

Tato zpráva shrnuje nejdůležitější aktivity komise TNK136, která integruje činnosti Evropské komise CEN/TC278 a světově působící ISO/TC204.

**Výroční zpráva předsedy
technické komise
TNK136
„Dopravní telematika“
za rok 2013**

Intelligent Transport Systems

prof. Pavel Příbyl



ROAD TRANSPORT and TRAFFIC TELEMATIC



OBSAH:

1. Organizace komise TNK 136 „Dopravní telematika“	2
1.1 Organizace práce	2
1.1.1 Národní plán zavádění mezinárodních norem	4
1.1.2 Terminologická norma ČSN 73 6100	4
1.1.3 Projekt STANDARD	4
2. Významná zasedání CEN TC278 a ISO TC204.....	5
Jarní zasedání – CEN TC278.....	5
Podzimní zasedání – CEN TC278	5
Jarní zasedání - ISO TC204.....	7
Developing of a Data Registry/Data Dictionary (WG1)	7
Podzimní zasedání - ISO TC204	9
3. Znalostní systémy – další rozvoj	9
3.1 Projekt ESSENCE/STANDARD	9
3.2 Ontologie a znalostní systémy	9
3.3 Co se udělalo v letošním roce.....	10
3.3.1 Semiautomatická tvorba ontologie.....	10
3.3.2 Ontologie vytvářená expertem z databáze termínů.....	15
3.4 Znalostní systémy – závěr.....	18
3.4.1 Projekt EXTRACT/ESSENCE.....	18
3.4.2 Ontologie	18
4. Restrukturalizace činnosti ISO TC204	19
5. Závěrem.....	19

1. Organizace komise TNK 136 „Dopravní telematika“

Komise TNK 136 „Dopravní telematika“ vyvíjí od ustavující schůze ze dne 4. března 2003 samostatnou činnost, výhradně se zaměřením na standardizaci v rámci evropské CEN v technické komisi TC278 „Road Transport and Traffic Telematics“¹ a stále hlouběji i na světovou standardizaci ISO v TC204 „Intelligent Transport Systems“.

1.1 Organizace práce

TNK 136 pravidelně organizuje čtvrtletní zasedání jejích členů na ČNI, tedy čtyřikrát ročně se pravidelně schází s velmi dobrou účastí okolo 20 členů. V letošním roce se podařilo s jistými problémy stabilizovat finanční situaci, takže byla zajištěna základní činnost komise s omezením zahraničních cest na úkor tvorby další skupiny Extraktů.

Letošní zasedání komise se konala:

7. března 2012

Jubilijní 40. zasedání komise bylo výjezdní a se konalo v Brně a bylo spojeno s prohlídkou vily Tugendhat, památky UNESCO. Z významnějších bodů jednání:

- Ing. Z. Švédová převzala plně agendu gestora WG3 po Ing. S. Bartákovi, který zůstává nadále členem TNK 136;
- s poděkem bylo přijato bezkolizní předání agendy výkonného tajemníka komise a CTN Mgr. Lence Svorové od Mgr. Davida Bárty;
- bylo představeno Nařízení EU 1025/2012 o evropské normalizaci, které podporuje její rozvoj a klade větší důraz na přínosy národních norem včetně jejich převodu na řešení v úrovni EN;
- TNK 136 schválila písemné podklady prof. P. Příbyla k reorganizaci ISO/TC 204 a doporučila jejich předání jako oficiálního stanoviska ČR;
- byl odsouhlasen návrh finančního plánu na rok 2013 (dok. č. 136/764). Smluvně sjednané řešení projektu pokračuje v rámci MD ČR v intencích reálného naplnění plánu na rok 2013 včetně zařazení další práce na extraktech (STANDARD III B);
- za 1. kvartál byly odevzdány 4 normy jako konečné verze překladem: CEN TR 16152, CEN TR 16092, CEN TS 28701 (130 stran), ISO TS 24533 a další 4 normy převzaté originálem s přeloženou terminologií (ISO 21210, CEN ISO 17263, EN ISO 17262, EN ISO 17261);
- pokračování projektu STANDARD III B 2013 (extrakty): pokračování prací z roku 2012. Bude zpracováno dalších 25 extraktů;
- diskutovány plánovací podklady k řešení nového normalizačního úkolu, který tvoří samostatnou pátou část sady ČSN 73 6100 „Názvosloví pozemních komunikací – část 5: Dopravní telematika“. Práce navazují na předchozí dva rozborové úkoly (2004 a 2009) a dokončí tak legitimní kodifikací názvosloví nového oboru v podobě ČSN

¹ Po schválení orgány CEN se na Londýnském zasedání komise přejmenovala na „Intelligent Transport Systems“, přičemž číslo komise je beze změn

- CDV představilo rozsáhlou práci na národní normě ČSN 01 8245 „*Informační systémy ve veřejné dopravě osob – celostátní systém informací v reálném čase (CISReal)*“. Tato norma navazuje na evropské normy, ale plně zohledňuje národní podmínky. Je hrazena z výzkumných úkolů CDV.

6. června 2012

Z významnějších bodů ze zasedání:

- předseda TNK seznámil členy s jarními zasedáními v Seattle (ISO) a v Bruselu (CEN). Z obou zasedání mají naši gestoři kompletní dokumentaci. Zvláště nutné je sledovat zprávy konvenorů;
 - o na zasedání ISO bylo představeno naše stanovisko k reorganizaci ISO navržené na zasedání v Moskvě delegací USA;
 - o trvá problem nižší účasti našich gestorů na zasedáních; bylo by vhodné zajistit alespoň tři cesty ročně;
- za 2. kvartál bylo odevzdáno 8 norem v rámci národního plánu přebírání norem;
- byla provedena kontrola odevzdávání extraktů za 2. kvartál;
- gestoři byli požádáni, zda by nepřipravili myšlenkové mapy pro slova, které připravují; prezentace o myšlenkových mapách XMind: cílem je najít hierarchické uspořádání slov;
 - o Ing. Vlčínský nabídl převod z Excel do FreePlane;
 - o termín zpracování do 31.8

5. září 2012

Z významnějších bodů:

- zasedání se zúčastnil doc. P. Bouchner z Fakulty dopravní ČVUT, který by měl zajišťovat gesci skupiny WG10 “HMI”, která v komisi nemá zastoupení;
- zasedání byl přítomen interní auditor z ÚNMZ Ing. Josef Boháček a Ing. Stefan Krebs, který pracuje jako senior consultant na odboru technické normalizace;
- byla diskutována otázka legálnosti extraktů vyplývající z diskuse na zasedání CEN v Bruselu. ÚNMZ vydal k tomuto tématu kladné stanovisko a práce na tvorbě dalších extraktů v projektech STANDARD 3A, 3B potvrzuje funkčnost tohoto formátu v rámci práce TNK 136;
- i v období letních měsíců a dovolených gestoři včas zaslali svá vyjádření ke 22 dokumentům, stejně jako korektury řešených NÚ. Díky tomu bylo za 3.Q bylo odevzdáno 15 návrhů norem, z toho 9 konečných návrhů;
- bylo konstatováno, že za 2. a 3. Q chybí pouze 4 extrakty;
- Mgr. Bárta představil prezentaci „*Návrh struktury názvoslovné telematické normy*“. Gestoři mají za úkol zkusit uspořádat termíny dle tohoto principu;
- podstatné úsilí bylo věnováno dokončení terminologické normy ČSN 73 6100-5. Společné celodenní připomínkové jednání k názvoslovné ČSN 73 6100-5 se uskutečnil ve středu 23. října na ÚNMZ, v 9:30;
- norma ČSN 1 8245 „*Informační systémy ve veřejné dopravě osob – Celostátní systém informací v reálném čase*“ (CIS Real) již prošla připomínkovým jednáním a dokončují se poslední vypořádání se zaslánými připomínkami.

