
PROJEKT HUMBOLDT: CESTA K IMPLEMENTACI ESDI

Otakar Čerba¹, Petr Horák², Karel Charvát³, Jan Ježek⁴, Štěpán Kafka⁵, Tomáš Mildorf⁶, Pavel Sedlář⁷

^{1,4,6} Katedra matematiky, Fakulta aplikovaných věd, Západočeská univerzita v Plzni,
Univerzitní 22,
306 14, Plzeň, Česká republika

¹ota.cerba@seznam.cz

⁴h.jezek@centrum.cz

⁶mildorf@centrum.cz

^{2,7} Help forest, spol. s r. o., Slovanská 21,
787 01, Šumperk, Česká republika
helpforest@helpforest.cz

³ České centrum pro vědu a společnost, Radlicka 28,
150 00, Praha 5, Česká republika
charvat@ccss.cz

⁵ Help service - remote sensing s.r.o., Černoleská 1600,
256 01, Benešov u Prahy, Česká republika
kafka@email.cz

Abstrakt.

Projekt Humboldt je výzkumný projekt Evropské unie, jenž přispívá k implementaci evropské infrastruktury prostorových dat (European Spatial Data Infrastructure, ESDI). ESDI integruje všechny rozličnosti prostorových dat dostupných z mnoha evropských organizací. Hlavní úkolem projektu je harmonizace prostorových dat a příslušných metadat poskytovaných různými organizacemi evropských států. Harmonizace umožní přístup k datům pomocí síťových služeb ve formě, která umožní jejich sdílení s jinými daty ESDI koherentní cestou pomocí použití společné množiny souřadnicových referenčních systémů, klasifikačních systémů, aplikačních schémat atd.

Výsledkem projektu by měla být metodika (sekvence harmonizačních procesů) umožňující jednotlivým subjektům sdílet a publikovat data bez ohledu na jejich původ. Kromě metodiky by měly být v rámci projektu vytvořeny příslušné nástroje integrované do softwarového prostředí, které zjednoduší a zautomatizují harmonizační procesy. Dojde k vytvoření softwarových nástrojů a procesů, které budou ukázkou proveditelnosti a výhod infrastruktury pro prostorové informace v Evropě plánované iniciativou INSPIRE (Infrastructure for Spatial Information in Europe).

Klíčová slova: Humboldt, ESDI, harmonizace, INSPIRE, prostorová data, územní plánování, lesní hospodářství.

Abstract.

Project Humboldt is the research design of European Union, which goes towards implementation of European Spatial Data Infrastructure (ESDI). ESDI integrates all diversities of spatial data, which are accessible from many european organizations. The main task of this project is harmonisation of spatial data and metadata provided by different institutions of european countries. Harmonisation makes possible an access to data through web services in the form, which enables its sharing with other data by using general sets of coordinate systems, classification system, application schemas etc.

The main goal of Humboldt project would be the methodology (sequence of harmonisation procedures) to make an access to data possible to all subjects. Except the methodology there would be created different tools integrated to software framework, which simplify and automate harmonisation processes.

Keywords: Humboldt, ESDI, harmonisation, INSPIRE, spatial data, urban planning, forest management.

1 Úvod – problematika různorodosti informací, její příčiny a důsledky

Současný svět je postavený před řadu naléhavých otázek z oblastí ochrany životního prostředí, vyšší bezpečnosti, udržitelného a spravedlivého rozvoje. Pro řešení těchto otázek jsou nezbytné informace, které umožní efektivní cílení opatření a rozhodování.

V současnosti již není hlavním problémem kvantita (množství) informací, ale především jejich kvalita. Ta spočívá v dostatku metainformací (zdroj informací, jejich aktualizace, vhodnost pro daný účel, možnosti získání informací apod.) a možnosti kombinace dat s jinými datovými sadami a jejich implementace do používaných softwarových i hardwarových řešení bez ohledu na zdroj dat, státní hranice, měřítko dat, používané technologie, datové modely, organizační strukturu, legislativní pravidla, uživatelské potřeby, typ dat, formát dat apod.

