ONTOLOGICAL MODELLING AND COMMITMENT TO COMPARATIVE LEGAL THEORY. A CASE STUDY

Burkhard SCHAFER
Wim VANDENBERGE
John KINGSTON

SUMMARY: I. *Introduction*. II. *FF Poirot.* III. *Selected user requirements and resulting demands on the ontology.* IV. *Some theorical reflections.* V. *Conclusion.*

I. INTRODUCTION

Ontology based approaches have become increasingly widespread in the computer science community in general and legal information system design in particular. Their importance has been recognized in fields as diverse as knowledge engineering,[1] knowledge representation,[2] qualitative modelling,[3] language engineering,[4]

[1] Uschold, M, Gruninger, M., Ontologies: Principles, Methods and Applications, *"The Knowledge Engineering Review"*, (11), 1996, pp. 93-136.

[2] Guarino, N., Formal Ontology, Conceptual Analysis and Knowledge Representation, *"International Journal of Human and Computer Studies*", (43), 1995, pp. 625-640.

[3] Borgo, S, Guarino, N, Masolo, C., An Ontological Theory of Physical Objects, "*Proceedings of the 11 International Workshop in Qualitative Reasoning",* Corona, Italy, IAN-CNR, Pavia, 1997, pp. 223-232.

[4] Bateman, J.A., On the Relationship between Ontology Construction and Natural Language, *"International Journal of Human-Computer Studies",* (43), 1995, pp. 929-922.

database design,[5] object-oriented programming,[6] information retrieval[7] and agent based system design[8]. Applications span from enterprise integration,[9] natural language translation,[10] medicine,[11] e-commerce,[12] geographical information systems,[13] and of course law[14]. One of the main attractions of ontologies is their promise of simplicity and certainty in an ever more complex an ambiguous world. Global markets and the ubiquitous interconnectivity of systems and information processes in cyberspace that they bring with them have dramatically increased our awareness of the problems created by conceptual mismatches and failing system interoperability. For example, to fully exploit the potential of an autonomous agent that searches the internet to identify the

5 Burg, J.F., Linguistic Instruments in Requirement Engineering, IOS Press, Amsterdam, 1997.

6 Wand, Y., "A Proposal for Formal Model of Objects", Kim, W., Lochovsky, F. H. (eds.), *Object Oriented Concepts, Databases and Applications,* Addison Wesley, Reading, 1989, pp. 537-559.

7 McGuiness, D., "Ontological Issues for Knowledge Enhanced Search", N. Guarino (ed.), *Formal Ontology in Information Systems,* IOS, Amsterdam, 1998, pp. 302-317.

8 Franklin, S., "*Cybernetics and Systems*", Autonomous Agents as Embodied AI, (28), 1997, pp. 499-520.

9 Gruninger, M, Fox, M. S., "The Logic of Enterprise Modelling", J. Brown, D. Sullivan (eds.), *Reengineering the Enterprise,* Chapman and Hall, 1995.

10 Mahesh, K., *Ontology Development for Machine Translation,* New Mexico State University, MCCS-96-292, 1996.

11 Rector AL, Nolan WA, Kay S., "Methods Inf Med", Foundations for an electronic medical record, 30, 1991, pp. 179-86.

12 Lehmann, F., "Machine Negotiated Ontology Based EDI", *Proceedings of CIKM-94 Workshop on Electronic Commerce*, Spirnger, Berlin, 1995.

13 Casati, R, Varzi, A., "Spatial Entities", O. Stock (ed), *Spatial and Temporal reasoning*, Kluwer, Dordrecht, 1997.

14 Hage, J., Verheij, B., "*International Journal of Human Computer Studies*",The law as Dynamic Interconnected System of States of Affairs, A Legal Top Ontology (51), 1999, pp. 1034-1077; Visser, P.R.S., R.W. Van Kralingen and T.J.M. Bench-Capon, A Method for the Development of Legal Knowledge Systems, in: Proceedings of ICAIL'97, Melbourne, 1997.

best buy for its use, prices need to be communicated in a unified format, without the need to decide from the context if VAT is included or not.

The idea to agree on explicit and unambiguous subject taxonomies resonates particularly well with lawyers. Much of European Union legislation can be understood as the legal equivalent to ontology integration, most problems of private international law as partial responses to the problem of ontology mismatch where such higher level agreement can not be reached. Ontology based solutions have therefore unsurprisingly attracted the attention of lawyers working in multi-jurisdictional contexts. In the absence of supranational harmonization, these contexts are also particularly knowledge intensive, making the use of AI solutions even more plausible.[15] A UK lawyer might just about know the relevant EU regulations, but he can not be an expert for all the legal systems and their specific classificatory schemata he might encounter. The project in the centre of this paper, FF POIROT, is situated in just such a context. It addresses issues of cross-border financial fraud in the EU, a domain for which there is some harmonization already in place (in the form of the EU VAT directive for instance) but large areas are still governed by national law and its diverse conceptual and regulative approaches to these problems. In a typical application, an investigator in Germany might want to find out if his British suspect has relevant previous convictions. In accessing a UK database for this purpose, he risks however a mismatch between his conceptual definition of "previous conviction" and the one underlying the UK database, which for instance will not have a direct equivalent to the notion of "Strafbefehl".