28. listopadu 2013: závěrečné hodnotící zasedání, ke kterému se vztahuje i tato zpráva.

Obecně platí, že jarní a podzimní zasedání jsou plánována tak, aby se konala vždy před plenárním zasedáním CEN TC278 resp. ISO TC204, protože jsou přípravou na tato zasedání. Na těchto jednáních předkládají gestoři a členové komise své náměty, dotazy a návrhy, které jsou pak tlumočeny jako postoje České republiky. Zimní zasedání v listopadu je vždy konáno jako příprava ke zpracování a bilancování činností za uplynulý rok.

Vždy před zasedáním jsou v dostatečném časovém předstihu zasílány všem členům pozvánky s programem nadcházejícího jednání. Je důsledně dodržován program jednání, jehož struktura se během mnoha jednání optimalizovala na následující body a zůstane zachována i nadále:

- Shrnutí a diskuse připomínkování dokumentů CEN, ISO v uplynulém období od posledního setkání;
- Prezentace a komentáře ke všem došlým dokumentům sekretariátů CEN a ISO;
- Revize plánu výjezdů a komentář gestorů z výjezdů;
- Informace ČNI o zařazení nových norem do Věstníku;
- Informace o novinkách webu TNK 136 a publikační činnosti;
- Prezentace aktivit pracovních skupin (pouze novinky, ke kterým došlo od posledního zasedání).

Z jednání jsou pravidelně pořizovány záznamy, které jsou vždy rozesílány všem členům komise, navíc je jejich znění zveřejňováno na webovém portálu Silmos a ČNI. V letošním roce se jedná o dokumenty: TNK 136/768 (březen); TNK 136/780 (červen) a TNK 136/804 (září).

1.1.1 Národní plán zavádění mezinárodních norem

I v letošním roce platí rozdělení na kategorie normativních dokumentů podle klasifikace A – D, kdy kategorie A i B požadují převzetí do ČSN překladem, kategorie C převzetí originálem s překladem terminologie a kategorií D gestor nedoporučuje normu momentálně do ČSN přejímat. Zároveň bylo nutné, u norem navrhovaných k převzetí do ČSN překladem v kategorii A, dodat argumentaci a důvody pro toto převzetí, které slouží k podpoře stanoviska MD ČR jako ÚSU.

Národní norma pro veřejnou hromadnou dopravu: norma ČSN 1 8245 „*Informační systémy ve veřejné dopravě osob – Celostátní systém informací v reálném čase*“ (CIS Real), která je produktem kolektivu okolo CDV a její tvorba byla hrazena z prostředků na výzkum je v době psaní tohoto dokumentu před vydáním. Její význam je v tom, že je zaměřena na národní prostředí s plným respektováním zahraničních standardů.

1.1.2 Terminologická norma ČSN 73 6100

Pátá část národní terminologické normy ČSN 73 6100 „*Názvosloví pozemních komunikací – část 5: Dopravní telematika*“ navázala na předchozí dva rozborové úkoly (2004 a 2009) a tato část uzavírá řadu terminologických norem. Práce byly organizovány Silmos ve velmi detailním členění, s přesnými a nepřekročitelnými termíny. Díky tomuto tlaku se podařilo zvládnout velkou práci, která spočívala na gestorech TNK 136. Norma je připravena k závěrečné redakci a vydání.

1.1.3 Projekt STANDARD

Dále pokračuje úsilí v tvorbě extraktů norem, protože jejich zpracování představuje nejefektivnější způsob seznámení české technické veřejnosti s obsahem českých a mezinárodních norem ITS. V loňském roce 2012 bylo, jako nouzové řešení díky nezajištění financování, zpracováno 25 extraktů. Stejně tak byl splněn cíl zpracovat v roce 2013 dalších 25 extraktů, za cenu úsporných opatření z limitovaného rozpočtu projektu mezinárodní spolupráce. Jednalo se opět o velmi vstřícný krok komise, kdy jsou extrakty vytvářeny navíc, nad rámec činnosti gestora.

2. Významná zasedání CEN TC278 a ISO TC204

Jarní zasedání – CEN TC278

Jarní zasedání CEN/TC278 se konalo v Bruselu s programem:

- 19. března ... 44. zasedání konvenorů
- 20. března ... 50. plenární zasedání

Podrobná zpráva ze zasedání je přílohou zprávy předsedy TNK136. Na zasedání jsem požádal o možnost prezentovat náš přístup k zavádění norem do praxe, což bylo akceptováno. Mluvil jsem jednak o projektu ESSENCE, což je jinak nazvaný STANDARD a dále o nutnosti tvořit ontologii pro oblast ITS.

Výsledkem bylo, že se tento přístup líbí předsedovi L. Eggingovi, kterému jsem slíbil, že začneme pilotně vytvářet ontologii pro EFC. Na druhé straně jsem měl dojem, že R. Bossom, který vytváří slovník termínů v rámci WG13 si myslí, že mu to zasahuje do jeho práce. Na zasedání slíbil, že vytvoří pilotní vzorek slovníku do zasedání ISO v Seattlu. I po jednání jsem ho kontaktoval a snažil se vysvětlit, že ontologie má právě pomoci v hledání termínů a kontextově pojímaných definic, ale zdálo se mi, že to nechápe a nebo to nechce zvažovat.

Podzimní zasedání – CEN TC278


Podzimní zasedání CEN/TC278 se konalo v Londýně. Program zasedání byl následující:

- 11. září ... 45. zasedání vedoucích pracovních skupin (Convenors meeting)
- 12. září ... 51. plenární zasedání (Plenary meeting)

V rámci zasedání se projednávalo oživení skupiny WG7: skupinu představil p. Kees Wevers (Navtech). Jedním z cílů je do map přenést dynamická data, např. změny symbolů dopravních značek, které dělají různé autority. Je potřebné vyjmát tato data a vkládat je do aplikací.

Podrobná zpráva ze zasedání je přílohou zprávy předsedy TNK136.

Opět jsme prezentovali tvorbu ontologie, tentokrát již konkrétněji, na příkladu EFC. Dokument byl oficiálně distribuován jako N 2969 a je přílohou této zprávy. Předseda L. Egging, který si uvědomuje důležitost ontologie poslal po zasedání konvenorovi WG1 J. Eghdalovi dopis, se žádostí o spolupráci s námi. Jesper odpověděl dlouhým mailem diplomaticky, že to je pěkné, ale že na to nemají peníze ani čas, ale že nás budou podporovat. Zároveň poslal dokument N2799 s termíny a definicemi, které vytvořili ve WG1 a jehož úvodní list je na následujícím obrázku:

	DATE	REFERENCE NUMBER
	2012-11-19	CEN/TC 278/N 2799
SUPERSEDES DOCUMENT		
N/A		
CEN/TC 278 Road transport and traffic telematics		
<p>TITLE EFC terms central repository</p> <p>SOURCE Working Group 1</p> <p>NOTE(S) Please find enclosed document WG1 N1677 for distribution in CEN/TC278 and in ISO/TC204.</p> <p>As we have informed in the latest status reports to TC204 and TC278, WG1(WG5) has started work with harmonising the terms defined in our EFC-standards. The first step, now concluded, is a full inventory of all our terms (in the attached document). All in all 685 terms were identified. The next step is to start work on new common definitions of (EFC) terms to be used in upcoming versions of our standards.</p> <p>By providing this document to the members and WGs of TC204 and TC278, we hope that we can give an opportunity for the editors of standards in other WGs to take into account the already defined EFC-terms if such are used</p>		

Temínů je sice 685, ale jsou často duplicitně a nemají žádnou hierarchii. Celý dokument působí značně chaoticky. Na základě jeho, spíše negativní odpovědi, jsem mu poslal následující dopis:

Dear Jesper,

Thank you for your kind comments and enclosed documents.

