

VELKÁ DATA V ZEMĚDĚLSTVÍ

Karel Charvát, Pavel Hájek, Michal Kepka,
Martina Kepka Vichrová, Dmitrij Kozhukh,
Jiří Kvapil, Hana Kubíčková, Petr Uhlíř,
Jaroslav Šmejkal, Antonín Kubíček,
Heřman Šnevajs, Stein Runar Bergheim,
Karel Charvát, jr.,
Šárka Horáková

2021



VELKÁ DATA V ZEMĚDĚLSTVÍ

Karel Charvát, Pavel Hájek, Michal Kepka,
Martina Kepka Vichrová, Dmitrij Kozhukh,
Jiří Kvapil, Hana Kubíčková, Petr Uhlíř,
Jaroslav Šmejkal, Antonín Kubíček,
Heřman Šnevajs, Stein Runar Bergheim,
Karel Charvát, jr.,
Šárka Horáková



OBSAH

1. ÚVOD	7
2. ZNALOSTNÍ ZEMĚDĚLSTVÍ	9
2.1. EXTERNÍ VLIVY	9
2.2. CÍLE UDRŽITELNÉHO ROZVOJE A ZELENÁ DOHODA PRO EVROPU	10
2.2.1. CÍLE UDRŽITELNÉHO ROZVOJE (SUSTAINABLE DEVELOPMENT GOALS)	10
2.2.2. ZELENÁ DOHODA PRO EVROPU (ZELENÝ ÚDĚL - GREEN DEAL)	12
2.3. ZNALOSTNÍ PYRAMIDA	14
2.4. CHUDÁK FARMÁŘ	15
2.5. DESTINATION EARTH: DIGITÁLNÍ DVOJČE ZEMĚ	17
3. CO JSOU BIG DATA (VELKÁ DATA)	18
3.1. ZÁKLADNÍ DEFINICE VELKÝCH DAT	18
3.2. KDO SKUTEČNĚ POTŘEBUJE DATA	20
3.3. ZPĚT KE ZNALOSTNÍ PYRAMIDĚ	21
4. OTEVŘENÁ A SDÍLENÁ DATA	22
5. CO JE TO INTEROPERABILITA	24
5.1. DATOVÉ TYPY A DATOVÉ FORMÁTY	24
5.2. STANDARDIZAČNÍ ORGANIZACE A INICIATIVY RELEVANTNÍ PRO ZEMĚDĚLSTVÍ	25
5.3. STANDARDIZACE GEOPROSTOROVÝCH INFORMACÍ	27
5.4. PROPOJENÍ S NÁRODNÍMI NORMAMI	30
5.5. OTEVŘENÁ PROPOJENÁ DATA – LINKED OPEN DATA	31
6. BEZPEČNOST DAT A JEJICH OCHRANA	32
6.1. BLOCKCHAIN	33
7. NÁVRATNOST NÁKLADŮ (ROI NEBOLI RETURN OF INVESTMENT)	34
8. CO JSOU METADATA A OPRAVDU JE POTŘEBUJEME?	36
8.1. TECHNOLOGIE PRO METADATA	36
8.2. (NE)VIDITELNÁ METADATA	37
9. DATOVÉ ZDROJE A JEJICH HARMONIZACE	40
9.1. FOODIE DATA MODEL	40
9.2. OTEVŘENÁ DATA VYUŽITÍ PŮDY – OPEN LAND USE	43
9.2.1. POSTUP TVORBY OLU V2:	48
9.3. CHYTRÉ BODY ZÁJMU – SMART POINT OF INTEREST	51
9.3.1. DATOVÝ MODEL SPOI	51
9.3.2. SPARQL ENDPOINT	52
9.3.3. MAPOVÝ KLIENT	52
9.4. PROPOJENÍ DATOVÝCH SAD	53
10. DÁLKOVÝ PRŮZKUM ZEMĚ	54
10.1. LETECKÉ A UAV PLATFORMY	56
11. SENZOROVÉ TECHNOLOGIE	57
12. OBČANSKÁ VĚDA A CROWDSOURCING	58

