	4567	79	382	11	550	99	278
Greece	775	20	107	6	92	0	38
Hungary	9	3	7	8	0	0	154
Ireland	130	3	33	11	12	0	31
Italy	2509	9	245	8	187	0	146
Latvia	0	7	4	1	0	0	68
Lithuania	0	6	23	0	0	0	174
Luxembourg	259	2	0	3	0	0	7
Malta	0	2	0	3	0	0	19
Netherlands	1884	54	174	5	51	6	126
Poland	24	8	28	1	0	1	342
Portugal	282	55	18	5	9	0	135
Romania	1	0	52	7	0	0	85
Slovakia	19	7	0	3	0	0	80
Slovenia	5	0	0	6	0	0	172
Spain	1387	41	104	7	11	0	220
Sweden	139	11	59	8	39	0	21
UK	1125	22	312 ^{b)}	9	123	62	30
EU CoJ	0	0 ^{a)}	172	0	130	0	207
Iceland	0	0	6	0	0	0	1
Liechtenstein	13	0	0	0	0	0	39
Norway	4	0	10	0	0	0	78
Switzerland	19	0	0	0	0	30	182
EFTA Court	0	0	13	0	0	0	0
Other European countries	0	0	0	0	0	0	1356
Other non-European countries	0	0	0	0	0	21	762
Total	19035	504	2244	171	1732	230	5645

a) Decisions in preliminary proceedings of the European Court of Justice are registered at the Member State.

b) Totalized. England/Wales 252, Northern-Ireland 13, Scotland 25, UK 22.

* Counts done on 15-11-2008, except for Codices, which was counted in aug. 2008.

Figure 1: number of cases in the reviewed databases

Field	Dec.Nat		JuriFast		Caselex		FRA		JURE		IARLJ		Codices	
	Srch ^{a)}	Dspl ^{b)}	Srch	Dspl	Srch	Dspl	Srch	Dspl	Srch	Dspl	Srch	Dspl	Srch	Dspl
Country	+	+	+	+	+	+	+	+	+	+	+	+	+	+
Court	+	+	-	+	+	+	-	+	+	+	+	+	-	+
Chamber	-	+	-	-	-	+	-	+	-	-	-	-	-	+
Date	+	+	+	+	+	+	+	+	+ ^{c)}	+ ^{c)}	+	+	+	+
Parties	+	+	-	+	+	+	-	-	+	+	-	-	-	+
Party-role	-	-	+	-	+	+	-	-	-	-	-	-	-	-
Decision type	+	+	-	+	+	+	-	-	-	-	-	-	-	-
Irrevocability	-	-	-	-	-	-	-	-	-	-	-	+	-	-
Database ID	-	-	-	+	+	+	-	+	+	+	-	+	+	+
National ID	+	+	-	+	+	-	-	+	+	+	+	+	+	+
Common name	+	+	-	-	-	+	-	-	-	-	-	-	+	+
Language	+	+	-	-	-	+	-	+	-	+	-	-	-	+
Headnotes	-	+	+	+	+	+	-	+	-	-	-	-	-	+
Summary	-	+	-	+	+	+	-	+	+	+	+	+	+	+
Thesaurus	+	+	-	-	+	+	+	+	+	+	+	-	+	+
EU-provision	+	+	+	+	+	+	+	+	+	+	+	-	-	-
idem, Celex nr	+	+	-	+	+	-	-	-	-	-	-	-	-	-
Nat. provision	+	+	-	-	-	+	-	-	-	-	-	-	-	-
Internat. prov.	+	+	-	-	-	-	-	-	-	-	-	-	-	-
EU case law	+	+	-	+	-	+	-	-	-	-	-	-	-	-
nat. case law cited	-	-	-	-	-	+	-	-	-	-	-	-	-	+
Cited by nat. case law	-	-	-	-	-	+	-	-	-	-	-	-	-	-
Full text	-	-	-	+	+	+	+	+	+	+	-	+	+	+

- a) Search criteria
- b) Metadata displayed
- c) Only the year.

Figure 2: search criteria and metadata in the reviewed databases

2.2 Common Portal of National Case Law

The *Common Portal of National Case Law*⁷ of the Network of the Presidents of the Supreme Judicial Courts of the European Union offers a metasearch engine which permits simultaneous searches of almost all the case law databases of the Supreme Courts of the EU. Using the Eurovoc-thesaurus⁸ the user-query can be translated into the languages of each of the chosen databases. In the public access version the records found can only be accessed in the original language; users having a login code are also offered machine translations in various language pairs. No metadata search is possible, and there no search filters to limit the time range. The Common Portal uses just the identifiers supplied by the connected systems, which makes it hard to cite.