Potřeba harmonizace dat je dána nutností spolupráce poskytovatelů a zpracovatelů geoprostorových dat na národní i internacionální úrovni. V důsledku postupné globalizace ekonomiky se vzájemná provázanost týká stále většího množství ekonomických subjektů i vlastních datových sad. V současnosti je aktuální především vazba dat na úrovni Evropské Unie. Důležitá je provázanost na úrovni státní správy i na úrovni tematických vrstev. Přičemž je potřeba si uvědomit, že většina dat je vytvářena a spravována na lokální úrovni, poté jsou data poskytována na další vyšší úrovně.

Heterogenitu dat kromě rozdílných potřeb a požadavků jednotlivých uživatelů zapříčiňuje odlišné

- uplatňování norem a standardů (INSPIRE, ISO 19100),
- technologické a technické standardy (např. datové formáty),
- protokoly,
- aplikační schémata,
- datové a metadatové modely,
- rozsah témat,
- katalogy prvků,
- nomenklatury, taxonomie, klasifikace,
- terminologie, thesaury, ontologie,
- referenční systémy,
- souřadnicové systémy a používané jednotky,

- měřítko,
- organizační a legislativní aspekty,
- licenční politiky,
- způsoby prezentace (legenda, kartografický styl apod.),
- procesní funkce,
- kvality dat,
- rozšíření,
- používané jazyky,
- kulturní prostředí. [Čer2007b]

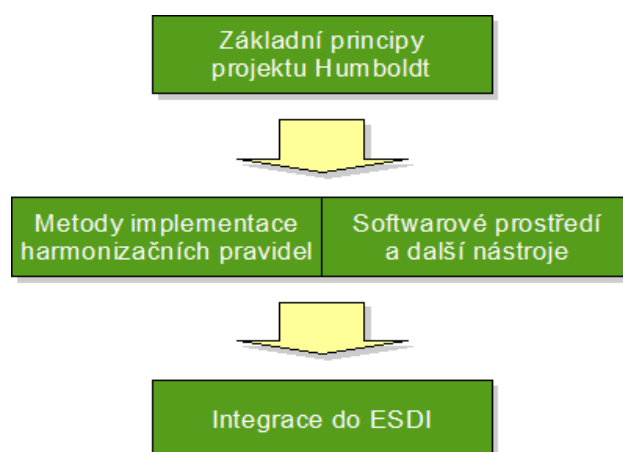
Důsledkem heterogenity a obtížného sdílení datových sad je především ekonomická a organizační náročnost spojená s tvorbou, udržováním a aktualizací podobných nebo téměř shodných dat. Se stále rostoucím počtem datových sad, které mají sice podobný, ale nikoli zcela shodný, obsah se stává stále obtížnější vyhledávání relevantních informací.

2 Popis projektu Humboldt

Projekt Humboldt je výzkumný projekt Evropské unie, jenž přispívá k implementaci evropské infrastruktury prostorových dat (European Spatial Data Infrastructure, ESDI). ESDI integruje všechny rozličnosti prostorových dat dostupných z mnoha evropských organizací. Hlavním úkolem projektu je harmonizace prostorových dat a příslušných metadat poskytovaných různými organizacemi evropských států [Öst2006]. Harmonizace dat poskytuje přístup k datům pomocí síťových služeb ve formě, která umožní jejich sdílení s jinými daty INSPIRE (Infrastructure for Spatial Information in Europe) koherentní cestou pomocí použití společné množiny souřadnicových referenčních systémů, klasifikačních systémů, aplikačních schémat apod. v ESDI [Dat2006].