I would like to say that my presentation was future oriented and it is result of research and the aim of it is not to complicate present status in the creation of dictionary. I understand that preparation of dictionary required great deal of work.

Nevertheless, current ITS community does not speak only of the need to ensure the interoperability of data. A growing problem is semantic interoperability – the people in our domain do not interpret meaning of a word (generally term) on the same way. The reason is that the specific term is not placed in the concrete position taking in account the parent (superior) term. It is known since Aristotle time that some organization of terms is needed.

Arrangement of terms in some skeleton, advanced form of it is known as ontology, is practically used in biology, medicine, palaeontology, for traffic sensors or in automotive industry. The fundamental advantage is twofold:

1. There is arrangement of terms in a form of logical hierarchy. Each term has superior term and thus its context (and definition) is fixed and clear.
2. Graphical view of skeleton enables to uncover missing terms and the ontology sequentially creates semantic model of given domain

We have elaborated transport dictionary which contains 22 000 word from different modes of traffic (book, see picture + internet application). The words and definitions were selected from a huge set of different standards. To my surprise, the dictionary is not too much used in praxis. I asked several people - the main problem is that there



are words arranged in alphabetical order and some of them are generic (ITS, vehicle, ...) and some are very detailed (some parameter in DSRC telegram). The people are not sure if the searched term will be found there or not.

Dear Jesper, our “project” is not actually financially supported but we continue freely three years project ZNALSYS (Knowledge Systems) which was fully covered by ministry of transport. We – it means I and some my Ph.D. students. They are very active and would like to show something practical. As you know, our Japanese colleagues working on the same topic and we share our results.

If you agree we can continue by following steps:

1. The basis for the field test will be your database of terms and definitions (it was enclosed to your mail)
2. We try to arrange these words according their meaning into hierarchical skeleton – it will be fully heuristic approach. We still not handle method for automatic hierarchical ordering, but these methods exist.
3. Missing terms at different levels will be added. We will use software for natural language processing.
4. You and your colleagues will check it and you will assess the potential benefit thus formed ontology

My deep belief is that ITS standardization community is most progressive group and that we can show to another groups that ontology and semantic model could help to understand better not only problems but this approach will also help to disseminate knowledge about standards.

Takto by tedy měl vypadat další postup tvorby ontologie pro WG1, viz dále.

Jarní zasedání - ISO TC204

Jarní zasedání ISO/TC204 se konalo v Seattle USA ve dnech 18-19. 4. 2013. Program zasedání byl následující:

- | | |
|-------|---|
| 19.4. | ... 4 th Cross cutting meeting |
| 20.4. | ... 39 th ChOD/Plenary meeting |

Byla zvolena nová organizace, tzv. Cross-Cutting Meeting², kterého se při minulých zasedáních účastnilo vždy kolem 50-70 lidí, byl zrušen pro malou efektivitu a místo toho se konaly paralelní sekce pro sdílení informací mezi několika vybranými WG. Já jsem se zúčastnil první sekce Data registry/data dictionary, kterou vedl Dick Schnacke, ale hlavní slovo měl Richard Bossom (WG1) a to hlavně kvůli našim aktivitám ve znalostních systémech. Dále potom sekce vedené H.J.Schadem „Message sets“ reflektující aktivity WG18 kooperativní vozidla.

Developing of a Data Registry/Data Dictionary (WG1)

Účastnil jsem se této sekce ve čtvrtek od 9-10:30 h, přítomno bylo okolo 20 lidí. Moderoval D. Schnacke (USA). R. Bossom, konvenor WG1/ISO nejprve mluvil o projektu CHIRD orientovaném na slovníky, který ale nebyl přijat Evropskou komisí. Nově uvedl projekt „Steward“ předpokládající databázi na webu, nesmí se opakovat slova a měla by se automaticky aplikovat syntaktické vazby.

² proti tomu jsem písemně protestoval, protože na Cross-cutting každý z gestorů podrobně předkládal zprávu o činnosti a byla to vlastně jediná možnost, kdy bylo možné získat podrobnosti z činnosti jednotlivých skupin. Stejně tak to chodí na zasedáních CEN a tyto schůzky jsou mnohdy důležitější, než vlastní plenární schůze, které jsou velmi formální ... nicméně moje a dalších kolegů snaha byla převálcována

Většinou mluvil jenom o hardware a o výhodách užívání toho systému. Dále uváděl dokumenty ISO, ETSI TS 102 894-2, IS 14817 (víc jsem nezachytil), které se slovníky zabývají. Potěšující bylo, že zmínil, že případnou ontologii by dělala univerzita v Praze a Yokohamě. Jeho kolega potom předvedl web orientovaný databázový systém, popisující každé slovo v různých detailech, vč. sémantiky. Po jednání jsem se snažil Richarda vyzpovídat, jak jsou daleko, tvrdil, že předvedená databáze je jen ukázka mající jen několik slov.

V diskusi jsem opět upozornil, že databáze je čistě horizontální a není hierarchická. Slova různé „síly“ jsou stále na stejné úrovni, což je problém takovýchto databází. Ontologie tento problém omezuje.

Na závěr bylo dohodnuto předložení návrhu rezoluce, že se 18 měsíců bude pracovat na projektu databázové podpory datového registru. Někteří kolegové správně namítali, jak se ale pozná, že projekt byl úspěšný a k něčemu vedl.

Závěr po zasedání: Po jednání jsem široce diskutoval o možnostech co dál s prof. Kawashima (KWA) s následujícím výsledkem, jehož závěry byly vysvětleny, snad chápajícím, R. Bossomovi:

- (1) JPN a CR strana zpracují vzorovou ontologii, např. pro EFC. Není na to žádný projekt, takže to bude každá strana financovat ze svého. Je to ale prestižní záležitost. K tomuto, na místě, připravil KWA rukou psaný zápis, kam jsem doplnil následující:
 - o obě strany vyjdou ze stejných podkladů, PŘI připraví základní schéma³, které projednáme s p. Bossomem
 - o strany si nebudou konkurovat, ale budou se doplňovat
 - o vzhledem k tomu, že PŘI nebude na podzim v Kobe⁴ (a pokud tam nebude nikdo z ČR), prof. Yamaguchi případně představí i naši prezentaci.

První verze, resp. kostra ontologie by měla být do podzimního zasedání TNK 136 v Kobe.

- (2) Snažil jsem se přesvědčit, aby nám JPN pomohli s vyšší formou projektu STANDARD, kdy by DB našich extraktů byla převedena do angličtiny, byla vytvořena ontologie a vyhledávací web mechanismus. Představoval jsem si finanční podporu od japonské strany. Prof. KWA argumentoval, že po diskusi s B. Williamsem se domnívají, že naše extrakty nejsou legální z hlediska autorských práv. Doporučená naše TNK akce:
 - o Silmos vybere několik typických extraktů a zkontroluje, zda nechybí odkazy na zdroje a necháme je přeložit (nebo můžeme vzít již přeložené, ale nutně zkontrolovat)
 - o PŘI následně pošle návrh výše uvedeného projektu a vybrané extrakty vedoucímu TC278 (Lex Eggin), se kterým o projektu mluvil, aby nechal posoudit v CEN⁵ legálnost.

Totální závěr: zdá se, že je právě nyní doba, kdy můžeme zúročit oba projekty STANDARD i ZNALSYS. Pokud neučiníme kroky směřující ke konkrétnímu výstupu, je nutné stopnout další prezentace a přesvědčování, protože by to bylo opakování známého.