⁷ <http://www.reseau-presidents.eu/rpcsjué>

⁸ <http://europa.eu/eurovoc/>

2.3 Caselex

The commercial initiative Caselex⁹ gathers case law from EU and EFTA Member States which are deemed relevant for the interpretation and implementation of Community law. The search interface is very comprehensive; extensive metadata are supplied and decisions are summarized in English – with a clear division between the facts of the case, the legal issues and the final decision.

The judgments are gathered via a network of national contactpoints, but in some countries they are more active than in others (Germany is represented with 382 cases, while i.a. Belgium has none).

Caselex developed its own identifier for all cases in its database, e.g. CASELEX:NL:2008:10, the first part being the name of the identifier ('Caselex'), the second the EU-countrycode¹⁰, the third the year of the decision, and finally an ordinal number (related to country and year). National identifiers are also displayed and can be searched for.

2.4 Dec.Nat

The Dec.Nat-database¹¹ of the Association of the Councils of State and Supreme Administrative Jurisdictions of the European Union contains ± 19.000 national decisions regarding Community law. The references and the legal analyses of the decisions have been supplied by the Research and Documentation Service of the European Court of Justice. Search possibilities and metadata (mostly in French) are quite extensive. The usability varies though: despite superfluous metadata on relevant legislation the actual dictum is often not clear, summaries are lacking and there are no links to the original texts. Dec.Nat uses its own identifier (e.g. 'NL.20060127/03') – consisting of a country code, year of decision and an ordinal number – but this is merely a technical number, not explicitly presented to the user.

2.5 Jurifast

The Association of the Councils of State and Supreme Administrative Jurisdictions of the European Union is also responsible for *Jurifast* ('Fast information system for case law')¹², which contains preliminary questions submitted to the European Court of Justice, the Court's answers, and the subsequent national decision(s). Although *Jurifast* contains much less decisions than Dec.Nat – at the moment 504 – usability is considerably better. Just because the database contains many cases which haven't been decided by the Court yet, it can be effective in preventing the submission of redundant preliminary questions to the Court. *Jurifast* uses an identifier (e.g. 'UK02000031') comparable to the one used by Dec.Nat.

2.6 FRA

The EU Agency for Fundamental Rights (FRA) has a database¹³ of important national decisions of courts and special bodies related to issues of discrimination law. All cases are summarized, and have a rich set of metadata. Some decisions are also available in full text. The database itself though contains only 170 cases, the latest from 2007. FRA assigns its own (opaque) identifiers, e.g. '135/1'. National identifiers are displayed in the 'title' or in the 'subtitle', but can not be searched for.

⁹ www.caselex.com

¹⁰ <http://publications.europa.eu/code/pdf/370000en.htm> 'EU' added for the courts of the European Union and 'EF' for the EFTA-court.

¹¹ http://www.juradmin.eu/en/jurisprudence/jurisprudence_en.lasso

¹² http://www.juradmin.eu/en/jurisprudence/jurifast/jurifast_en.php

¹³ <http://raxen.fra.europa.eu/1/webmill.php>

2.7 JURE

The JURisdiction Recognition Enforcement database (JURE)¹⁴ was developed by the European Commission. It contains 1732 cases of the European Court of Justice and the Member States' courts on the interpretation of Council Regulation (EC) no 44/2001 of 22 December 2000, on Jurisdiction and the Recognition and Enforcement of judgments in civil and commercial matters and on the interpretation of the 1968 Brussels Convention and the 1988 Lugano Convention. Apart from the decisions in the original language, the database offers summaries in English, French, German and the original language. JURE assigns its own (opaque) identifiers (e.g. '38621'), and many metadata are added. It is also possible to search for national identifiers (case or docket numbers) and publication references.

A drawback of the JURE-database is its up-to-dateness: the most recent cases are from 2005. Also the distribution of the cases over the Member States is unequal: Germany has 550 cases, while many others have none.

2.8 IARLJ

The International Association of Refugee Law Judges established, in close cooperation with the *Europäische EDV-Akademie des Rechts*, the database Court Decisions in Asylum Law.¹⁵ It doesn't have the friendliest user interface around, and contains 230 cases. The database assigns its own (opaque) identifiers (e.g. '137'). National case numbers are displayed, but can not be used for searching. The IARLJ-database is the only database having a metadata-field for irrevocability.

The database does not seem to be maintained anymore.

2.9 CODICES

The European Commission for Democracy through Law (also known as Venice Commission) maintains an extensive database on constitutional case law: CODICES.¹⁶ The database is worldwide, but has a strong European focus (4883 European cases on a total of 5645). The metadata are superfluous, including an extensive thesaurus; the search interface is very rich. CODICES constructed its own identifier, e.g. 'FRA-2007-1-003', being the IOC country code, the year of the decision, an indicator for the printed copy of the Bulletin on Constitutional Case Law, and finally an ordinal number. In the result list also a national identifier is shown (in this example 2007-551 DC), in conjunction with the CODICES-identifier.