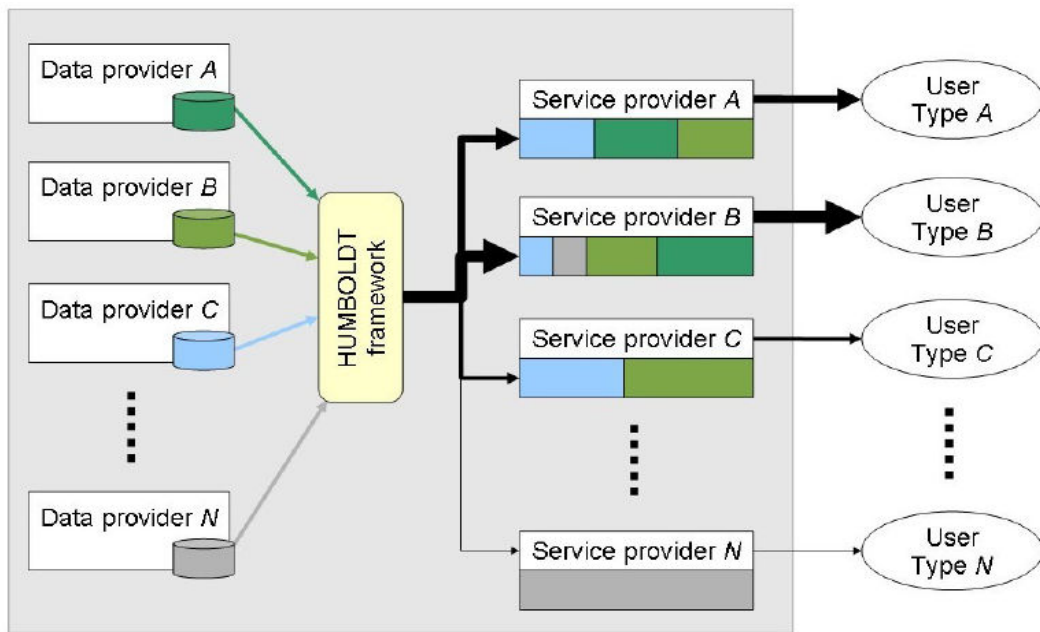
Výsledkem projektu by měla být

- metodika (sekvence harmonizačních procesů) umožňující jednotlivým subjektům sdílet a publikovat data bez ohledu na jejich původ [Čer2007b],
- softwarové prostředí a nástroje umožňující realizaci jednotlivých harmonizačních procesů.



Obr. 1. Schéma projektu Humboldt (podle [Čer2007b]).

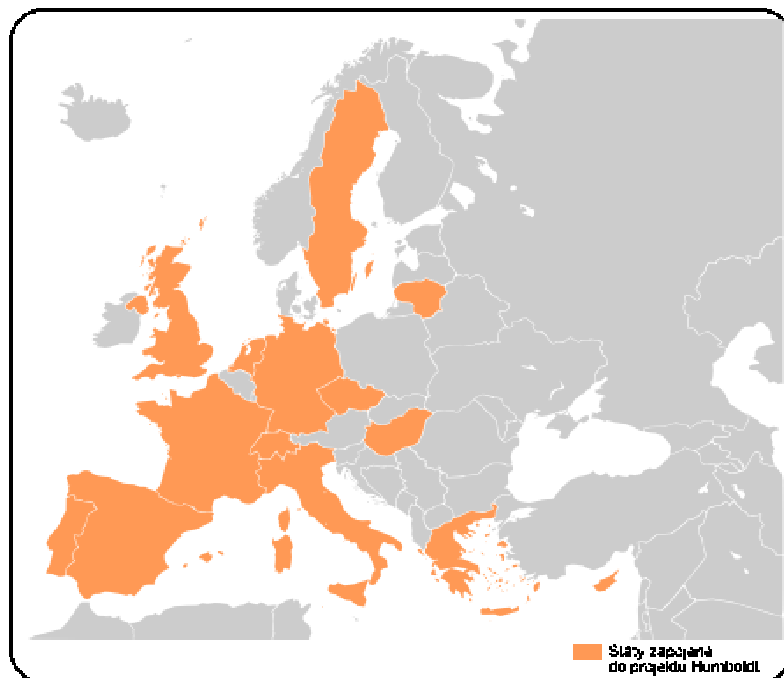
V rámci projektu dojde k vytvoření softwarových nástrojů a procesů integrovaných do softwarového prostředí, které budou ukázkou proveditelnosti a výhod infrastruktury pro prostorové informace v Evropě plánované iniciativou INSPIRE a zároveň zjednoduší a zautomatizují harmonizační procesy.