13. VIZUALIZACE VELKÝCH DAT	59
13.1. PŘÍKLADY POKROČILÝCH VIZUALIZACÍ	59
13.1.1. KOMPLEXNÍ INTEGROVANÁ VIZUALIZACE DAT	59
13.1.2. VIZUALIZACE LINKED OPEN DATA	61
13.1.3. VYSVĚTLUJÍCÍ VIZUALIZACE (EXPLANATORY VISUALISATION)	61
13.1.4. KOLABORATIVNÍ VIZUALIZACE	61
14. ANALÝZY, MODELOVÁNÍ A UMĚLÁ INTELIGENCE	64
14.1. DATA MINING	64
14.2. STATISTICKÉ METODY	65
14.3. UMĚLÁ INTELIGENCE	65
15. ŽIJEME V DOBĚ CLOUDU	66
15.1. DATOVÉ A ZPRACOVATELSKÉ PLATFORMY	67
15.2. PLATFORMY PRO VYUŽÍVÁNÍ DAT	67
15.3. STANDARDY PRO EARTH OBSERVATION CLOUD ARCHITECTURE	68
15.4. ŘEŠÍ DNEŠNÍ CLOUD SITUACI "CHUDÁKA ZEMĚDĚLCE"?	68
16. PŘÍKLADY ŘEŠENÍ Z PROJEKTŮ HORIZONT 2020	70
16.1. DATABIO A NÁSTROJE VELKÝCH DAT	70
16.1.1. ZEMĚDĚLSKÉ PILOTY	70
16.1.2. DATABIO ARCHITEKTURA	71
16.1.3. ČESKÁ PILOTNÍ APLIKACE	73
16.2. EUXDAT	74
16.2.1. VÝSTUPY PROJEKTU EUXDAT	75
16.2.2. PILOTY A SCÉNÁŘE	75
16.3. DEMETER	77
16.3.1. ČESKÝ PILOT	78
16.4. INNOVAR	79
17. DIGITÁLNÍ INOVAČNÍ HUBY - BUDOUCNOST, JAK TRANSFORMOVAT VELKÁ DATA DO ZNALOSTÍ	81
17.1. DEMONSTRAČNÍ APLIKACE S DATY DRUŽIC SENTINEL-1 A SENTINEL-2	82
18. NĚCO NA ZÁVĚR	87
SEZNAM OBRÁZKŮ	88
SEZNAM ZKRATEK	90
ZDROJE	92



1. ÚVOD

Cílem této publikace je, co nejsrozumitelnější formou poskytnout informaci o tom, co to jsou velká data a kde nám mohou pomoci v zemědělství a regionálním rozvoji. Jedná se o ne úplně jednoduchý úkol a v některých kapitolách je těžké se zcela vyhnout techničtějším detailům. Naší snahou je poskytnout ucelený pohled na problematiku velkých dat (včetně prezentace vybraných souborů velkých dat) a pokusit se nastínit, jak mohou velká data v zemědělství pomáhat. Cílem je i poukázat na to, že velká data představují komplexní problém, který obvykle přesahuje schopnosti jednotlivého farmáře. Pro jejich využití je třeba zajistit účinnou spolupráci expertů z mnoha oborů a samozřejmě účinnou spolupráci s vědou a výzkumem. Publikace se opírá především o výsledky velkého evropského projektu **DataBio (Data-Driven Bioeconomy)**, využívá výstupy i dalších evropských projektů, nejen ze vzdálenější minulosti, ale i ze současnosti.

Projekt **DataBio** byl jedním z prvních dvou klíčových projektů tzv. "lighthouse projects (majáku)" v oblasti velkých dat, které byly financovány v rámci programu Horizont 2020 (další dva byly uděleny o rok později) probíhající od ledna 2017 do prosince 2019. DataBio byl projekt zaměřený na využití velkých dat v zemědělství, lesnictví a rybnářství. Cílem bylo navrhnout obecnou architekturu a implementovat pilotní řešení. Projekt DataBio přinesl nejen řadu výsledků, které jsou využitelné v současné době, ale nastínil i scénáře budoucího rozvoje velkých dat v zemědělství. Naše kniha tudíž ukazuje i zárodky mnoha technologií, které budou pravděpodobně hrát významnou roli v datovém prostoru pro zemědělství. Popisuje i vztah těchto technologií s řešením problémů, které před naším evropským, ale i globálním zemědělstvím vyvstávají. Jedním z těchto problémů je sdílení dat. Schopností sdílet data způsobem, který zachovává nejen soukromí osob, ale i obchodní tajemství (obojí je nezbytným předpokladem pro řádné fungování a správu datového prostoru) je jedním ze základních požadavků na budoucí systémy. V současné době narůstá důležitost měření ze senzorů, využití dat Dálkového průzkumu Země a ekonomických a statistických dat.