2.10 Other websites

Apart from the mentioned systems many others exist. They range from broad commercial aggregators like V|Lex¹⁷, to niche systems like Oxford Reports on International Law (formerly International Law in Domestic Courts)¹⁸ and the legal tools database of the International Criminal Court.¹⁹ Lacking a specific European focus, they are not discussed here.

¹⁴ <http://ec.europa.eu/civiljustice/jure/index.htm>

¹⁵ <http://iarlj.eear.eu/IARLJ/>

¹⁶ <http://www.codices.coe.int>

¹⁷ <http://vlex.com>

¹⁸ <http://ildc.oxfordlawreports.com/>

¹⁹ <http://www.icc-cpi.int/legaltools/>

2.11 Conclusions

The overview shows a wide range of different databases. Some of them are very general in nature, some of them focus on specific fields of law. The metadata and interfaces, although sometimes due for improvement, are designed with a view to the specific needs of the intended users. Although some websites might better be discontinued due to a lack of maintenance, it might not be advisable to replace the existing initiatives by one new database.

Still though, there is room for improvement and better cooperation. Nearly every system has developed its own identifier – the problems of which will be discussed in the next section. And although sometimes the same kind of metadata are used, it does not look like any site has tried to borrow already existing field labels from other sites. Generally accepted metadata standards like Dublin Core (*infra, section 4.4*) are not used on any of the discussed websites.

3 Identifiers

3.1 A broad Variety

Every country has its own tradition and – sometimes extensive, sometimes minimal – rules on citation. These rules and traditions mostly stem from the pre-internet era, and therefore cause problems in a rapidly changing legal information landscape.

A more extended description of these issues can be found at [*Van Opijnen 2008*] and literature cited there, but they may be summarized as follows:

- judicial decisions are often identified / cited by their ‘attributes’ (court, date, case number); which do not have a fixed notation and are therefore not suited in an electronic environment. Moreover, these attributes do not state anything on the exact location (law journal, database record, URL) where the judgment is to be found.
- Therefore, often a ‘publication reference’ is added – an identifier pointing to such a specific location. Such a reference (e.g. ‘NJ 2008/37’) though is mostly vendor-specific: if you do not have (fee-based) access to the journal or database, you will not find the judgment (easily), even if it’s also available in a source you do have access to.
- Also, the traditional identifier often is not medium neutral: it is built on a concept (pages) which not only falls into oblivion, but might also lead to ambiguous citations: when there are three judgments printed on one page, a paper based search will still lead quite easily to the right case, while in an electronic environment this will lead to insurmountable problems: the uniqueness of identifiers is a crucial pillar in the information architecture.

3.2 Solutions: Neutral Identifier and Index

A neutral identifier is necessary to enable citing case law in a unique, medium-neutral and vendor-neutral way. To stress the fact that the courts themselves are responsible for the numbering of their decision, it might also be called a ‘court designated identifier’ [*Mowbray 2000*].

The coming into existence of case law databases of the judiciaries themselves, has already given an impulse to the use of such court designated identifiers, but in many countries not all cases which are published in legal journals are published on the judiciary’s website.²⁰ Also, the court designated identifier is often viewed by the legal community as just another publication reference, instead of ‘the’ identifier.

²⁰ A study in the Netherlands [*VanOpijnen 2006a*] showed that in 2004 more than 50% of cases of district courts and courts of appeal which were published in legal journals were not published on the judiciary’s website.

The court designated identifier differs in form from country to country. In some countries there are completely opaque, in other countries they are composed of meaningful elements, like name of court (mostly abbreviated), year of decision and an ordinal number. In common law countries names of parties are mostly added.

To give an example: in Finland a case might be identified as ‘KKO:2008:97’. This is a meaningful identifier, holding an abbreviation (KKO) for the name of the court (*Korkein oikeus* – the Finnish Supreme Court), year of decision and an ordinal number.

In the Netherlands a case might be identified as ‘LJN AB1234’, where ‘LJN’ is the name of the identifier itself (infra, section 3.3). Other information can be added, but is not necessary for identification.

Still though, just adding a court designated identifier to a decision is not sufficient to tackle the problems of citation.

- 1) When you have the identifier, you still do not know where to find the case – or where to find summaries, annotations or translations.
- 2) If the issuing of the court designated identifier depends on the publication of the decision on the court’s website, a lot of cases might go without this identifier if not all cases are published.
- 3) The number might be regarded as just being reference number, and not really be used for citation.
- 4) A lot of older cases which were never published still don’t have the neutral identifier.

Probably the best solution attainable is to build an index of identifiers, as was done in the Netherlands.