Obr. 2. Prostředí projektu Humboldt.

V projektu Humboldt je zapojeno 27 partnerů pocházejících ze 14 evropských států (13 členů Evropské Unie a Švýcarsko). Vedení projektu zajišťuje Fraunhofer Institute for Computer Graphics (Darmstadt, Německo). Mezi partnery se nachází

- komerční firmy (např. LogicaCMG),
- národní mapovací organizace (např. Institut Geographique National / French National Geographic Institute),
- výzkumná centra (např. Sveriges Meteorologiska och Hydrologiska Institut / Swedish Meteorological and Hydrological Institute),
- zástupci univerzitní sféry (např. Technische Universiteit Delft / Delft University of Technology).



Obr. 3. Státy zapojené do projektu Humboldt.

Českou republiku zastupují tři subjekty:

- Help service – remote sensing spol. s.r.o. (HSRS),
- Ústav pro hospodářskou úpravu lesů (ÚHUL),
- Intergraph, Czech republic.

Projekt Humboldt je rozčleněn do 12 pracovních skupin (Workpackages, WP) a 8 scénářů (Humboldt Scenarios, HS).

Work packages (WP):

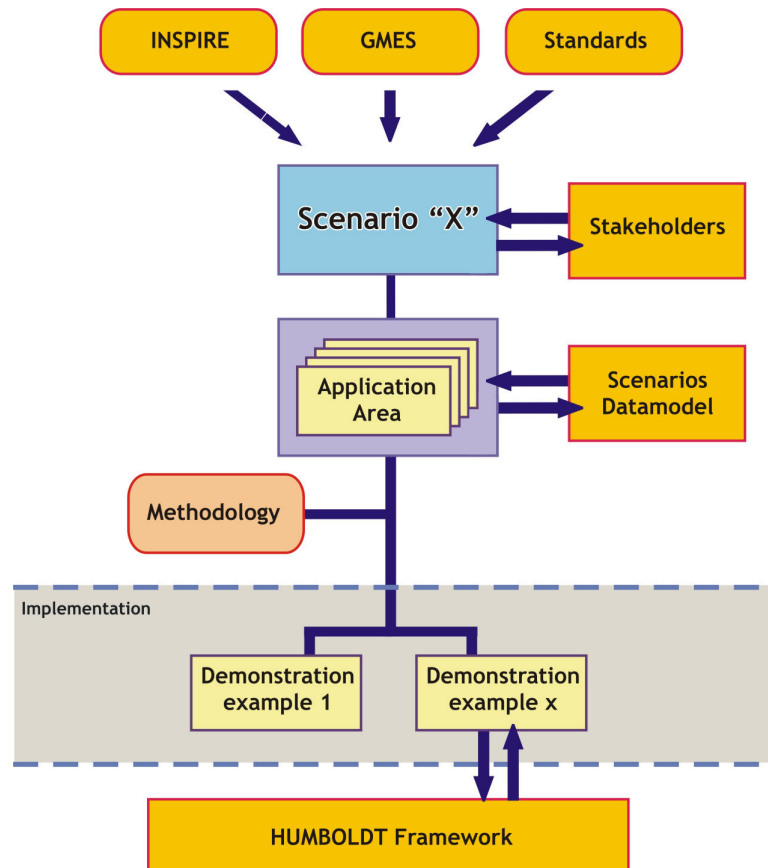
1. Administration,
2. Cost & Process Analysis,
3. State of the Art,
4. Development Process Specification,
5. Framework Interface, Models and Architecture,
6. Component Validation Process,
7. Harmonized Data Profiles,
8. Framework Development,
9. Scenario Applications,
10. Evaluation,
11. Dissemination and Exploitation,
12. Training.