15 Schafer, B., Grounding Legal Information Systems, "*Global Review of Cyberlaw*", (1), 2001, pp. 83-132; Breuker, J., Elhag, L., Petkov, E., and Winkels, R., (2002), Ontologies for legal information serving and knowledge management, T. Bench-Capon, A. Daskalopulu and R. Winkels (eds.), *Legal Knowledge and Information System,* IOS Press, Amsterdam, 2002, pp. 73-82.

However, before we describe the POIROT project, we will put its very preliminary findings into a broader theoretical context, and address what can be called the "paradox of the ontology movement". As noted above, ontologies should facilitate communication across different classificatory schemata. In its most basic form, an ontology is an agreed upon concept of domain specific knowledge. This formulation, which does not refer to any specific way in which this agreement is incorporated, allowed us to think of the work of the European Parliament as an exercise in ontology engineering. However, at the same time as ontology oriented modelling is developing into the most successful paradigm in AI, the field of ontology research itself is rapidly loosing its "agreed upon conceptual classification", including an understanding of what "ontology" stands for. "Reference ontology", "foundation ontology", "application ontology", "core ontology", "top/upper level ontology" are just some of the expressions used in the literature, and their precise meaning and the relation between them often remains obscure. Two opposite schools of thought are at the centre of this paper. On the one hand is the school which focuses primarily on the realism or adequacy of an underlying ontological theory. Researchers in this school often, but not always, come with a background in philosophy. On the other hand is the (much larger) school which focuses primarily on the construction of ontologies as working applications at the expense of ontological realism, and which is associated with current developments under the heading of the Semantic Web. Nicola Guarino, in one of the numerous attempts to introduce generally accepted classifications in ontology research, draws a distinction between "reference" and "application" ontologies, concepts which he defines thus:

> *...what kinds of ontologies do we need? This is still an open issue. Some people believe that very general ontologies involving rich axiomatic characterizations are important, others think they are a waste of time, and prefer to concentrate on lightweight*

> *ontologies, focusing on the minimal terminological structure (often just a taxonomy) which fits the needs of a specific community.*

Guarino uses "reference ontology" (which he now calls "foundational ontologies") to refer to ontological theories whose focus is to clarify the intended meanings of terms used in specific domains. Smith has proposed a more focused definition within the framework of philosophical realism:

> *a reference ontology is a theory of some independently existing domain of entities, a theory which maximizes descriptive or representational adequacy to the maximal degree compatible with the constraints of formal rigour and computational usefulness.*

In this paper, we will address the relation between the theory of comparative law and the choice of the appropriate notion of ontology in designing multi-jurisdiction computer support systems. Are "light" or "application" ontologies sufficient, or is there an identifiable need for "foundational" or "heavy" ontologies? We will develop several interrelated arguments for the development of computer systems that support lawyers who work in a multi-jurisdictional environment:

— Our analysis of user requirements indicates that "bare" knowledge of a foreign legal system is insufficient for many applications. By this, we mean a system that uses the national experts to develop conceptual classifications of their own legal system in isolation from each other, making their models accessible to other users only after they completed this process for their national system. We also include under this notion systems that in addition use thesauri to provide ad-hoc translations between terms of different legal systems. A user of such a system might learn for instance that for a specific application, "contract" is a sufficient translation of "Vertrag" or "contratto", but

he can't "drill down" into the ontological structure of the "contract-objects" denotated by these terms to analyse the adequacy and limits of this translation. In the vocabulary of Zweigert and Koetz, the system models knowledge of foreign law, but not comparative legal knowledge. Instead, we will argue that in addition modeling knowledge of the individual legal systems, genuine comparative law insights have to be represented as well.

— "Heavy" upper level ontologies are particularly well suited to model certain methodological approaches to comparative law. In particular, they sit well with functionalist accounts of comparative legal methodology in the tradition of Zweigert and Koetz, or economic analysis and comparative law as proposed by Ugo Mattei.[16]

— However, they stand and fall with the plausibility of these methodological approaches. Arguments made within comparative law against functionalism translate directly into arguments against certain approaches to ontology design, even if these designs are chosen on purely technological grounds. Legal ontologies in the form proposed by Breuker *et al* become as a result susceptible to two lines of criticism: Their structural similarity to functionalist models of comparative law opens them to criticism by structuralists, constructivist and cultural comparative law theories.[17] The realist philosophy underpinning "heavy" top level ontologies opens them to criticism by constructivist schools of epistemology and philosophy. This does not mean that this approach is misguided or wrong. It does however mean that the commitment to an underlying comparative legal and philosophical methodol-

16 Mattei, Ugo, 1997, "Comparative Law and Economics, Ann Arbor", MI, The University of Michigan Press.

17 For an overview see Schafer, B., "Form Follows Function Fails as a Sociological Foundation of Comparative Law", 1999, *Social Epistemology*, (13), 1999, pp. 113-128.

ogy needs to be made more explicit, and requires more in terms of justification, than technology driven approaches to legal knowledge modeling normally provide. The rest of the paper is therefore a case study on how this interaction between commitment to comparative legal theory, commitment to philosophical theories and user requirements can interact and inform each other.