³ což jsem udělal v psané formě ještě po jednání, kopii dostal KWA a RB

⁴ podklady jsme poslali a dokument byl představen, zatím jsem jenom v písemném kontaktu s KWA, který je poměrně skoupý na slova, takže nevím, jak byly naše úvahy přijaty (listopad 2013)

⁵ v době psaní dokumentu máme stanovisko k legálnosti z ÚNMZ, další akce na úrovni CEN jsme nedělali

Podzimní zasedání - ISO TC204

Podzimní zasedání ISO/TC204 se konalo v japonském KOBE a bylo časově v koordinaci se světovým ITS kongresem v Tokyu. Zasedání jsem se z osobních důvodů neúčastnil, ale prezentace připravená pro toto zasedání byl přednesena prof Yamaguchi z Yokohamské univerzity.

3. Znalostní systémy – další rozvoj

Česká republika a naše komise je stále propagátorem implementace znalostních systémů hlavně proto, aby se znalosti obsažené ve standardech šířily do praxe. Úvahy postupují dvěma směry:

- I. Znalostní systém založený na projektu STANDARD
- II. Znalostní systém využívající ontologie pro daný obor (na bázi projektu ZNALSYS)

Protože byli členové komisí CEN i ISO při prezentacích patrně mateni názvem Standard byl pro zahraničí formálně změněn název na ESSENCE, který vyjadřuje, že se z normy vytvoří jakási esence informací, kterou představuje námi používaný Extrakt. Také byl vytvořen inovovaný popis tohoto projektu, který je v příloze tohoto dokumentu.

Vytvoření ontologie se ukázalo větším problémem, než bylo předpokládáno. To je patrně také důvod, proč zatím ontologie pro obor ITS nikde ve světě nevznikla. V době psaní tohoto dokumentu se konala již 9. schůzka neformálního týmu tvořeného dvěma pracovníky Fakulty dopravní – dr. Faltuse a Ing. Plánkové a dále externím specialistou/programátorem dr. Pavelkou. Ze schůzek jsou pořizovány záznamy. Pro tyto aktivity není zatím žádná finanční podpora a tým pracuje na dobrovolné bázi.

V dalším textu jsou komentovány dílčí výsledky.

3.1 Projekt ESSENCE/STANDARD

Projekt, který je u nás doveden do praxe se zdál uchopitelný již v roce 2011, kdy jsme ho představili na jarním zasedání ISO v Praze. Patrně díky tomu, že nebyl dobře pochopen jeho princip a hlavně proto, že nebyly prostředky na jeho anglickou mutaci však nebyla větší odezva.

Projekt byl představen i v Bruselu v sídle ERTICO v roce 2012. Letos v září jsem osobně projednal s panem Mayerem, prezidentem ERTICO, při jeho návštěvě fakultní konference v Praze. Předal jsem mu svázaný popis projektu ESSENCE.

Zatím se nedaří tento projekt prosadit na mezinárodní scéně. Patrně by bylo nutné zpracovat demonstrační část a snažit se to prosadit přes nějakou silnou firmu. Možná by v podpoře mohlo pomoci SDT Česká republika, neboť je členem mezinárodní sítě ITS, viz závěry.

Problém zpochybnění díky autorským právům byl snad překonán stanoviskem ÚNMZ

3.2 Ontologie a znalostní systémy

Největší zájem o znalostní problematiku a tedy i ontologii projevují Japonci, se kterými jsem ve stálém kontaktu. Pro zasedání v KOBE, kterého jsem se nemohl zúčastnit, jsem připravil prezentaci, kterou doplnil a přednesl prof Yamaguchi.

Původní očekávání bylo, že budeme vytvářet znalostní systém, který bude na základě ontologie vyhledávat v dokumentech konkrétní znalostní jednotky. Díky tomu, že není jakákoliv finanční podpora, jsme zatím tuto myšlenku opustili a soustředujeme se na tvorbu ontologie ITS, resp. vzorově na ontologii pro segment EFC. Důvody jsou následující:

1. V ITS komunitě se prosazuje názor, že není jediným problémem datová interoperabilita, ale že je vážným nedostatkem sémantická interoperabilita. Jinak řečeno si nerozumí odborníci, kteří dělají EFC například s lidmi zabývajícími se kooperativními systémy. Již dva roky se intenzivně mluví o nutnosti termíny a definice nějak uspořádat.
2. Problém vzájemného neporozumnění se řeší tvorbou slovníků s definicí významu termínu. Jeden slovník vznikl ve skupině EFC – 650 termínů, často duplicitních a navíc s různými významy. Další slovník řeší WG1/ISO. To je internetový vyhledávač termínů a definicí. Zásadní nevýhodou těchto slovníků je:
 - i. Termíny nejsou řazeny v kontextu, tj. jeden termín nemůže mít víc definicí, podle toho, kam je zařazen. To toto uspořádání neumožňuje a je nutno mít „univerzální“ definici, platnou pro jakýkoli kontext;
 - ii. Termíny jsou řazeny horizontálně dle abecedy – najdeme zde „silné“ (důležité) termíny, ale i nějaké detailní pojmy, protože není hierarchie ani jiná organizace, nelze usuzovat, co ještě ve slovníku bude a co ne.

3.3 Co se udělalo v letošním roce

Systematicky pracovala neformální skupina na půdě Fakulty dopravní, která se v pravidelných čtrnáctidenních termínech schází a snaží se postupovat ve dvou směrech:

- A. Matematické dolování četnosti slov a jejich spojení z dokumentů, což by mělo dát obraz o významu slova (četnost = význam?) v dokumentu = semiautomatická tvorba ontologie.
- B. Ze známé DB slov expertně vytvářet hierarchické nebo síťové uspořádání slov.

3.3.1 Semiautomatická tvorba ontologie

Následující kroky charakterizují proces, který byl dosud proveden (a není zcela dokončen):

1. Vyhledání standardů, které budou analyzovány metodou NLP (Natural Language Processing) a jejich převedení do XML;
 - a. Byl zvolen standard ISO/TS 17575 „*EFC – Application interface definition for autonomous systems*“, který má 4 části o 186 stranách
 - b. Standard byl složitě ručně převeden do XML, na formu vhodnou pro strojového čtení (dokumenty jsou v pdf). Příklad „přeloženého“ textu je v následujícím obrázku:

```

file:///D:/Aktuell/Normy%20program/Zdroj%201.htm - Původní zdroj
Soubor  Úpravy  Formát
5345 </p> </td>
5346 <td width=80 valign=top style='width:60.0pt;border-top:none;border-left:none; border-
bottom:solid windowtext 1.0pt;border-right:solid windowtext 1.0pt; padding:0in 0in 0in 0in'>
5347 <p class=MsoNormal align=center style='margin-top:3.0pt;margin-right:0in; margin-
bottom:3.0pt;margin-left:0in;text-align:center'>
5348 <span>optional
5349 </span>
5350 </p> </td>
5351 <td width=93 valign=top style='width:70.0pt;border-top:none;border-left:none; border-
bottom:solid windowtext 1.0pt;border-right:solid windowtext 1.0pt; padding:0in 0in 0in 0in'>
5352 <p class=MsoNormal align=center style='margin-top:3.0pt;margin-right:0in; margin-
bottom:3.0pt;margin-left:0in;text-align:center'>
5353 <span>Yes/No
5354 </span>
5355 </p> </td>
5356 </tr>
5357 <tr>
5358 <td width=340 valign=top style='width:255.15pt;border:solid windowtext 1.0pt; border-
top:none;padding:0in 0in 0in 0in'>
5359 <p class=MsoNormal style='margin-top:3.0pt;margin-right:0in;margin-bottom: 3.0pt;margin-
left:2.85pt'>
5360 <span>serviceProviderContract
5361 </span>
5362 </p> </td>
5363 <td width=80 valign=top style='width:60.0pt;border-top:none;border-left:none; border-
bottom:solid windowtext 1.0pt;border-right:solid windowtext 1.0pt; padding:0in 0in 0in 0in'>
5364 <p class=MsoNormal align=center style='margin-top:3.0pt;margin-right:0in; margin-
bottom:3.0pt;margin-left:0in;text-align:center'>
5365 <span>mandatory
5366 </span>
5367 </p> </td>
5368 <td width=93 valign=top style='width:70.0pt;border-top:none;border-left:none; border-