3.3 Dutch Solution: an Index with Case Law Identifiers

In 2000 the Dutch judiciary decided to build *Porta Iuris*: an intranet portal with an integrated search interface for both internal and commercial case law databases. Firstly, because the judiciary wanted more control over the selection of, and access to their own case law, instead of being dependent on commercial publishers. And secondly, because the need was felt for one single integrated user interface for legal source material, replacing all different databases, websites, cd-roms and paper case law periodicals.

The architects of *Porta Iuris* faced a major challenge: because many decisions exist in more than one database, the user runs the risk of having a search result of e.g. 20 cases, from which only twelve are actually different – to find out though, he has to open all 20 of them. And also, when the user finds an interesting case, e.g. in a summarized version, he doesn’t know whether the full text is available in another database, or whether the case is annotated in some law review.

Therefore, it was decided to build an index of all decisions which were integrated in *Porta Iuris*. This index is built separately from the databases containing the cases themselves, and only contains:

- the unique case law identifier (hereafter: ‘LJN’, the abbreviation of ‘*Landelijke JurisprudentieNummer*’ = National case law number);
- the attributes of the decision (court, date, case number);
- publication references (the identifiers used by case law journals);
- existence (and author name) of annotations.

Using this index the user can search for any kind of number (attributes, LJN or publication reference), and have all instances of the case presented to him. From there he can view the

instance he prefers. When he is performing an advanced search, results are grouped by decision.

Figure 3 shows the results of a search for jurisprudence of the European Court of Justice: apart from the instance of the decision in EUR-Lex, the user can see that the case is also published in three Dutch periodicals, two of which also have an annotation.

Home Regelgeving Jurisprudentie Vaknieuws Lit. & Docu. Help

Zoeken op nummer Resultaatlijst Uitspraak

Jurisprudentie: resultaatlijst

Gezocht op: Bron: EURLEX 62004J0150;

Er is één uitspraak die voldoet aan uw zoekcriteria.

[Print deze pagina](#)

[aanpassen](#) [opnieuw](#)

1. LJN: BA1051, Hof van Justitie van de Europese Gemeenschappen, 30-01-2007, Zaak C-150/04.

EUR-Lex 62004J0150
Arrest van het Hof (grote kamer) van 30 januari 2007.
Commissie van de Europese Gemeenschappen tegen Koninkrijk Denemarken.
Niet-nakoming - Vrij verkeer van werknemers - Vrij verrichten van diensten - Vrij verkeer van kapitaal - Vrijheid van vestiging - Inkomstenbelasting - Ouderdom...
[\[Meer, Volledig\]](#)

NJ 2007, 283 m. nt. MRM
Beroep wegens niet-nakoming, ingesteld op 23 maart 2004.
Vrij verkeer van werknemers - Vrij verrichten van diensten - Vrij verkeer van kapitaal - Vrijheid van vestiging - Inkomstenbelasting - Ouderdomspensioen. Aansluiting bij in andere lidstaat gevestigde pensioeninstelling. Belastingwet...
[\[Meer, Volledig\]](#)

RvdW 2007, 273
Commissie van de EG tegen Denemarken
Beroep wegens niet-nakoming, ingesteld op 23 maart 2004.
Vrij verkeer van werknemers - Vrij verrichten van diensten - Vrij verkeer van kapitaal - Vrijheid van vestiging - Inkomstenbelasting - Ouderdomspensioen - Aansluiting bij in andere lidstaat g...
[\[Meer, Volledig\]](#)

VN 2007/14.12 m. nt. Red.
INKOMSTENBELASTING. EUROPEES RECHT
Zaak Commissie/Denemarken. Gemeenschapsrecht. Pensioen. Denemarken. Uitsluiting van fiscale faciliteiten bij pensioenopbouw op grond dat de pensioenregeling is ondergebracht bij een buitenlandse pensioeninstelling is in strijd met het gemeenschapsrecht
[\[Meer, Volledig\]](#)

Jur. 2007, I-01163

Figure 3 ECJ-case found in Porta Iuris by searching on Celex code. The user can access the case in EUR-Lex (as integrated within Porta Iuris), or in one of three Dutch law reviews, in two of which ('NJ' and 'VN') also a scholarly note can be found. Also the reference of the European Court Reports ('Jur. 2007, I-01163') is displayed.

At the moment the index contains 350.000 unique decisions, which have in total 460.000 instances.

They include:

- the case law database of the judiciary’s public website;
- 20 internal databases which can only be viewed by the court supplying the content;
- an internal database for the whole judiciary;
- the internal case law database of the Council of State;
- 40 databases of commercial publishers;
- all 19.000 judgments of the European Court of Justice (imported from EUR-Lex).

It should be stressed that the LJN and the index are not dependent upon the case law database on the judiciary’s website: only 130.000 cases out of the 350.000 are published on the internet. By now, nearly all law reviews are included in the index, including historic files (dating back to 1913). For technical reasons – extensively elucidated in [*Van Opijnen 2006b*] – some law reviews are still missing though.