II. FF POIROT

1. *A comparative look at VAT fraud*

It is estimated that the EU loses several billion euros per year due to financial fraud. Therefore it should not come as a surprise that prevention and early detection of fraudulent activity is an increasingly important goal for the EU and its Member States.[18] This paragraph will introduce a short overview of the development in Europe and indicate differences and similarities to the situation in NAFTA, although the reader is mainly refereed to the excellent comparative papers by Cockfield and Cnossen

After decades of effort, the member states of the European Union accepted a diminution of their sovereignty over tax policy when they implemented value-added tax (VAT) harmonization in order to promote economic efficiency. Beginning in 1967, the member states of the European Community instituted a common VAT system. The tax rates of the European VATs were harmonized on December 31, 1992; a minimum VAT rate of 15% is required, with reduced VAT rates of at least 5% for certain

[18] See Communication from the Commission, Protecting the Communities' Financial Interests. Fight Against Fraud, Action Plan for 2001-2003, COM, 2001, 254 final, A recent effort by the UK government is the initiative by HM Customs & Excise, titled "Protecting Indirect Tax Revenues", designed to save £2 bn a year.

goods and services.[19] The changes were most pronounced in the UK, which had to replace its existing Retail Sales Tax by the model preferred in continental Europe. While harmonization of tax law was achieved, as we will see the more basic divisions between common law jurisdictions (the UK) and civil law countries continue to affect the functioning and interpretation even of the "harmonized" legislation. To some extend at least, harmonization remains superficial.[20] Unlike the European Union, NAFTA is merely a trade pact; it contemplates only negligible political integration by its membership. Nevertheless, a good argument can be made that the movement toward freer regional trade and investment under NAFTA ought to be complemented by the gradual harmonization of North American tax regimes.[21] Increasingly, the problem of taxation of cross-border internet transactions adds further pressure on tax harmonization.[22] In the context of NAFTA however, Mexico is the exception, following with its VAT tax continental Europe, while the preference for a retail sales tax is deeply entrenched in Canada and the US.[23] While political resistance to tax harmonization within NAFTA might be insurmountable, the problem of cross border tax fraud impacts equally on NAFTA countries, and to this extend, the present analysis is applicable in this context as well. In addition to the divide between the European continental legal tradition, represented in NAFTA by Mexico, and common law, extending our analysis to NAFTA would face the additional problem of the

19 Sijbren Cnossen, "Coordination of Sales Taxes in Federal Countries and Common Markets", 9 *Conn. J. Int'l L.* 741, 1994.

20 Legrand, P., "*International and Comparative Law Quarterly*", European Legal Systems are not Converging, (45), 1996.

21 Cockfield, A., Tax Integration Under NAFTA, "*Stanford Journal of International Law",* Resolving The Conflict Between Economic And Sovereignty Interests, (34), 1989, pp. 39ff.

22 Cockfield, A., "Balancing National Interests", The Taxation Of Electronic Commerce Business Profits, "Tulane Law Review", (74), 1999, p. 133ff.

23 Laird, S., Latin American Trade Liberalization, *Minnesota Journal of Global Trade* 1995.

integration and harmonization of tax law in the federal states of Canada and the US *internally*. Problems of harmonizing tax within Canada or the US have been widely acknowledged, and it would be interesting to see if the present analysis could make a contribution to this debate.[24]

2. *VAT Fraud*

The impetus for building a financial fraud ontology results from the need to supplement the efforts of EU Member States to combat financial fraud, and especially to improve the cross-border co-operation between investigating and prosecuting agencies. FF POIROT is focusing on two areas of fraud: VAT (sales tax) fraud, which is the example used in this paper, and unauthorised online investment solicitation. Two examples of VAT fraud are "carousel fraud" and "missing trader intra-community fraud". In carousel fraud, goods are apparently exported, thus making them exempt from VAT, but the export does not actually take place; the transaction may then be reversed in the "receiving country", thus doubling the fraud while leaving stock levels in both countries unaffected. Missing trader intra-community fraud (MTIC fraud) is carried out (in general terms) by setting up fake or minimum-sized companies, which trade across borders in such a way that one becomes liable to repay VAT to the government, while the other has already claimed a VAT refund and these companies then disappear without trace. In practice, several layers of buffer companies may be used.

Complying with VAT regulations requires knowledge of a vast web of regulations, and is therefore a heavily knowledge-based task. *Detecting* VAT fraud appears to require less knowledge; once all the information is available, a numerical comparison between transaction sizes or identification codes on invoices, or a failure to match one invoice with another, will reveal the fraud.

[24] Ulbrich, *State and Local Taxation of Out-of- State Mail Order Sales,* Report to the US Advicory Committee on Intergovernmental relations 1985.

The European Community has developed a cross-border database (known as VIES, the VAT Information Exchange System).[25] The problem is in making the information available in a timely fashion. The current operation of VIES does not allow sufficiently early availability of such information.

In what follows, we introduce a few of the requirements our analysis has discovered so far. They are chosen here because they indicate needs for variously detailed axiomatic descriptions of the concepts in question, and illustrate the range of theoretical choices and commitments that have to be made. At present, these requirements and the ensuing theoretical commitments are seen in isolation, as specific technological problems to be solved by technicians. This paper then is a first attempt to see them interrelated through more basic, theoretical or philosophical choices, amongst which the decisions between reference vs. application ontology, description vs. revision ontology and "upward constructive" (functionalist) and "downwards constructive" (formalist) theories of law figure prominently. To anticipate somewhat: The multi-jurisdictional setting of the project means that we are confronted with several mutually incompatible conceptualisations of the VAT domain. This indicates the need for a "heavy", axiomatic characterization to minimise the potential for misunderstandings between users from different jurisdictions. This then raises the question of the ontological status of legal concepts. In one approach, legal concepts are ontologically dependant on the social reality they regulate. This is a functionalist approach to law, and there are well developed comparative law theories (Zweigert, Koetz)[26] and jurisprudential theories (Searle) that support this view. Decisions made on the modelling of evidence then have a direct impact on the modelling of legal regulations.