```

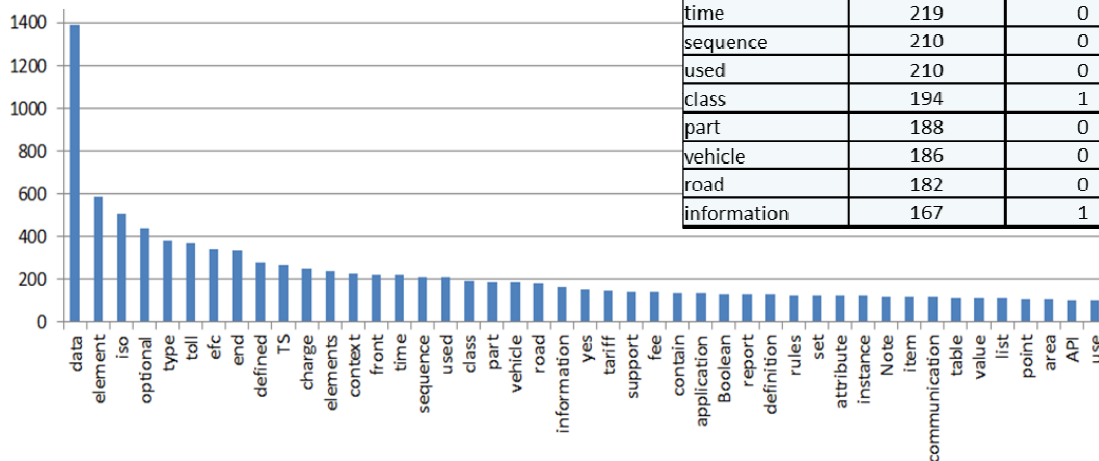
c. Takto zpracované čtyři části dokumentu měly 43 802 řádek, 49 88 slov a 9 716 vět!!!

d. Počítačovou filtrací byla vyloučena krátká slova, diakritická znaménka apod.

2. Softwarem pro zpracování přirozeného jazyka byla analyzována četnost slov. Pro tento účel byl

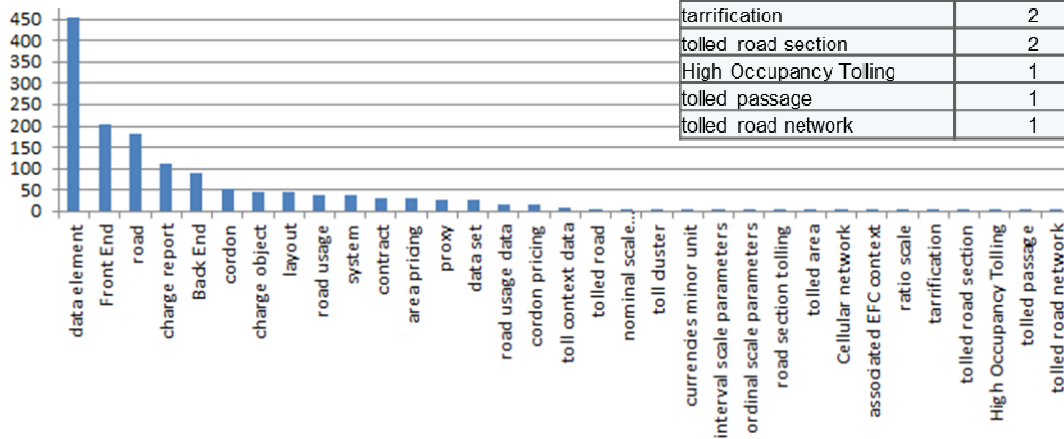
vyvinut program v MATLAB. Komerční programové produkty pro Natural Language Processing se pohybují v řádu stů tisíc korun. Výsledkem byla databáze frekvencí slov pro slova vyskytující se v nadpisech a v textu analyzované normy, viz následující tabulka a grafická interpretace:

The word	Frequency of words in the source text	Frequency of words in headlines
data	1394	38
element	589	12
iso	509	0
optional	436	0
type	383	22
toll	369	0
efc	344	14
end	336	1
defined	276	0
TS	265	0
charge	249	3
elements	240	5
context	226	3
front	224	1
time	219	0
sequence	210	0
used	210	0
class	194	1
part	188	0
vehicle	186	0
road	182	0
information	167	1

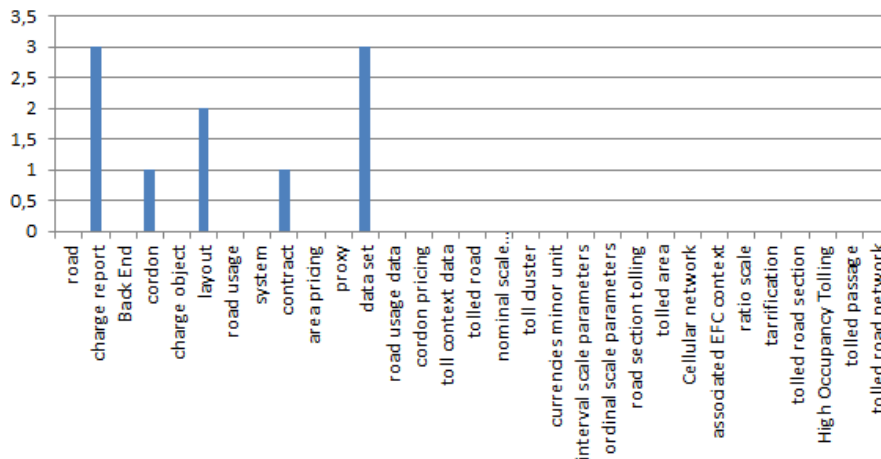


3. V dalším kroku bylo provedeno srovnání termů, které vybrali experti do DB pro WG1 s tím, jak často se vyskytují tato slova v textu. Následující tabulka a graf ukazují, že zde není žádná velká shoda mezi tím, jaké (klíčové) slovo vybere expert a tím co vybere program:

Terms defined by experts (chap. 3)	Frequency of terms in the source text
data element	455
Front End	202
road	182
charge report	111
Back End	91
cordon	54
charge object	46
layout	45
road usage	38
system	37
contract	31
area pricing	29
proxy	27
data set	26
road usage data	17
cordon pricing	16
toll context data	8
tolled road	5
nominal scale parameters	4
toll cluster	4
currencies minor unit	3
interval scale parameters	3
ordinal scale parameters	3
road section tolling	3
tolled area	3
Cellular network	2
associated EFC context	2
ratio scale	2
tarrification	2
tolled road section	2
High Occupancy Tolling	1
tolled passage	1
tolled road network	1



Z uvedeného je patrné, že není velká shoda v tom, jaká slova považují experti za důležitá a jaká slova se vyskytují v textech. Stejně tak je velmi malá shoda mezi výběrem slov experty (slova na vodorovné ose) a tím, jak se tato slova vyskytují v nadpisech:



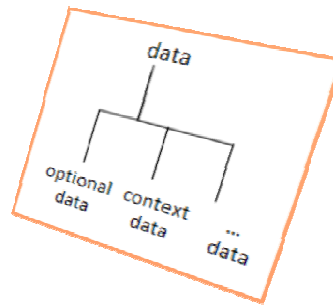
4. V dalším kroku se strojově hledaly asociace dvojic a trojic slov. Výsledkem je rozsáhlá databáze 8 328 dvojic a trojic slovních spojení. Příklad DB je v tabulce napravo:

Díky strojovému zpracování se vyskytuje i celá řada spojení, která nedávají smysl nebo nejsou důležitá a je nutno je ručně třídit, což vyžaduje další ruční práci. Například od řádku 1259 je 26 dvojic slov začínajících na „data“:

data defined
data defining
data element accountStatus
data element addToAccount
data element applicableTimeClasses
data element basicFeePerChargeUnit
data element chargeDistance
data element chargeUnit
data element cordonLayout
data element localVehicleClasses
data element mileage
data element obeStatus
data element regimeld
data element reloadAccount
data element tollCharger
data element transactionCounter

Selected phrases	Frequency in the source document
data element	455
data type	322
iso TS	263
front end	202
data elements	187
charge report	111
Back end	91
context data	79
optional data	76
efc context	74
efc attribute	70
optional data element	60
toll charger	60
service provider	57
FC application	52
fee collection	52
time class	51
charge object	46
electronic fee	46
electronic fee collection	46
PICS proforma	45
toll regime	44
vehicle class	44
tariff class	43
toll context	40
road usage	38
autonomous systems	36
efc contexts	34
roaming rules	32
efc attributes	30
area pricing	29
data types	29
efc regimes	27
iso IEC	27

5. Na základě takto vytvořených dvojic či trojic je možné vytvářet základní prvky hierarchie analogicky s diagramem tříd v UML, jak ukazuje obrázek:



V principu se jedná o hledání vztahu rodič-dítě. V této etapě prozatím skončilo bádání našeho týmu. Další postup:

6. Z dílčích uspořádání termů musí expert vytvořit celou kostru a to vycházejí z kontextu slov a jejich pozice v kostře.
- nevádí různé definice slov vycházejících z různého kontextu;
 - vadí různé definice dané různými přístupy autorů: možná eliminace a sjednocení definic pomocí WordNet nebyla zkoumána, ale je v záloze.
7. I při takto vytvořené kostře se nejedná o ontologii, která by měla zahrnovat i relace mezi slovy.

Dílčí závěr k semiautomatické tvorbě

- A. Přes automatizované vyhledávání slov se jedná se o časově náročný proces, jehož rozsah lze zatím jen odhadovat. Hlavním problémem je:
- a. Převod Word nebo pdf do XML;
 - b. Nutnost manuálně třídit a uspořádávat dvojice a trojice slov.
- B. Na druhé straně metody přirozeného zpracování řeči NLP, zvláště sofistikované programy, mohou pokrýt oblast standardů rozhodně lépe, než pouhé definování termů různými experty.
- C. Ontologie vytvořená tímto způsobem je více autentická než ontologie vytvářená ručně na základě slovníku termínů.

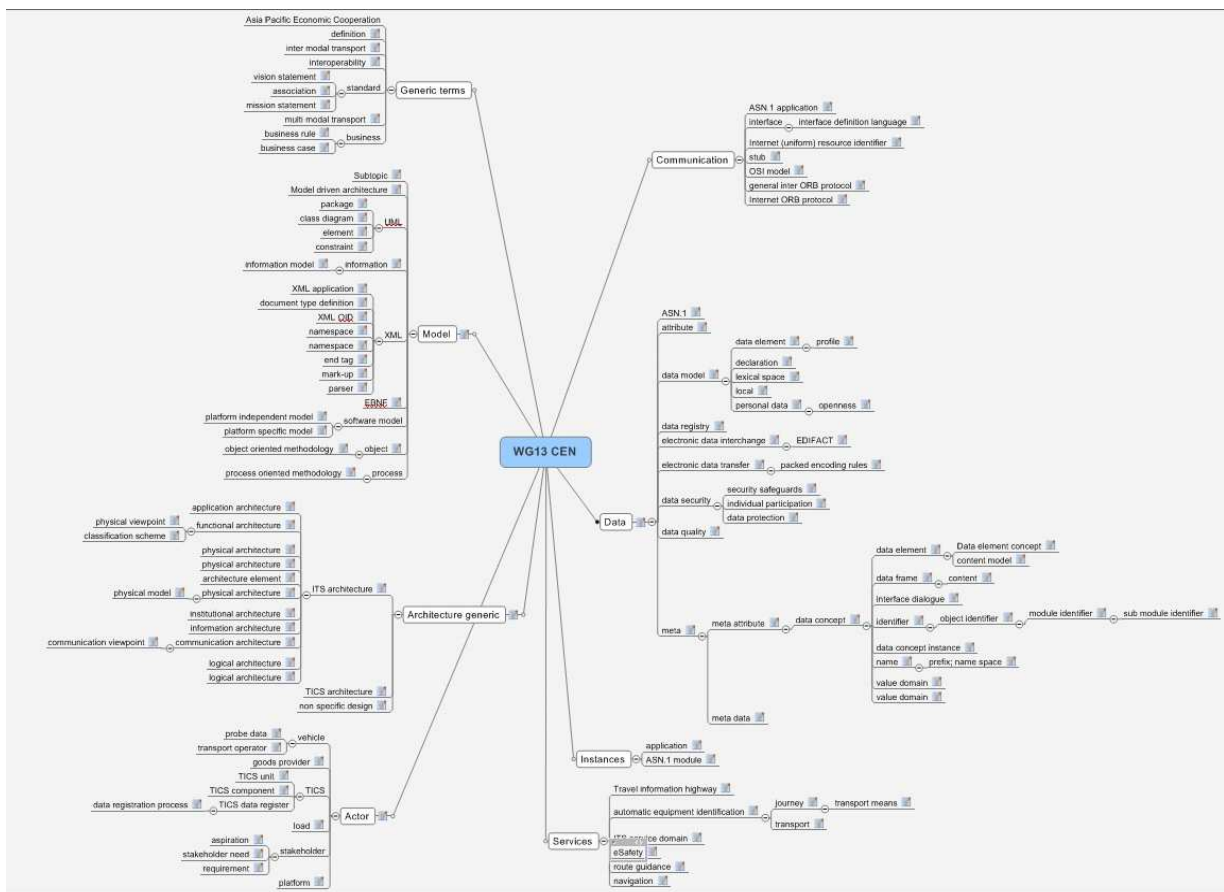
3.3.2 Ontologie vytvářená expertem z databáze termínů

Za základ byla brána databáze termínů, které byly „minovány“ z norem experty komise TNK 136. Termíny byly experty zpracovány, tříděny, byly odstraněny duplicity apod.

Počty takto zpracovaných termínů je možné vidět v příložené tabulce.

Working group	Number of terms
WG1	356
WG2	269
WG3	442
WG5	103
WG8	76
WG9	50
WG12	228
WG13	206
WG14 ISO	177
WG15	139
WG16 ISO	330
WG17 ISO	50
summary	2429

Pro ruční uspořádání termínů do hierarchie či struktury byl zvolen nástroj Mind Map Organizer a jako pilotní test 120 termínů z WG1 (celkově 356) a cca 100 termínů z WG13. Pro ilustraci je část takto vytvořené struktury na následujícím obrázku.