The matching of the cases is partly automated, but still requires some manual labour. Daily updates are done by the content managers of Porta Iuris, but the initial work required assistance of quite some law students.

Because the index – developed with public means – was considered to be of interest for the whole legal community in the Netherlands, it was decided to make it freely available on the internet. And to comply with the Directive on the re-use of public sector information²¹, also a webservice was developed, enabling heavy users to query the index from within their own network or to maintain a synchronized copy of the index. While in Porta Iuris users can access the decisions themselves, in the public version of the index this is only possible for those databases which are accessible free of charge (the Dutch judiciary’s website and EUR-Lex).

As a result of the adoption of the LJN by the public-private standardization body ‘Juriconnect’²² [*Rikhof 2008*], the LJN is now also included in (paper and electronic) publications of commercial publishers.

3.4 The European Level

Citing case law on the European level can be as difficult as on the national level. This may be illustrated by Case C-299/02 from the European Court of Justice of 14-10-2004.

Apart from citing the case by these attributes, there are some other useful identifiers:

- European Court reports 2004, Page I-09761
- Celex-number: 62002J0299
- LJN:AR7319
- NJ 2005, 421
- CASELEX:EU:2004:8
- BAILII: [2004] EUECJ C-299/02
- V|Lex: VLEX-17284540

Other, national, examples can easily be gathered. E.g. in the Dec.Nat-database one can find the decision of Tribunal da Relação de Lisboa, 20-05-2004, 2320/04-8²³. Under ‘remarks’ one reads (figure 4): “*See also IH/04896-A*”. One can suppose – though not even be sure – that

²¹ Directive 2003/98/EC of the European Parliament and of the Council of 17 November 2003 on the re-use of public sector information,

<http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32003L0098:EN:NOT>

²² www.juriconnect.nl

²³ http://www.juradmin.eu/en/jurisprudence/jurisprudence_en.lasso?page=detail&doc=P..20040520/01

it's a judgment being referred to, but from which country? Which database to search? Google isn't of much help here ...

ANALYSIS

Matters (CELEX codes) : Judicial cooperation in civil matters - Miscellaneous

Matters (index marks) : EC - Judicial cooperation in civil matters

Keywords:
EEC - Free movement of persons; - Judicial cooperation in civil matters - Judicial jurisdiction and enforcement of decisions in civil and commercial matters - Council Regulation (EC) No. 44/2001 - Agreement conferring jurisdiction Formal conditions International maritime transport contract - Clause indicated on the back of the bill of lading Form permitted via use in the commercial sector in question

GENERAL COMMENTS

Remarks :
See also IH/04896-A

Figure 4 Mystery citation in the Dec.Nat-database.

Or, because Caselex uses its own identifier and (for Dutch decisions) the LJN, while JURE uses the attributes and the publication references, there is no number to match a Caselex-case to a JURE-case, especially not because JURE doesn't display the exact date of the judgment, only the year.

So, how to cite a national case for an international audience?

- The citing lawyer has to know whether the country he is citing a case from, has a neutral identifier at all;
- Then, he has to be able to recognize this neutral identifier between all possible numbers and codes displayed in the interface;
- Then, he has to somehow add a countrycode, to make clear to the reader of the citation where to look for the case;
- Anyway, one ends up with a citation that can not be easily processed automatically.

The solution might be to develop a European Case Law Identifier (ECLI). This ECLI is a standard method to construct an identifier from a national neutral identifier in such a way that it can be used by any lawyer or citizen from every Member State to cite or find a case in every system where it is stored.

While [*Van Opijnen 2008*] discusses the more theoretical framework, the requirements for the construction of the ECLI can be summarized as follows:

- **Neutrality**
It has to be medium- and vendor neutral, preferably court designated.
- **Identifiability**
Case law identifiers are used in everyday legal writings. Therefore, a legal practitioner has to be able to recognize the identifier as such.
- **Meaning**
Apart from the identifier one shouldn't need any other data to assess the relevance of the citation. So, country, court and year should be part of the identifier.

- **Error-proof**
 - Codes for courts are preferable to full names;
 - The use of characters which are not common to lawyers should be avoided and preferably not more than one type of non-alphanumeric character is used;
 - The shorter the better – information which is not necessary for identifying should not be included, but left to the metadata (e.g. the type of procedure);
- **Compliant with other standards**
Compliance with international standards or EU-practice (European country codes preferable to ISO 3166-1).
- **Complimentarity**
For acceptance of the identifier it shouldn't be necessary to replace already existing national court designated identifiers.