25 Certain parts of this common computer network can be accessed by businesses across Europe to check the validity of their customers' VAT identification number on the Internet since 14 June 2002. See http://europa.eu.int/comm/taxation_customs/vies/en/vieshome.htm.

26 Zweigert, Koetz, Einfuehrung in the Rechtsvergleichung, Muenchen, 1986.

However, the realistic assumptions underpinning this view are a two-edged sword. On the one hand, they give a good explanation why legal comparison is possible: two concepts from different jurisdictions are the same if they regulate (are dependent upon) the same bits of social reality. On the other hand, this approach almost inevitably (and comparative law research in this tradition bears witness to this) results in a revisionist ontology. In a medical expert system, to get the right result as opposed to the result accepted by the medical fraternity seems a desirable goal. In a legal expert system however, this might well contradict a user requirement. To get the "right" answer is of little help if courts remain unconvinced and indeed it is questionable what the "right" answer could mean if social institutions with the appropriate authority disagree with this result. As an alternative, there are equally well-developed "downwards constructive" jurisprudential (Teubner)[27] and comparative legal (Samuel)[28] theories. In these approaches, legal systems and concepts create their own (social and physical) reality. Decisions made about the correct representation of law then have a direct bearing on the correct modelling of evidence and other non-legal entities. However, this approach raises the question if two concepts of different jurisdictions can ever mean the same, and what "sameness" could possible mean. According to some writers in this approach, cross-legal correspondence of terms is merely conventional, not based on any "real" and objective correspondence, and is ultimately policy driven. As a result, a lightweight application ontology might after all be necessary to represent the *ad hoc*, merely conventional and highly context sensitive correspondences between legal vocabularies of different jurisdictions in a thesaurus. The legal philosophy of Reinach might provide a compromise solution between these two that combines upward

27 Günther Teubner, *Law as autopoetic System,* London, Blackwell, 1992.

28 Geoffrey Samuel, "*International Comparative Law Quaterly*", Comparative Law and Jurisprudence, (47).

and downward construction. Legal concepts exist independently and irreducibly from the physical world that they regulate, and it is indeed a function of the law to "cut out" pieces ("Sachverhalte") of this reality according to criteria intrinsic to the law. However, ultimately non-legal objects form parts of these Sachverhalte, which act as a bridge between law and social reality. As we will see, this approach has its own problems especially regarding the question of description vs. revision.

III. Selected user requirements and resulting demands on the ontology

The ideal system for supporting VAT compliance would be a knowledge based system that asks the user a few questions (in a choice of languages) and then supplies all the necessary information and forms for VAT compliance. This is very similar to a system that is already marketed by VAT Applications BV, a member of the FF POIROT consortium. If such a system were to be supported by an ontology, the requirements on the ontology would be:

— To represent legislation and legal rules from multiple countries;
— To support reasoning about compliance with these legal rules;
— To associate natural language terms in various languages with key concepts, so that (for example) key rules from one country can be viewed in the language of another country;
— To represent interpretation of legislation and legal rules useful to the user.[29]

[29] For example, HM C&E have refused to zero-rate supplies from companies which were complying with UK VAT law (notice 703). See, Case Study "How HM Customs & Excise misinterprets EC & UK VAT law to the detriment of British business trading in Europe", http://www.bavat.co.uk/cases/ec.htm.

As stated above, the most important asset in fighting cross border VAT in the EU would be an effective system of mutual assistance and information exchange in order to ensure the proper functioning of the VAT system. A possible use case would be the automatic and preferably spontaneous exchange of information to help in the detection of fraud in intra-Community trade. The ontology should support this system, providing a solid basis for the monitoring and enforcement of non-compliance of VAT laws. For this use case, the primary role of the system would be to enable two-way co-operation within and across agencies and within and across borders. This should include:

— Data integration within the same agency;
— Data integration between different (national) agencies;
— Data integration between two or more EU Member State agencies;
— Knowledge about legal methods of fraud investigation (which will differ between countries);
— Knowledge of typical indicators of fraud. Note that knowledge of methods of fraud is not essential, only of the indicators that point to the use of those methods;
— The system must deal with multiple data sources (VIES, ICT listing, etcetera), each having different functions and user interfaces;
— The system should know about law related to fraud investigation as well as fraud identification. For example, Customs & Excise in the UK have power of criminal investigation, but only the police have the power to investigate in Belgium.
— The system should automate repetitive tasks such as checking to ensure VAT numbers are valid.

VAT fraud prevention requires rather more wide-ranging knowledge than VAT fraud detection, since it is searching for indicators of possible future fraud rather than indicators of exist-

ing fraud. Its knowledge may therefore include past histories of individuals, psychological profiles, typical company organizations for fraud, and typical commodities traded in certain types of fraud.

Since the goal of such a system would be to direct VAT inspectors to more likely cases of fraud, there is also a need to represent national and supranational laws relating to VAT fraud and national laws relating to investigation of VAT fraud.

Listed below are the top-level requirements of the ontology to address the demo showcases on this project (VC indicates a compliance-related requirement, VD is related to fraud detection, and VP to fraud prevention). It is not yet clear whether all these requirements can be met with current technology, within the timescales of this project. However, in order to support these requirements, the ontology must be able to describe concepts in the domains listed below.