Humboldt Scenarios:

- HS Forest,
- HS Urban Planning,
- HS Border Security,
- HS Protected Areas,
- HS ERiskA (European Risk Atlas),
- HS Water,
- HS Ocean,
- HS Galileo.

3 Scénáře HS Forest a HS Urban Planning

Subjekty zastupující Českou republiku (konkrétně HSRS a ÚHUL) především participují na aktivitách spojených s problematikou prostorových dat v lesnictví a územním plánování. Konkrétně se jedná o scénáře HS Forest a HS Urban Planning. Úkolem scénářů je testovat a také prezentovat vytvořené prostředí a jednotlivé komponenty v reálných podmínkách na konkrétních případech. Scénáře se většinou zabývají oblastmi, které souvisí s problematikou ochrany životního prostředí a trvale udržitelného rozvoje. [Čer2007b]



Obr. 4. Model vývoje scénářů.

Scénáře jako jednotlivé složky WP9 jsou součástí tzv. kritické cesty projektu Humboldt. Kritická cesta se skládá ze tří základních prvků WP2 (analytické činnosti), WP5 (architektura systému) a WP9 (scénáře jako konkrétní vazba mezi uživateli a tvůrci systému), na jejichž úspěšné realizaci je závislý úspěch celého projektu.

Základním prvkem scénářů jsou tzv. Use Cases, které reprezentují vybrané případy využívání prostorových dat při činnostech typických pro jednotlivé scénáře. Scénáře HS Forest a HS Urban Planning zahrnují následující Use Cases:

HS Forest

- Forest typology – data for field survey of forest typology for RPFD
- EAFRD - Support european countryside
- Transportation network database maintenance and update
- Interconnection between data layers in the project “Virtual Forest Authority“ in cooperation of Germany and Czech Republic

HS Urban Planning

- Providing of preliminary information – territorial planning information (based on the building law)
- Providing data of territorial planning bases and territorial planning documentation to a planner of the territorial plan (TP) – (in terms of building law)
- Providing data of territorial planning bases and territorial planning documentation to a planner for a territory plan (without a TP officer)
- Data providing from territorial planning bases and territorial planning documentation (Agregated A13)
- Providing data of structural ecological