The proposed form of the ECLI is:

ECLI:[EU-countrycode]:[court code] :[year of decision]:[ordinal]

To give a few examples for converting national numbers:

Country	National case	ECLI
Italy	Corte Costituzionale 322/2008	ECLI:IT:CC:2008:322
Netherlands	LJN BC8581 (Hoge Raad, 01-04-2008)	ECLI:NL:HR:2008:BC8581
Belgium	Raad van State/Conseil d'État, 185.273, 09-07-2008	ECLI:BE:RSCE:2008:185273
Finland	KKO:2008:97	ECLI:FI:KKO:2008:97

Figure 5 Translating national identifiers to European Case Law Identifiers.

As with the national level, a unified and unique identifier doesn't say anything about the findability of the case. If in a legal document a Greek judgment is cited, a Spanish reader being interested in this case not only has to be able to find the case on the website of the Greek Supreme Court, but probably also wants to know whether this case is in the collection of Caselex, JURE, Jurifast, or any other international database.

So, apart from the ECLI as an identifier, we also need an index, containing:

- ECLI
- the attributes of the case: country, court, date, case number
- the databases / websites / law reviews where the case can be found

Also, the index has to contain a multilingual glossary explaining the country and court codes (to make clear to a user that 'HR' in 'ECLI:NL:HR:2008:BC8581' stands for 'Hoge Raad', which means 'Supreme Court').

Every database connected to the system would need to have a 'resolver' to redirect a hyperlink that is displayed by the ECLI-index to its own database. So, if the ECLI-register would be at www.ecli-site.eu then for ECLI:NL:HR:2008:BC8581 a deeplink can be constructed: www.ecli-site.eu/ECLI:NL:HR:2008:BC8581.

At the website of the Dutch judiciary this case might then be found at www.rechtspraak.nl/ECLI:NL:HR:2008:BC8581 and at a commercial site maybe at www.commercial-caselaw-site.eu/ecli/uri:ECLI:NL:2008:HR:BC1233 – the part preceding the ECLI is the resolver-address.

Moreover, the ECLI-index should have two interfaces: a graphical user interface (GUI) and a webservice interface, to be used by other databases and case law systems.

In the GUI a user has to be able to search by ECLI or attributes. The result should be an overview of all sites where the case is published – including subscription-based sites (figure 6).

ECLI:DE:BVerfG:2008:2556
Germany, Federal Constitutional Court (Bundesverfassungsgericht)
BVerfG, 2 BvR 2556/07, 29.1.2008

Versions found:

- [Site of Law Review X](#) – subscription-based
- [Site of Bundesverfassungsgericht](#) – free access

Figure 6 Search result in the graphical user interface of the ECLI-index.

Preferably a user with a subscription for ‘Law review X’ should be able to access the decision concerned without entering an access code. However, this requires the implementation of specific protocols²⁴, which might be too complicated in a first increment – for the moment indicating the availability of the decision on the referred site might be enough.

By using the webservice-interface national websites, search engines, content integrators and commercial publishers should be able to make use of all information in the ECLI-database.

Maintenance of the index shouldn’t be done by any central body, but – completely automated – by the content holders themselves.

A prerequisite for such an automated data-entry would be the existence of a national neutral identifier of the cases which are fed to the index – to uniquely identify the decisions and to minimize the risk of incorrect or double registrations.

The most essential part of the ECLI-index is the ECLI itself and the index-information on where the case can be found. Still though, other data – such as those described in the next section – might be added to the index.

4 Metadata

4.1 Introduction

Metadata are used for describing content, thereby enabling enhanced search and providing a better and faster access to information. Because searching by using plain text often doesn’t lead to useful results, case law search interfaces have specific filters – based on metadata – enabling searches not offered by Google.

To facilitate cross-border access to national case law, metadata should be standardized as much as possible. To illustrate just some of the advantages:

- By labeling the summary field in a unified way, search interfaces with a translation module (like the Common Portal of National Case Law, *supra* 2.2) could search

²⁴ OpenURL, ANSI/NISO Z39.88-2004

summary and full text of the judgment separately, as to improve ranking of the results (words found in summaries are more important than words only found in full text).

- By standardizing the legal subject of a case by using Celex-codes, searching for decisions that interpret a specific regulation or directive could be facilitated. E.g. with one search action all national cases on the European Arrest Warrant²⁵ could be retrieved.
- By standardizing metadata fields for decisions taken in appeal or first instance (preferably with an ECLI), the course of the case can easily be traced and displayed.

4.2 *Intermezzo: Bibliographical Entities*

Metadata are mostly added after a decision is rendered, although they are sometimes extracted from the original document or stem from the production process. For an adequate use of metadata it is important to distinguish between different documentary levels.

For this, a useful framework is offered by the Functional Requirements for Bibliographic Records (FRBR) of the International Federation of Library Associations and Institutions (IFLA).²⁶ The FRBR has four levels:

– **Work**

Definition: a distinct intellectual or artistic creation.