VC1: The law of various countries regarding offering VAT, and supranational laws governing cross-border VAT transactions in the EU;

VC2: Knowledge of products, goods and transactions that is relevant to the various laws;

VD2: The law regarding investigation of VAT fraud in various countries

VD3: Different databases and their communication formats

VD4: The law of various countries regarding offering VAT, and supranational laws governing cross-border VAT transactions in the EU;

VD5: An understanding of corporate structure (for tracking down missing trader frauds)

VP1: Knowledge of specific methods of fraud and related indicators

VP2: Knowledge of indicators of fraud relating to products and transactions

VP3: An understanding of companies and corporate structure, including:

— Management structure (partnership, etcetera);
— Managers (the director might be someone with a criminal record);
— Corporate structure (Does it have subsidiaries in other EU countries? Did head of Company A found other companies in neighbouring countries?);
— Does the company actually exist (or is it PO Box address only; answering machine, etcetera?);
— Who are its trading partners? Where are they located?

VP4: Knowledge of indicators of fraud relating to individuals and their history.

VP5: Links between individuals and others via past or present relationships, or via types of evidence.

— VP5. 1 Relationships.
— VP5. 2 Certain types of evidence.
— VP5. 3 Evidence databases.

VP6: The law of various countries regarding offering VAT, and supranational laws governing cross-border VAT transactions in the EU.

VP7: The law regarding investigation of VAT fraud in various countries

Many of these top level requirements must be broken down into more specific requirements. For example, requirement VP1 (Knowledge of specific methods of fraud and related indicators) requires that the ontology must be able to represent the intention of actors. The legal definitions of fraud require the presence of an element of intentionality on the part of the individual. It's quite possible to perform a crime unintentionally for non-compliance, this is normally due to ignorance of the correct procedures, while for fraud, it may occur if a legal trader is used as an unwitting link in a VAT carousel fraud. This leads to the following more detailed requirements:

a) *Ontology requirement 1.2:* The ontology must be capable of representing *plans* (modus operandi).

— The ontology must be able to represent actions. This should include identifying the actor(s); the resources created, modified, consumed or employed; and the resultant changes to the state of the world. Thus key elements to be represented by the ontology might include the fraudsters, their actions; the assets at risk; other assets required; and the indicators that a fraud has taken place.
— The ontology must be able to represent sequences of actions. Sequences of events reveal how related events unfold over a time period. The timing is important; for example, when did the suspected fraudster establish the involved companies?, When did he "sell" goods?, When were these goods transported?, When did he send invoices to the authority? From a legal viewpoint, there are three time intervals that are important: before, during and after the action (be it a fraud or the publishing of a non-compliant website) was committed. Because the focus of this project is on cross-border financial fraud, the location of the fraud and its aspects has to be determined. This is important mainly for the *assertion of jurisdiction* by the relevant investigative and prosecuting authorities. Questions to be asked are: Where is the company incorporated?, Where are the victims (tax authority, private investor) situated? The ontology must be able to represent that some actions are "before" or "after" others. It's not yet clear whether precise time intervals between actions need to be represented.
— The ontology must be able to represent where the fraud happened (where harm is done), fraudsters are located, etcetera. Initially, the level of detail need be only the name of the nation(s) in which the individual resides/company is registered.

Another example is the representation of laws regarding VAT, which is required for all three possible scenarios (see requirements VC1, VD4 and VP6). This requires that:

b) *Ontology requirement 1.3:* The ontology must be able to represent *legal rules*:

— The ontology must be able to represent legal and administrative rules in different countries across the EU;
— The ontology must be able to represent definitions of fraud, sanctions for fraud etc. in different countries across the EU;
— The ontology must be able to support matching of plans against multiple legal rules from a single jurisdiction.

This last requirement is important, for there are few laws that consider common fraud plans as crimes directly; once such laws are made, the fraudsters simply change their plans. For example, VAT carousels do not have their own statutory penalization; but they can be prosecuted as an amalgam of criminal offences, such as specific fiscal offences and forgery.

Some of the other specific requirements that arise are listed below.

c) *Ontology requirement 2.1:* The ontology must be able to represent practical ("on the field") knowledge about authorities involved in investigation of fraud, derived from legal rules. For example: Who are the relevant investigating, prosecuting authorities? What is the structure of the investigating department?

d) *Ontology requirement 3.1:* The ontology must be able to represent products as concepts with multiple properties. The properties may be symbolic or numeric, single or multi-valued, may contain links to other concepts or properties, and may be derivable from other properties.

Commercial transactions require a sale, a vendor and a purchaser. A sale is an agreement between two legal entities for the supply of a product or of a service for a price.

e) *Ontology requirement 4.1:* The ontology must be capable of representing a Sale, an action that transfers ownership of a Product from a Vendor to a Purchaser (who must be a Legal Entity), and transfers a Sale Price from the Purchaser to the Vendor. A knowledge of the structure of companies is also needed.

f) *Ontology requirement 7.2:* the ontology must be able to represent links between companies, both parent/subsidiary links and trading links.