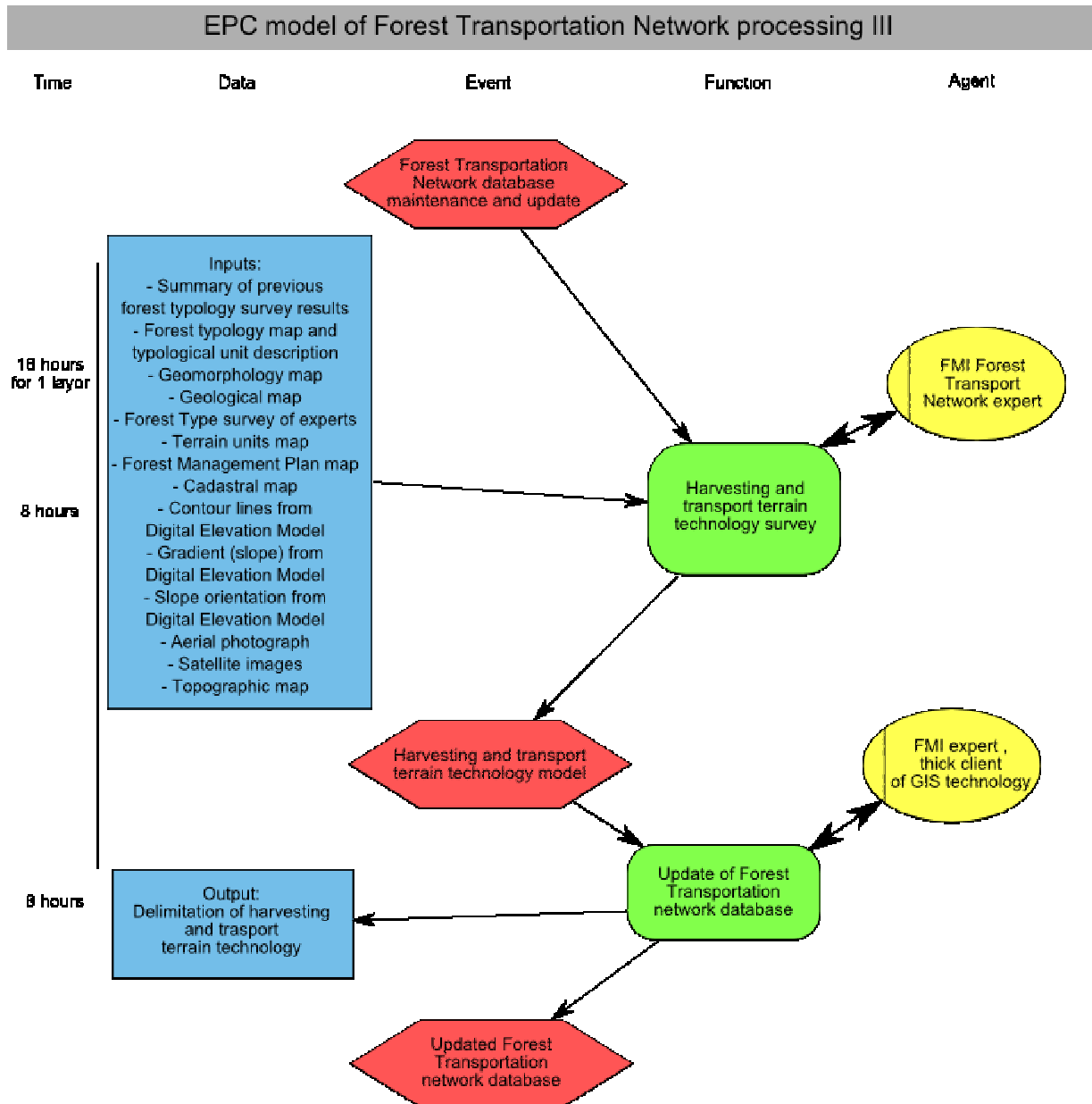
- network parts – SSES (spatial system of ecological stability) to applicant – expert (state administration bodies, designers, expertly erudite users)
- Providing data of structural ecological network parts – SSES (spatial system of ecological stability) to applicant – nonexpert (general public, expertly less erudite users)
- Equalization data by mains administrator to accredited authority

V prvním roce projektu byla většina aktivit zaměřena na analytické činnosti, na které bude navazovat vývoj prvních prototypů. Základem analýzy uživatelských požadavků je popis každého scénáře, který se v případě projektu Humboldt skládá ze tří částí:

1. Popis ve formě User Stories (Požadavky systému formulované běžným jazykem uživatele, jednou či dvěma větami, které by měly charakterizovat očekávání uživatelů od systému),
2. UML (Unified Modeling Language) Activity diagram (postupné zobrazení jednotlivých kroků celého procesu),
3. UML Use Case diagram (vztahy mezi uživateli a Use Cases uvnitř systému).

Systém User Stories se skládá z následujících kroků, které přesně popisují jednotlivé elementy celého systému (definice uživatelů a typických uživatelů reprezentující širší skupinu zákazníků, popis scénáře pro každého uživatele – individuální User Stories, výběr uživatelských požadavků na základě User Stories).

Dalším typem grafického znázornění, který byl využitý pro procesní a nákladové analýzy, jsou EPC diagramy (Event-driven Process Chain). EPC diagramy se skládají se sekvence vzájemně navazujících konkrétních událostí a funkcí (s důrazem na riziková místa procesu a na místa s potencionální redukcí nákladů), přičemž každá událost představuje spouštěcí mechanismus příslušné funkce. Tato základní kostra EPC diagramů je doplněna vstupy, výstupy a účastníky jednotlivých částí procesu a časovou osou ukazující časovou náročnost jednotlivých funkcí.