Pro sdílení dat je nutným předpokladem existence standardů, které jsou nezbytné k tomu, aby spolu mohly různé

systémy spolupracovat a aby farmáři mohli sdílet data s dalšími účastníky jako jsou poradci, organizace zemědělských služeb, veřejná správa, potravinářský průmysl a další. [1]

V publikaci se opíráme i o řadu dalších evropských výzkumných projektů, ukončených nebo stále probíhajících. Jedná se o tyto projekty:

- **VOICE** – Projekt Evropské kosmické agentury (2003-2004) - Cíle projektu THE VOICE byly:
 - identifikace společných a obecných technologických prvků nezbytných pro vytvoření prostředí pro spolupráci, které podporuje řetězce přidaných hodnot v zemědělství
 - identifikace společných mechanismů rozhraní pro data, aplikace a vytváření služeb, včetně "výměnných formátů" pro interakci a využívání dostupných zdrojů
 - implementace prototypů, tj. implementace kolaborativních prostředí s reprezentativními aplikacemi a službami pro řetězce přidaných hodnot v zemědělství.
- **Ami4For** – Projekt Evropské kosmické agentury (2005-2008) - Hlavním cílem projektu bylo vytvořit novou koncepci mobilní inteligence (AMI) pro řízení lesnictví, vinařství a zemědělství, která by zahrnovala mobilní komunikaci, nové metody navigace (GPS, EGNOS, GALILEO) a integraci prostorových informací včetně satelitního snímkování (SPOT, IKONOS, EROS, PROBA).
- **FUTUREFARM** – Integration of Farm Management Information Systems to support real-time management decisions and compliance of management standards (2008-2011). Cílem projektu FUTUREFARM byla identifikace současných a budoucích potřeb v oblasti správy dat, informací a znalostí v zemědělských podnicích, jakož i způsobu, jakým se tyto potřeby budou v budoucnu vyvíjet a který ovlivní systémy správy dat, informací a znalostí v zemědělských podnicích.
- **agriXchange** – A common data exchange system for agricultural systems (2009-2012) - Projekt agriXchange pracoval na vytvoření platformy pro výměnu dat v zemědělství v EU, která se skládala z technické infrastruktury a společenství odborníků z praxe a která vypracovala

referenční rámec pro interoperabilitu výměny dat v zemědělství a identifikovala hlavní výzvy pro harmonizaci výměny dat v zemědělství, což vyústilo ve strategický výzkumný program (SRA).

- **FOODIE** – Farm-Oriented Open Data in Europe (2014-2017) – Klíčovým bodem projektu FOODIE bylo vytvoření cloudové platformy, kde budou k dispozici prostorová a neprostorová data týkající se zemědělského sektoru pro skupiny zainteresovaných stran v zemědělsko-potravinářském odvětví. Základem projektu FOODIE byla standardizace těchto informací.
- **SDI4Apps** – Uptake of Open Geographic Information Through Innovative Services Based on Linked Data (2014-2017) – Hlavním cílem projektu SDI4Apps bylo propojit shora dolů řízený svět INSPIRE, Copernicus a GEOSS a zdola nahoru mobilní svět dobrovolných iniciativ a tisíců malých a středních podniků a jednotlivců, kteří vyvíjejí aplikace založené na zeměpisných informacích.
- **IoF2020** – Internet of Food and Farm 2020 (2017-2021) – Projekt IoF2020 byl zaměřen na urychlení zavádění internetu věcí pro zajištění dostatečného množství bezpečných a zdravých potravin a na posílení konkurenceschopnosti zemědělství a potravinových řetězců v Evropě.
- **EO4Agri** – Bringing together the Knowledge for Better Agriculture Monitoring (2018-2020) – Hlavním cílem projektu EO4AGRI bylo urychlit rozvoj evropských kapacit pro zlepšení operativního monitorování zemědělství od místní po globální úroveň na základě informací získaných z dat programu Copernicus prostřednictvím využívání souvisejících geoprostorových a socioekonomických informačních služeb.
- **EUXDAT** – European e-Infrastructure for Extreme Data Analytics in Sustainable Development (2017-2020) – projekt EUXDAT navrhl e-infrastrukturu, která se zabývá zemědělstvím, monitorováním půdy a energetickou účinností pro udržitelný rozvoj, jako způsob podpory plánovacích politik.
- **SmartAgriHubs** – Connecting the dots to unleash the innovation potential for digital transformation of the European agri-food sector (2018-2022) – Cílem SmartAgriHubs je urychlit digitální transformaci evropského zemědělsko-potravinářského sektoru.