Traditional example: Shakespeare's Hamlet

Case law example: the decision of the European Court of Justice in case C-299/02

– **Expression**

Definition: the intellectual or artistic realization of a work

Traditional example: the translation of Hamlet in the Dutch language by a specific translator.

Case law example: the aforementioned decision in the Italian version. For case law three kinds of expressions exist:

- a. language, in case there are more than one authentic language versions;
- b. temporal, only on the rare occasion that a revised version is issued;
- c. editorial, the version of a judgment as edited for different audiences. E.g.:
 - i. the authentic version as handed out to the parties and archived in court;
 - ii. the full but anonymized version as published on the court's website;
 - iii. the shortened version as published in a law review, with all kinds of scholarly metadata added;
 - iv. a translated version.

– **Manifestation**

Definition: the physical embodiment of an expression of a work

Traditional example: the paperback edition from a specific publisher of the aforementioned Hamlet translation

Case law example: the PDF-version of this Italian ECJ-judgment.

– **Item**

Definition: the single exemplar of a manifestation

²⁵ <http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=CELEX:32002F0584:EN:NOT>

²⁶ <http://www.ifla.org/VII/s13/frbr/>

Traditional example: the specific Hamlet paperback I carry with me all day
Case law example: the PDF-document of the Italian ECJ-judgment on my desktop

It is important to realize that the first two levels are of an abstract nature. The work and the expression cannot be grabbed, they need the embodiment of the manifestation, and finally the item.

4.3 Which Metadata

To give input for a possible discussion, the following scheme gives an overview of metadata, relevant for national case law that is of interest for lawyers abroad.

Apart from the kind of metadata and a description also the documentary level (work – expression – manifestation – item) is indicated, together with the number of occurrences (0-1: the field is optional, but can not appear more than once in one case. 1-n: the field is mandatory, and can appear more than once, etc).

Metadata	Remarks	Level	Occurrence
ECLI	European Case Law Identifier	Work	1
National identifier	Identifier of the judgment	Work	1
Case number	Identifier of the case	Work	0-n
Date	Date of the judgment	Work	1
Member state	Could also be the international organisation (e.g. EU, EFTA)	Work	1
Name of court	Has to be in a multilingual glossary	Work	1
EU-law	Relevant EU-law. Has to be in Celex-format	Expression	0-n
National law	Preferably in a national electronic standard, possibly in the future in a European standard	Expression	0-n
Parties	Including their specific role	Work	1-n
Headnotes	A few words indicating the contents of the case	Expression	0-1
Keywords	Could be classification from EU-thesauri, Eurovoc, etc.	Expression	0-n
Summary		Expression	0-1
Formal relation to other cases	E.g. appeal. Preferably indicated with an ECLI, and including the kind of relationship	Work	0-n
Material relation to other cases	E.g. other opinion on the same question of law. Preferably indicated with an ECLI, and including the kind of relationship	Expression	0-n
Under appeal	To indicate whether the judgment is irrevocable	Work	0-1
Language(s) of the authentic judgment		Work	1-n
Language(s) of		Expression	1

expression			
Publisher		Expression	1
File format		Manifestation	1

Figure 7 Possible metadata for national case law on Community law.

4.4 Metadata Standards

There are different methods for describing metadata. For reasons of interoperability it might be advisable to make use of the most commonly used standard: Dublin Core.²⁷ This standard contains a basic set of fifteen descriptors, extendable with a more comprehensive set. To prevent this paper from becoming too technical, no mapping is made here between the abovementioned metadata and Dublin Core – this would be work for a standardization body, taking into account what already has been achieved on the national level (figure 8).

```
<meta name="DC.Title" content="FINLEX ® - Korkeimman oikeuden ennakkopäätökset: KKO:2008:97" />
<meta name="DC.Language" content="fi" />
<meta name="DC.Identifier" content="http://www.finlex.fi/fi/oikeus/kko/kko/2008/20080097" />
<meta name="DC.Creator" content="Korkein oikeus" />
<meta name="DC.Publisher" content="Oikeusministeriö" />
<meta name="DC.Subject" content="Oikeudenkäyntimenettely - Muutoksenhaku - Vastavalitus - Reformatio in peius" />
```

Figure 8 Code snippet from a decision on the Finnish website www.finlex.fi, where Dublin Core ('DC') is already implemented.

4.5 Metadata Search

Once there is agreement on which metadata to use, there are various possibilities to implement metadata searches:

- 1) Keep the metadata in the databases where the judgments are stored, and define an interface between a search portal and the connected databases. When a query is entered, a simultaneous search is performed on all connected databases. This might be a rather time-consuming action, because the number of connected databases might be quite substantial.
- 2) Pre-index the metadata by a search-engine on a central location. This solution is less dependant on the availability and performance of the connected databases.
- 3) Store the metadata in central repository, e.g. the ECLI-index. This solution will also perform quite fast, but has as a second advantage that no local storage of these metadata is needed. On the other hand one could wonder whether connected systems would supply metadata to a central repository without storing them in their own collection.