Obr. 5. EPC model procesů aktualizace transportní sítě.

Další kroky obou zpracovávaných scénářů budou směřovat především k dokončení analytických podkladů (procesní analýzy, analýzy uživatelských požadavků, analýza redukčních procesů apod.). V případě HS Urban Planning budou do zatím obecně pojatých Use Cases implementováno řešení následujících procesů, které negativně ovlivňují využívání a především sdílení prostorových dat v územním plánování:

1. Širší využívání metadatových záznamů, které mohou sloužit jako podklad pro pasporty pro územní plánování (nástroje pro generování pasportu prostřednictvím metadat byly implementovány do systému MICKA). Problematika metadat je základem datové harmonizace, proto se jedná o univerzální téma, které bude vyžadovat synergie v rámci všech scénářů projektu Humboldt.
2. Dalším problémem, který ztěžuje využívání prostorových dat v rámci různých subjektů je rozdílnost datových modelů využívaných v GIS a CAD. Ve scénáři Urban Planning bude aplikován softwarový prostředek MAGDA pro transformaci dat CAD a GIS.

3. Třetí konkrétní aktivitou bude propojení scénářů HS Forest a HS Urban Planning za účelem sdílení dat důležitých pro krizový management. Součástí této aktivity bude také využívání sensorových měření a návrh kooperace s dalšími scénáři.

4 Závěr

Projekt Humboldt je pojmenován podle německého přírodovědce Alexandra von Humboldta. Tento významný muž evropské vědy již v první polovině na přelomu 18. a 19. století navrhoval širokou integraci znalostí.

Projekt Humboldt integruje odlišnosti prostorových dat poskytovaných různými evropskými organizacemi. Hlavním přínosem projektu Humboldt je umožnit poskytovatelům dat dokumentovat, publikovat a harmonizovat jejich zdroje prostorových dat. [Čer2007a]

K využívání prostorových dat v mnohem větším rozsahu a mnohem efektivněji by měla napomoci celá řada procesů. Vyšší úrovní integrace prostorových dat brání řada nedostatků, jejichž odstranění by mělo být cílem projektu Humboldt:

- Chybí zdroje interoperabilních dat.
- Dochází k duplicitnímu uchování dat (vysoké náklady na aktualizaci a správu dat).
- Nevyjasněná pravidla pro sdílení dat.
- Heterogenita dat.
- Neexistence, neúplnost či nejednotnost metadat na všech úrovních (národní, regionální a lokální).
- Chybějící číselníky metadat.
- Číselníky nejsou zahrnuty do norem a standardů.
- Neexistence jednotného datového modelu.

Reference

- [Čer2007a] ČERBA, O.; FRYML, J.; POSPÍŠIL, M.; PODLENA, R.; CHARVÁT, K. HS Forest Scenario of Humboldt - Way to INSPIRE Implementation in Forestry. In Informační systémy v zemědělství a lesnictví. Praha: Česká zemědělská univerzita v Praze, 2007.
- [Čer2007b] ČERBA, O.; CHARVÁT, K.; KAFKA, Š.; MILDORF, T.; SEDLÁŘ, P. Harmonizace dat územního plánování. In Územní plánování a GIS. Bítov: Českou asociací pro geoinformace, 2007.
- [Dat2006] Data Harmonisation Requirements [online]. RISE: Reference Information Specifications for Europe, 2006. [cit. 2007-10-04]. URL: <https://intranet.esdi-humboldt.eu>.
- [Öst2006] ÖSTMAN, A.; OSTREIKA, A.; LOIDOLD, M.; CHARVÁT, K. A2.1-D1 Process Analysis (Interim Version) [online]. Humboldt Consortium, 2006. [cit. 2007-10-04]. URL: <http://intranet.esdi-humboldt.eu>.