IV. SOME THEORETICAL REFLECTIONS

We have indicated the wide variety of concepts that need to be represented in order to represent the workings of financial fraud. It can be seen that some requirements make high demands on ontologies: to give two examples[30], natural language "terms" cannot merely be seen as annotations to "concepts", but need to be looked at as concepts themselves (in order to satisfy the requirements to be able to detect fraud markers in the material provided by the companies like annual reports etcetera); and also, the system must allow universals to be related to particulars (*e.g*. "the Belgian VAT law" is a particular but a person violating it must be represented as a universal). The choice of a general ontology on which the FF POIROT will be based should therefore be made very carefully.

It is the cross-border, multi-language and multi-profession environment that makes ontology-based solutions attractive. For an investigation conducted in the UK, the question whether the director of the German supplier of a British company has relevant previous convictions can be an important indicator of fraudulent activities. The relevant information will be held in a German database, labelled in German and using a largely contingent format of recording. An ontology-based approach should enable the in-

30 With acknowledgements to Werner Ceusters of L and C Computing BV, from whom these examples are quoted.

vestigator to give the system a query in his language ("does the director of this company have previous convictions?") and get a reliable answer in English that is affirmative if the "Vorstand" of the "Gesellschaft" has a "Vorstrafe". The processes involved on successful fraud detection and prosecution therefore involves.

— To explain the meaning of a term to somebody outside the community for instance to explain to a UK lawyer that "Vorstrafe" covers (most of) "previous conviction".
— To negotiate meaning between agents belonging to different communities (professional and national). This task is in our context partly fulfilled by the EU directives that have exactly this function. However, the system needs to represent how different (and possibly mutually inconsistent?) national laws "implement" one and the same directive.
— (Ideally) to establish consensus to accept *e.g.* that a German "Haftbefehl" is similar enough to a UK "criminal charge" for the UK authorities to act on the German court order to seize assets held in the UK as evidence.

These three functions, explanation, negotiation and consensus are listed in (Borgo *et al*)[31] as indicators that an explicit representation of ontological commitments is necessary to exclude terminological and conceptual ambiguities. A rich axiomatisation can deliver that. In this specific example for instance, it is possible to represent axiomatically the mereological structure of companies in German and UK law.[32] The two concepts (or wholes) are by no means identical, the German company having a necessary part (the "Aufsichtsrat") that is missing in UK law. The UK lawyer would however see that for each of the parts in a

[31] Borgo, Gangemi, Guarino, Maselo, Oltramari, Wonderweb deliverable D15.

[32] Burkhard Schafer, Inheritance principles and the community of heirs, in N. Guarino (ed.), Formal Ontology in Information systems, IOS Press, Amsterdam, 1998.

UK company, there is a corresponding part on the German company and that in particular "Vorstand" corresponds to "Director" not only regarding his legal function, but also in respect of his place in the complex structure that makes a company.

Another of the ontology requirements introduced above makes this approach highly desirable. The graphic representation of specific networks as evidence was introduced above as one of the desiderata of the system. In this instance, it facilitates the communication of the investigative expert, for instance a forensic accountant, with the prosecutor and possibly a court or jury. However, Gestalt-psychology teaches us the intuitive compellingness of such a graphical arrangements which makes it almost inevitable to see groups in graph layouts even if they are not substantial enough to establish say a "conspiracy to defraud" for legal purposes.[33] The historical link between Gestalt psychology and mereology developed by Gurwitsch could allow us to capitalise on the axiomatic features of the links in the graph to minimise misunderstanding. For instance, it might be desirable to represent only those links whose formal features guarantee that the ensuing whole is a "pregnant" whole.

While shared ontological commitment facilitates understanding, we encounter here a serious difference between similar projects in the natural sciences or medicine, and the legal domain. In the sciences, the underlying philosophical realism assumes convincingly that there is the possibility to agree on the relevant classification. The one shared reality sees to this. In law, the parliaments of the individual member states can create divergent realities. While German and UK law will agree that an extended prison sentence handed out by a court is a "previous conviction"

[33] Sparrow, Malcolm K., 1991 b). "The application of network analysis to criminal intelligence: An assessment of the prospects", Social Networks 13:251-274, Jensen, David, 1997, "Prospective Assessment of AI Technologies for Fraud Detection: A Case Study", AI Approaches to Fraud Detection and Risk Management, Collected Papers from the 1997 Workshop, Technical Report WS-97-07, Menlo Park, CA, AAAI Press html.

and "Vorstrafe" respectively, it cannot be taken for granted that UK law recognises the equivalent of a German legal term such as "Strafbefehl"[34]. If such a "Strafbefehl" is recorded against the name of the company director, should the system return an affirmative or a negative answer to our initial query? In an application ontology, this decision is taken pragmatically based on the interests of a small group of users with clearly defined goals and extracted from the texts they produce. For the investigator, the question is if his suspect had previous problems with the law, not the precise legal classification of this problem, and the classification of the domain will represent this. For the prosecutor, a much more restricted reading is necessary that tells him that the "Strafbefehl" did not automatically disqualify a person from becoming a director, the way an equivalent conviction would have. The problem is moved to the lexicon. This however means that the ontology loses "portability" and with that one of the main advantages of an ontology based approach. While irrelevant for forming an initial suspicion, the legal difference becomes highly relevant when the decision is taken to ask the German authorities officially for help in securing the relevant evidence, and might be crucial once legal proceedings are started. The alternative is to re-introduce realism and to treat the issue not as one of translation but of theory formation in comparative law.