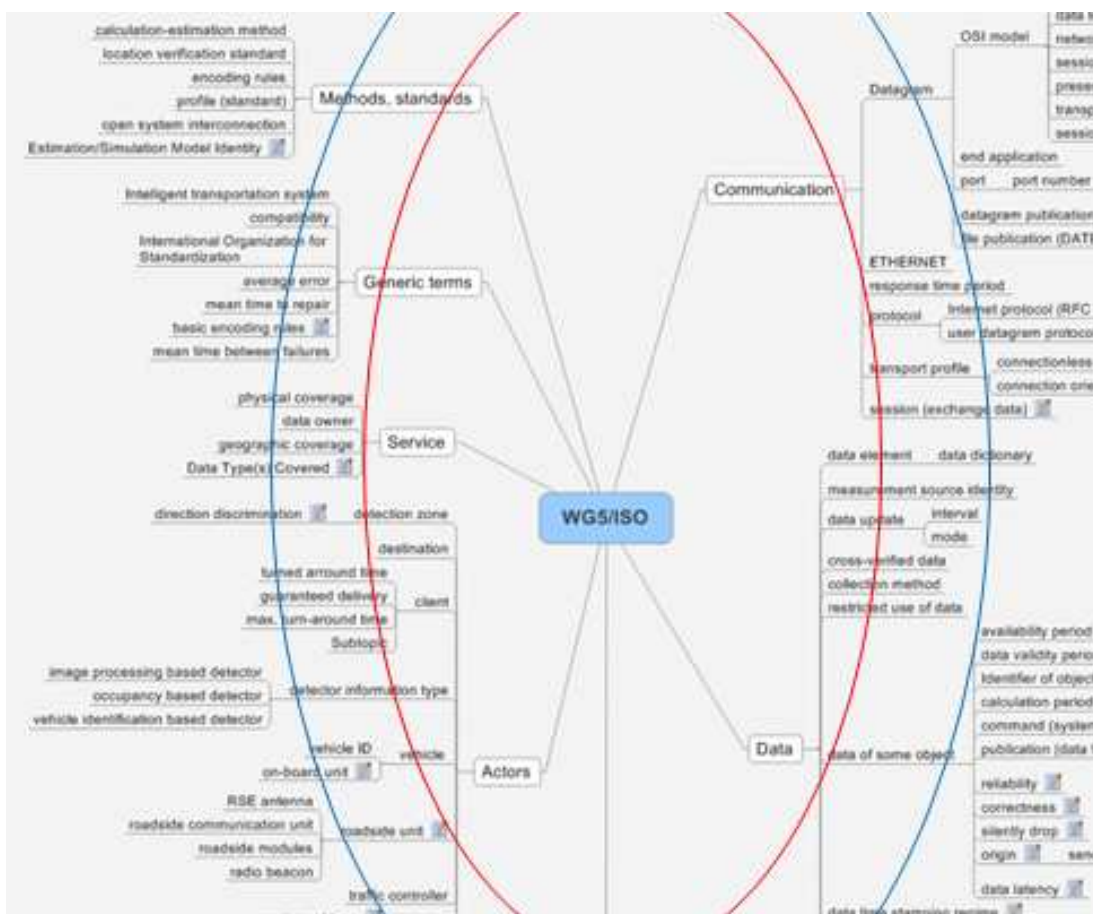


Je zřejmé, že do nějakých 100, možná až do 200 termínů je grafická interpretace celkem přehledná. Při více termínech by bylo nutné již sledovat konkrétní větev či větve. Velkou výhodou je, že po najetí kurzorem na jednotlivý výraz se otevře jeho definice, pokud je zadaná.

Bylo také možné vysledovat, že se zde uplatňuje individuální pojetí a přístup autora. To by bylo možné omezit, pokud by se nejvyšší třídy hierarchie pro všechny třídy ITS definovaly předem. Tento postup by bylo možné uplatnit alespoň do druhé hierarchické úrovně, neboť lze definovat:

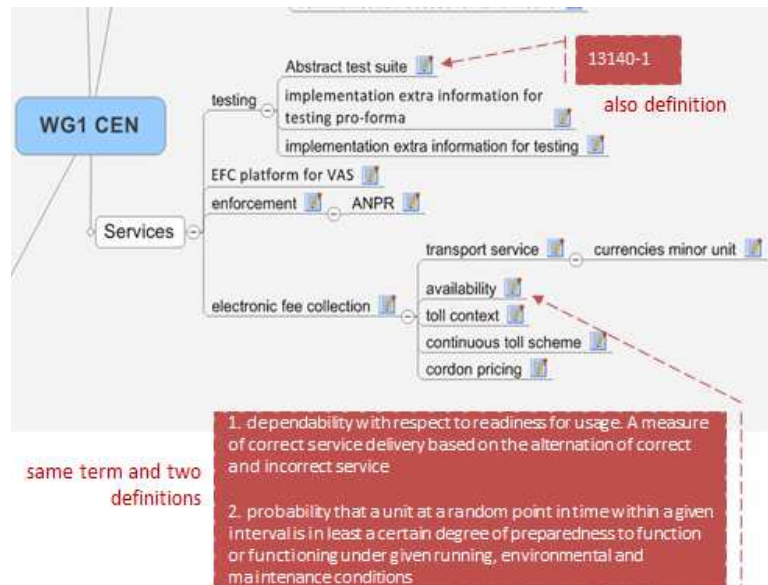
1. úroveň: obsahuje zcela obecné výrazy jako: aktor, služba, standard, ...
2. úroveň: obsahuje slova „vyšší“ úrovně, všobecně používaná, jako: vozidlo, palubní jednotka, data, komunikace, ...

V následujícím obrázku první a druhou úroveň přibližně vyznačují barvy:



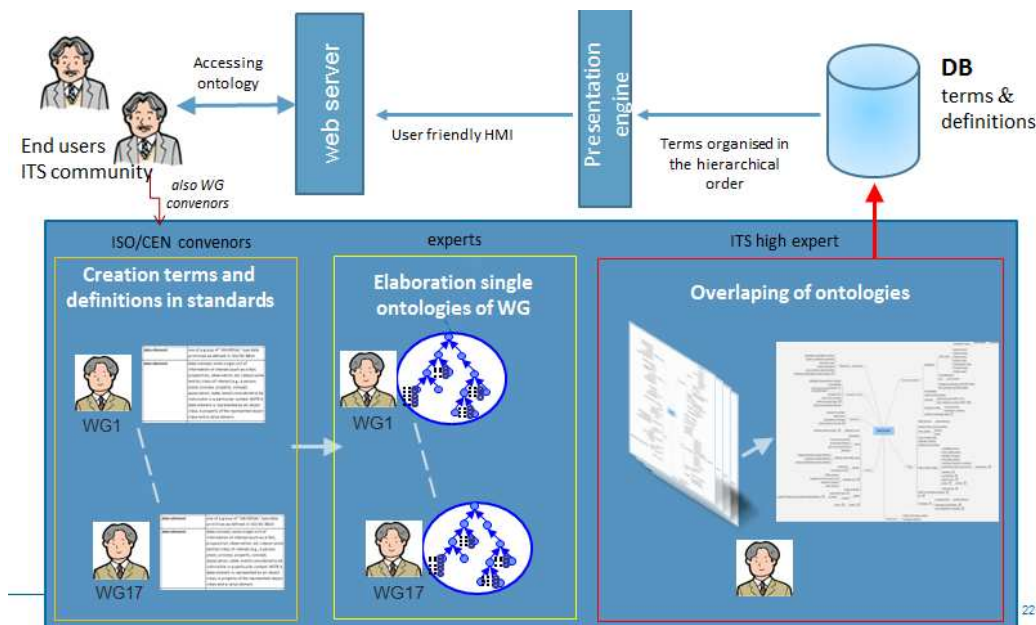
Díky využití programů Mind Map se odhalí i řada nepřesností či problémů, které by se v databázích hledaly hůře. Obrázek vedle ukazuje příklad, že pro slovo „availability“ jsou dvě definice a každá z nich vyplývá z jiného kontextu. Autor, který zařadil do databáze „Abstract test suite“ naopak žádnou definici nevytvořil a jen udal odkaz na normu.