Any more elaborated solution should offer maximum search flexibility, a standardized interface and a minimum of rework and maintenance on connected systems.

²⁷ <http://dublincore.org/>

5 Document Structures

5.1 Introduction

Roughly speaking, judicial decisions have a comparable structure everywhere: they comprise the parties and their respective roles, the date of the decision, the facts of the case, the considerations, the final decision, names of the judge(s) and clerk, citations of cases and paragraphs of law, etc.

Every lawyer, provided that he has mastered the language in which the judgment is written, is able to recognize these elements, although they are often not indicated explicitly.

Unlike the lawyer, a computer is not capable of parsing a judgment that was drafted without a structured template into its constituting parts.

Whether conceived by clerk, judge or computer, explicitly structured judgments offer several advantages:

- The numbering of paragraphs facilitates referencing specific paragraphs of the judgment (both in writing and by deeplinking);
- search results can be improved if searches can be performed on specific parts of the judgment;
- for a computer, understanding the syntax of a judgment is an indispensable first step for understanding the semantics, and subsequently the legal reasoning of the judgment. This enables sophisticated tools for legal reasoning, quality control and knowledge tools.

5.2 A Joint Direction

Although the structure of judicial decisions from different countries is comparable, national peculiarities and traditions will prevent the development of a European decision template – attaining this at the national level is already problematic enough.

But, keeping an eye on the goal as stated above, agreement on a lowest common denominator might suffice. This lowest common denominator might serve as an interchange format, without restricting national expressivity.

To illustrate this with an example: the interchange format might contain an indivisible element for <dictum>, while at the national level there might also be a <dictumOnSubstance> and a <dictumOnCosts>. In the same sense the interchange format could contain a <partyName>, while at the national level this element is subdivided in name-parts which are specific for that country.

The most important pan-European initiative at the moment is Metalex/CEN²⁸, which is developing a CEN Workshop Agreement (CWA) on an Open XML interchange format for legal and legislative resources. Metalex/CEN is not restricted to document structures, but also comprises a sound architecture for metadata and identifiers for the various documentary levels.

The latest proposal [*Metalex 2008*] is primarily drafted for legislative documents, but the intention is to extend it to case law.

²⁸ www.metalex.eu

6 Conclusion

During the last years, we witnessed the emergence of quite some initiatives for the cross-border dissemination of national case law, especially case law with a European connotation. These initiatives all developed their own solutions for some fundamental problems of legal knowledge management. Without questioning the sometimes remarkable achievements, it might be wise – for the sake of the uninitiated judge, student, lawyer or citizen – to take a step back, and contemplate on ways to share knowledge and resources for tackling challenges which are common to all projects.

For metadata it is advisable to use the already existing Dublin Core standard, which is also web 3.0-ready. For document structures and identifiers no off-the-shelf solutions are available. Newly developed standards, replacing existing templates and solutions at the national level, are candidates for failure. Instead, to achieve interoperability the aim should be to develop interchange formats which can be used on top of national formats.

7 Literature

- Metalex 2008 A. Boer, F. Vitali a.o., CEN MetaLex Workshop Proposal
<http://svn.metalex.eu/svn/MetaLexWS/documentation/2008proposal/2008proposal.pdf>
- Mowbray 2000 A. Mowbray, A Uniform Approach for Vendor and Media Neutral Citation - the Australian Experience, Citations Workshop: strategies for accessing law and legal information, Edinburgh, Scotland - 11th & 12th March 2000.
<http://www2.austlii.edu.au/~andrew/citation.html>
- Van Opijnen 2006a M. van Opijnen, Uitspraken op Rechtspraak.nl - een representatief beeld?, Trema, 2006/1. http://www.marcvanopijnen.nl/pubs/uitspraken_rsnl.html
- Van Opijnen 2006b M. van Opijnen, 'A Public Index of Case Law References - the End of Multiple and Complex Citations', *Legal Knowledge and Information Systems - JURIX 2006: The Nineteenth Annual Conference*, Edited by Tom M. van Engers, ISBN 1-58603-698-X
<http://www.marcvanopijnen.nl/pubs/citations.html>
- Van Opijnen 2008 M. van Opijnen, Finding Case Law on a European Scale - Current Practice and Future Work, *Legal Knowledge and Information Systems - JURIX 2008: The Twentyfirst Annual Conference*
<http://www.marcvanopijnen.nl/pubs/ecli.html>
- Rikhof 2008 F. Rikhof: 'Juriconnect, realisatie van openbare standaarden binnen het fiscaal-juridisch domein', in *Intellectueel Kapitaal 2008*, nr. 2, p. 8-12