(Breuker *et al*) have developed for similar applications an approach that explicitly incorporates philosophical commitments and comparative legal methodology.[35] In their "functional ontology", they develop a hierarchical model of reality in which legal concepts are made up of the elements of social or physical reality that they regulate. This functional approach to legal concepts has great intuitive appeal and has formed the basis of mainstream comparative analysis since the days of Zweigert and

34 Similar to a fixed fine, literary "punishment order".

35 Breuker, Elha, Petkov, Winkels, Ontologies for Legal Information Serving and knowledge management, Bench Capon (ed.), Legal Knowledge and Information Systems, IOS, Amsterdam, 2002.

Koetz and more recently in economic analysis and comparative law. In philosophical terms, Searle's constructivism seems to come very close to this approach. In this model, agreement between different legal conceptualisations is always possible in principle by recourse to the shared reality that the norms regulate. However, this theoretical commitment means that the criticism levelled against these schools also applies to the resulting formal system. In comparative law, it has been long recognised that the reductionism inherent in a functionalist approach can result in serious misrepresentations of foreign law. In philosophy, Smith has shown how Searle's approach either faces insurmountable problems accounting for crucial legal features or collapses into an unconvincing treatment of legal concepts as mere "*façons de parler*".[36]

While we agree with Smith's analysis of Searle, we think that some of the examples he uses point to even more serious difficulties that ultimately affect all realist approaches to legal ontology modelling. Of particular relevance to our setting is his treatment of a border conflict between China and Russia: both jurisdictions, on the basis of their law, might claim a specific piece of land. Searle has to treat these claims as valid even if they are not only mutually exclusive, but also solvable if higher-ranking law (*e.g.* international law) is employed. In our context, the presence of EU directives will sometimes solve similar disputes but not always. To complicate matters further, different jurisdictions might disagree on exactly when EU law is competent. In legal doctrine, this is the issue of "competence competence", and an as yet unresolved, and on legal-conceptual grounds possible irresolvable, question.[37] Furthermore, and again in stark contrast to realist assumptions, it might be possible

[36] Barry Smith and John Searle, The Construction of Social Reality: An Exchange, "American Journal of Economics and Sociology", (60), 2001.

[37] Neil MacCormick, Beyond the Sovereign State, *Modern Law Review,* 1993, pp. 1-18, Bankowski, Schafer Mistaken identities, in: Hoecke (ed), European Private law in context, Hart, Oxford, 2000.

that different national jurisdictions implement the same directive in a way that is only "upward consistent", that is, each national conceptualisation is consistent with the EU blueprint, but they remain mutually inconsistent. To "solve" contested legal issues through formal features of an ontology seems undesirable. Even though one could hope that such an analysis contributes to the quality of the discussion in that field, to impose them on a user who has to live with the relevant authorities seems overly ambitious. The requirement to represent jurisdictions as physical spaces as mentioned above therefore clarifies an important issue, but care must be taken that decisions on the correct modelling of space does not pre-empt the underlying legal question.

Another interesting group of examples in our context that is also discussed by Smith results from the requirement to represent temporal notions. Smith uses the examples of a legal claim whose documentary evidence had been lost together with any memory of the claim before it is "resurrected" by an heir centuries later. Another example he offers is the treatment of property in the former RDG ("East Germany") after reunification. For him, and against Searle, this shows that legal objects can exist even if they are not recognised by anybody. From a legal doctrinal point however, only his first example is unproblematic. It is however at least possible on legal grounds to analyse the second case differently, as introducing objects that exists intermittently in time, popping in and out of existence. Consequently, a realist ontology which rejects them (on good philosophical grounds) would have to revise the ontology assumed by the legal system.

Any realist ontology will on numerous occasions face similar problems. Legal concepts do not evolve in isolation. They are influenced by (often misunderstood) philosophical doctrines of the time when they were first recognised, and law in many ways resembles a scrap heap of discarded philosophies. Fraud is in many respects particularly fruitful (or difficult), because it typically involves an attempt to deceive someone about certain at-

tributes of an object which might or might not even exist, giving fertile grounds for legal systems to come to their specific answers on how to treat properties of non-existing objects, or to distinguish between attributes and essences. Certain distinctions in UK law on deception as to the identity or attributes of a fraudster (*e.g.* the difference between claiming to be Bill Gates, or to be just a very rich person) can't be found in German law. To be sound from the perspective of comparative law, our ontology must not pre-empt this question in favour of one or the other.

As a radical alternative, we find "downward" constructivist theories in both comparative law and legal theory. They still allow for a "heavy", axiomatic approach to legal ontology, and many formalist theories of law are actually much better suited for such an approach than some of the more pragmatic functionalist solutions. They do however turn the hierarchical conception of ontology upside down. The foundational ontology here is an axiomatic description of the more basic concepts of the respective legal system; the rest of the universe of discourse is derived from these legal conceptual commitments. Law creates its own realit(ies). An example in point is the legal treatment of electricity. Art some point in time, legal systems had to decide whether electricity is an object (and hence can be stolen) or a process (and hence something that can be fraudulently obtained). Different legal systems differ in their answer to this issue. Downward constructivist approaches take this difference seriously, and instead on deciding in the foundation ontology the status of electricity have multiple representations of it, one for each legal system.