Takovýchto problémů se díky grafickému zobrazení termů ve struktuře odhalí velké množství.



Dílčí závěr k expertní tvorbě ontologií ze slovníků

Na jednání CEN v Londýně byl nabídnut postup, který by nekomplikoval stávající činnost konvenorů a přinesl by poměrně rychle výsledky. Celý postup, spočívající v tom, že by se v každé skupině vytvořila dílčí ontologie a to dle předem definovaného schématu vyšších úrovní je naznačen v dalším obrázku:



Z dílčích ontologií by sjednocením a filtrací vznikla celková ontologie. Je patrné, že největší práci by měl expert, který by jednotlivě vzniklé ontologie sjednocoval.

3.4 Znalostní systémy – závěr

3.4.1 Projekt EXTRACT/ESSENCE

1. Projekt EXTRAKT je světově unikátní projekt, který by měl a mohl pomoci k šíření standardů do praxe.
2. Projekt byl opakovaně popularizován v obou komisích, které naše myšlenky publikovaly i v oficiálních dokumentech, ale odezva není prakticky žádná
 - přetrvává jakýsi problém s autorskými právy, byť jsme předkládali rozklad ÚNMZ a patrně bude nutné vyvinout další naše úsilí k eliminaci tohoto myšlení.
3. Problém je, že se z naší strany vždy jedná jen o jednoduchou demonstraci, které patrně členové komise dobře nerozumí. Pokud by se mělo jít dál, je nutné vytvořit pěkné internetové prostředí vyhledávače a hodně extraktů přeložit a tím demonstrovat využití. Cestou by mohlo být zadání diplomové práce pro web rozhraní a překlad extraktů z peněz sponzorů.
4. Do projektu je nutné zapojit SDT ČR (úloha pro prof. Příbyla a Ing. Večerku). Ideálně sehnat mezinárodního sponzora – našlo by to jako aplikaci na smart telefonu?
5. Projekt byl publikován na vědecké konferenci v Polsku s textem následně ve vydavatelství Springer, na světovém ITS kongresu ve Vídni a u nás v časopisu Silniční obzor. Bylo jednáno s Australany, prezidentem ERTICO apod. Hledat další kanály na popularizaci!!!

3.4.2 Ontologie

1. O nutnosti vytvořit ontologii pro ITS se stále mluví. Z několika prezentací a následných jednání se jeví, že členové komise ISO/CEN nejsou na tuto úroveň myšlení připraveni, že je problematika pro ně složitá a pouze z prezentací neuchopitelná. Zjistil jsem, že třeba nerozumí významu pojmu „ontologie“.
2. Z Japonska jsem dostal významný dokument od Toma Kurihara „*Report of study group on the feasibility of preparing an ontology for the SC7 domain and standards*“, který podporuje naše úsilí a mluví o nutnosti vytvořit skupinu pro tvorbu ontologie v IT světě. Ta byla také vytvořena a patrně pracuje.
3. Komunikace předsedy TC278 směřující k pilotnímu projektu ve WG1, vedla k tomu, že její konvenor sdělil, že nejsou fondy na tuto práci.
4. Ontologie byla představena na několika zasedáních, ale také na stejné povrchní úrovni a dál touto cestou není možné jít. Je nutné zpracovat konkrétní výsledek:
 - i. Jediné řešení je pokračovat v započatých pracích s doktorandy a pokusit se to uplatnit v projektech Technologické agentury nebo ve spolupráci s Japonci, kteří o to mají také velký zájem.
 - ii. Tato neformální skupina se schází každých čtrnáct dní v pozdně odpoledních hodinách a bylo by dobré, kdyby mohl být začleněn někdo z TNK.

4. Restrukturalizace činnosti ISO TC204

Již několik let se diskutuje o nutnosti restrukturalizovat činnost TC, protože se práce zdá neefektivní. Důvodem je, že schůzky se v podstatě, až na jisté výjimky, konají pouze dvakrát ročně, přičemž na zasedání třeba nepřijedou (nemohou přijet) rozhodující odborníci. Téměř vůbec neexistuje křížová korelace mezi pracovními skupinami.

Na zasedání v Moskvě přednesl vedoucí USA delegace p. Dick Schnacke prezentaci, jak zvýšit efektivitu práce. Jeho úvaha je založená na vytvoření sub-skupin, tvořenými různými WG, které by měly svého koordinátora. Vytvářela by se tedy ještě jedna úroveň řízení.

Proti tomuto návrhu naše komise protestovala, stejně jako celá řada dalších zemí. Vedoucí delegace USA D. Schnacke v Seattle poděkoval za podnětné návrhy ze 17 zemí, představil dokumenty shrnujícími náměty a konstatoval, že návrh na vytvoření další úrovně do struktury neprošel. Na druhé straně je ale nutné se zamyslet nad dalšími, moderními, způsoby komunikace mezi experty.

Naší komisi může těšit, že náš rozbor byl hodnocen vysoko, neboť se zamýšlel nad celou organizací a byl doplněn řadou schémat.

5. Závěrem

Tato výroční zpráva popisuje výstupy, kterých bylo dosaženo v roce 2013 v rámci činnosti TNK 136. Z obsahu zprávy je vidět, že tato komise je velmi aktivní a zpracovává mnoho dokumentů a norem týkajících se oblasti dopravní telematiky. Členská základna aktivně sleduje a spolupracuje na utváření evropských a světových norem, snahou je zapojit se co nejvíce do světové procesu ISO. Stále platí, že naše země je jediná z východních evropských zemí, která má aktivní zastoupení v obou komisích. Potěšením je, že se v téměř každé prezentaci konvenorů objevují jména našich gestorů.

V této zprávě není věnováno velké místo detailnějším popisu činnosti – kolik norem bylo připomínkováno, kolik jich bylo transformováno do naší standardizační soustavy apod. K tomu slouží velmi podrobné čtvrtletní zprávy předkládané Centrem technické normalizace, společností SILMOS.

Významné je, že se daří již patnáct let držet pohromadě skupinu špičkových odborníků v ITS, kteří se nejenom scházejí, ale i aktivně pracují. Podobnou myšlenku vyjádřil na zasedání v září auditor Ing. Krebs, který byl překvapen množstvím a rozsahem práce v naší komisi. Proto bych chtěl velmi poděkovat členům komise, kteří vytváří jednu společnou rodinu, snažící se podporovat standardizaci ITS.

Důležité je, že se nedílnou filosofií práce naší komise stalo i šíření standardů do odborné i laické veřejnosti. Tato myšlenka byla diskutována i na zasedáních CEN a ISO a zdá se, že projektem STANDARD a ZNALSYS jsme se dostali o velký kus dopředu a jsme v tomto oboru považováni za průkopníky. To je i důvodem, že je této otázce věnován velký prostor i v této zprávě.

V současné době je nutné najít bussiness model, udělat demo a projekty více propagovat a to včetně zahraničí. Hlavním úkolem je pracovat tak, aby jej bylo možné transformovat na evropský či světový projekt (přímý zájem projevil Japonsko a USA).

Na závěr bych rád touto cestou poděkoval za velmi aktivní pomoc při organizování práce celé komise Ing. Večerkovi a Mgr. Svorové.

Velmi významnou činnost v podpoře našich aktivit představuje zástupce ČNI Ing. Křivka, který se podílí na organizování celé naší činnosti. Věříme, že dosavadní spolupráce s ČNI bude pokračovat stejně úspěšně.

Za MD děkujeme garantce Ing. Křištofikové za podporu dopravní telematiky a tím i činnosti komise.



Prof. Ing. Pavel Příbyl, CSc.

předseda TNK 136

Praha, 22. listopadu 2013

Přílohy

1. N2969 „Presentation slides Ontology in ITS“
2. Basic decription of project ESSENCE