A consequence of this approach is an inflationary, multiplicative approach to ontology, again borrowing from (Borgo *et al.*). Regardless of the question whether one adopt this radical stance or not, a multiplicative ontology seems one way to solve most of the issues discussed here, and in particular the question of jurisdiction. In those cases where the law does not provide a clear-cut

solution, to assume that different and mutually contradictory legal entities can nonetheless have a co-localised existence seems a solution, which comes very close to the self-understanding of the law. As fraud under German law, a specific scenario might be punishable; as clever trading under UK law, it might not.

The problem with a downward constructivist approach is that it can't explain how legal system reach non-arbitrary agreement over meaning of terms, which in a context dominated by EU legislation is the norm and not the exception. Their problems go even further than that. One of the requirements from above is the need to represent intention in criminal law. This intention again involves some understanding of the law. Individual fraudsters might have understandings of the law which are very different from the "official" reading, and if it is possible (as required by law) to say that nonetheless, their understanding is "sufficiently similar" to the correct one to constitute legal intent, such an assertion seems to rely on the same mechanisms this school considers impossible for cross-jurisdictional analysis.

A possible compromise solution is a multiplicative reading of Adolf Reinach's legal ontology.[38] It showed as early as 1921 how we can think of a realist model of comparative law with jurisdiction-independent legal concepts, of which the national laws are but instantiations. The vocabulary to describe these supra-jurisdictional concepts cannot any longer be extracted from legal texts, but could be constructed *a priori* and axiomatically described. The corresponding axiomatic descriptions of national law are then partial models of these more general concepts, a relation that in turn can be axiomatically characterised. Law largely retains its autonomy in this approach, and the (legal!) notion of "Sachverhalt" becomes the crucial glue between legal concept and social reality.

[38] On the Cognition of States of Affairs, K. Mulligan, ed., *Speech Act and Sachverhalt: Reinach and the Foundations of Realist Phenomenology*, Dordrecht/Boston/Lancaster, Nijhoff, 1987, pp. 189–225.

However, this approach raises other philosophical and practical problems. In line with Reinach's philosophy, but in stark contrast in particular to the UK understanding of law, this would mean that national courts could get their conceptual definitions "objectively" wrong. While in a medical expert system, the correct answer is what the patient wants or needs even if it differs from the view of a specific human expert, in a legal context predicting the court might be more important than being "objectively" right. Since access to court decisions is a user requirement for our application, this issue would still need resolving. Unlike the functionalist approaches discussed above, "wrong" here only means internal conceptual contradictions or incompleteness. This indicates the inherent problem of this approach. As noted above, philosophy and law often co-evolve. Unlike in the cases mentioned above, here philosophy borrows from law. The problem remains the proximity between one specific legal tradition, here German law, and the resulting philosophy. Or put differently, a Reinachian ontology does not impose so much a philosophical perspective on the law, but German legal concepts on the rest. To give a quick illustration, for Reinach it is *a priori* self-evident that the legal notion of contract and *delict* are instantiations of the more abstract concept of obligation. This is indeed an organisational principle of the German (or more generally continental legal) Civil Code. The consequence of this is that contract law cases that clarify the notion of obligation impact with necessity on the understanding of the law of *delict.*[39] The ontology as a result produces potentially revisionist arguments, but at least in the case of German law, these revisions would be *legally* sound. Common law does not share this understanding, and here legal concepts can be essentially underdetermined if a court does not explicitly rule on them. The open is-

[39] B. Schäfer: Reinach, pragmatic universals and the methodology of comperative law, K. Dorn (ed.) *Pragmatics in Contemporary Philosophy*. Beiträge der Ludwig Wittgenstein Gesellschaft, Wien, LWS, 1997, pp. 850-856.

sue remains if a multiplicative reading of Reinach can avoid cases of "bad" revision. This in turns depends on the question how the ontologically problematic "incomplete" concepts of the common law can be represented, that is concepts that should apply to a given sachverhalt (by the standards of UK law), but in the absence of a formal court decision are as yet lacking appropriate authorization.

Legal concepts in common law countries, understood internally, display certain features that we also fin din objects of a very different kind, the fictitional objects of literature and film. "Sherlock Holmes" has certain properties that are explicitly introduced by the Conan Doyle stories. He is male, he is English, he smokes a pipe. However, while his author, as a human being, also has the property of having a specific blood group (even if nobody ever knows which one), it does not make sense to ask of his creation, Sherlock, which blood group he has until such time as this information is explicitly introduced in one of the stories. Fictional characters are incomplete at any given time, but gain more and more features as the story progresses. The only constraint is that new attributes must be consistent with attributes already introduced. To treat (in a common law context) legal objects as fictitional objects would account for instance of Dworkin's intuition of law as a chain novel. An ontology that can model both actual and fictitional objects, along the lines *e.g.* of Edward Zalta's axiomatic metaphysics, could encompass both common law and civil law ideas about the status of legal concepts.[40]

V. CONCLUSION

Ontology modelling of multi-jurisdiction problems requires sound underpinnings in comparative legal theory. However, the

[40] Zalta, E., *Abstract Objects: An Introduction to Axiomatic Metaphysics*, Dordrecht, D. Reidel, 1983.

conceptual vocabulary of computer science can also enrich the theoretical discourse in comparative law. We started our case study with the claim that existing examples of legal knowledge representation ignore or even contradict our best comparative law theories. However, in trying to incorporate these comparative legal insights into formal computer models, we also changed the nature of comparative law into an exercise in comparative formal ontology. If computer science should succeed in reconciling legal theory with comparative law, sadly separated at the birth of comparative law as academic discipline, this alone would be a major achievement.