- **DEMETER** – Building an Interoperable, Data-Driven, Innovative and Sustainable European Agri-Food Sector (2019-2023) – Projekt analyzuje údaje získané od široké škály subjektů (výrobních odvětví a systémů) s cílem poskytnout integrovaný interoperabilní datový model umožňující optimální řízení zdrojů v evropském zemědělsko-potravinářském odvětví.
- **Innovar** – Next generation variety testing for improved cropping on European farmland (2019-2024) – InnoVar vyvíjí novou generaci testování odrůd rostlin vytvářením nástrojů a modelů, které rozšíří současné postupy a využijí pokroky v genomice, fenomice, zobrazovacích technologiích a strojovém učení.
- **PoliRural** – Future Oriented Collaborative Policy Development for Rural Areas and People (2019-2022) – projekt poskytne srovnávací hodnocení výzev, kterým čelí venkovské oblasti a lidé, a opatření, která přijali regionální aktéři k jejich řešení. Spojí rozhodovací orgány, odborníky a obyvatele venkova a bude využívat pokročilé nástroje pro simulaci politik, aby lépe pochopili a řešili regionální výzvy a v konečném důsledku učinili venkovské oblasti a profese atraktivnějšími a vhodnějšími k životu pro starousedlíky, tak i pro nedávné nebo potenciální nově příchozí.
- **SIEUSOIL** – Sino-EU Soil Observatory for Intelligent Land Use Management (2019-2022) – projekt vyvíjí udržitelné a holistické postupy hospodaření s půdou založené na harmonizovaném informačním systému o půdě, který je vhodný pro různé klimatické a provozní podmínky v různých lokalitách EU a Číny.

Doufáme, že tato publikace poskytne zajímavý náhled do světa evropských projektů v oblasti informačních komunikačních technologií velkých dat v zemědělství a ukáže, kam dnes technologie směřuje. Cílem je, aby pomohla v orientaci v základních pojmech a napomohla pochopení vztahu jednotlivých technologií a zemědělství. Publikací chceme přispět k obecné informovanosti zemědělců, rádi bychom oslovili též veřejnou správu, vývojáře aplikací pro zemědělství, investory a případně i finanční sektor. Naše více než dvacetileté zkušenosti z evropského výzkumu ukazují, že pokud máme zajistit udržitelné zemědělství, nebude to možné bez spolupráce všech zúčastněných.



2. ZNALOSTNÍ ZEMĚDĚLSTVÍ

2.1. EXTERNÍ VLIVY

Zemědělství je důležité hospodářské odvětví produkující potraviny, zemědělsko-průmyslové suroviny a energii, poskytující environmentální služby prostřednictvím komplexního hospodaření s půdou, vodou, vzduchem a biologickou rozmanitostí. Ze zemědělství se tak v dnešní době stala komplexní činnost zahrnující mnoho subjektů a zúčastněných stran v řetězcích služeb a zemědělsko-potravinářských řetězcích, které produkují a dodávají potraviny a zemědělské komodity spotřebitelům. Kromě zemědělců jsou to dodavatelé zemědělských vstupů, zpracovatelé, dopravci a zprostředkovatelé na trhu, z nichž každý hraje svou roli v tom, aby byly řetězce efektivní.

Před více než 10 lety projekt FutureFarm [2] rozpoznal, že na zemědělský sektor má silný vliv několik různých vnějších faktorů. Mezi hlavní faktory budoucích změn podle [3] patří:

- Klimatické změny – a jejich vliv na složení plodin a způsobu hospodaření.
- Rostoucí populace – bude stimulovat rostoucí nároky na potraviny a energii.
- Náklady na energii – budou vyvolávat nové požadavky na nové metody výroby energie.
- Urbanizace a opouštění půdy – povede ke změnám ve společnosti a využívání půdy.
- Kvalita potravin – požadavky občanů a trhu na vyšší kvalitu potravinářské produkce.
- Stárnutí populace a zdravotní problémy – vyvolá specifické požadavky na výrobu potravin a stravování.
- Etnické a kulturní změny – vyvolají specifické požadavky na složení potravin (rostoucí muslimská populace v Evropě, rostoucí počet vegetariánů).
- Bioekonomika založená na znalostech – zavede nové produkty a plodiny, včetně GMO.
- Předpisy a normy – dohodnuté vládní normy pro výrobu a využívání energie a ochranu životního prostředí.
- Ekonomické nástroje – tržní nástroje (např. daně, obchodovatelná povolení) k internalizaci externalit a podpoře nákladové efektivnosti energetických a environmentálních politik a opatření.
- Dotace – postupné rušení neproduktivních produkcí a poskytování podpory tam, kde je třeba zmírnit environmentální a sociální náklady na změnu.
- Investice – skutečné náklady odrážející ceny na trhu, které vyšlou správné signály soukromým investorům.
- Partnerství a dobrovolné dohody – společné programy veřejného a soukromého sektoru pro rozvoj a zavádění udržitelné produkce.
- Výzkum a vývoj – vládní pobídky pro soukromý výzkum a vývoj na podporu inovací v oblasti zemědělství pro udržitelný rozvoj.
- Informace a komunikace – kampaně na podporu lepšího porozumění široké veřejnosti vnitrostátní a mezinárodní situaci v oblasti energetiky a životního prostředí.
- Hodnocení a scénáře – hodnocení udržitelnosti, která identifikují synergie a kompromisy napříč ekonomickými, environmentálními a sociálními dopady zemědělské politiky.
- Oceňování ekologických výkonů, dokud budou náklady převážně externalizovány směrem k životnímu prostředí (včetně zemědělství a lesnictví) a nebude existovat právní rámec. Potřebné změny budou probíhat pomalu.
- Národní strategie – přístupy dobré správy založené na rozhodování v rámci celé vlády, transparentnosti a pochopení politické ekonomie prosazování změn v zemědělských systémech.
- Politici - včetně jejich politického povědomí o stávající situaci a nutných změnách v budoucnosti, včetně dopadů a změn v důsledku životního prostředí nebo technologií.
- Tisk - vzhledem k jeho vlivu na politiky a jeho roli prostředníka mezi veřejností a politikou.
- Vzdělávání - včetně školení a přenosu know-how a povědomí o nezbytné rychlosti budoucích změn.
- Modely spolupráce a integrace - komplexní řetězec v zemědělství, biomase a životním prostředí i složitá struktura partnerů řetězce, jejich chování při řešení nových cílů včetně kontroly.
- Mezinárodní organizace - jako Evropská komise, World Bank, FAO, OECD atd. s jejich schopností/neschopností ovlivňovat potřebné změny.

Společný a budoucí vliv jednotlivých faktorů tvoří složitý proces, který je těžko popsitelný a vede ke složitým situacím. V mnoha případech mohou stát dva externí faktory proti sobě a jejich budoucí vliv na zemědělskou produkci

a trh s potravinami závisí výrazně na předpisech a společné politice. Například důraz na kvalitu a bezpečnost potravin může být v rozporu s požadavky na zvyšování produkce kvůli rostoucímu počtu obyvatel. Podobně je to s rostoucí produkcí a požadavkem na výrobu bioenergie. Některé faktory vedou ke zmenšování zemědělské půdy a zároveň k požadavku na zvyšování produkce. Potřebujeme zvýšenou kvalitu a bezpečnost, a to bude spojeno se snižováním počtu chemických vstupů. Tyto faktory také přinesou požadavky na snížení spotřeby energie a vody. Ukazuje se, že tato studie poměrně dobře analyzovala budoucí stav a problémy, které budou pro zemědělství vznikat. Projekt FutureFarm doporučoval dívat se na danou problematiku komplexně, zvažovat dopady jednotlivých vlivů a hledat optimální řešení.

2.2. CÍLE UDRŽITELNÉHO ROZVOJE A ZELENÁ DOHODA PRO EVROPU

Po více než deseti letech je zajímavé porovnat, jak a které vlivy skutečně působily/působí na naše prostředí, i na zemědělství, a zdá se, že většina těchto externích vlivů je stále platná. Je zajímavé se na toto podívat pohledem současných evropských a globálních strategií. Je také důležité, aby se k zavádění těchto strategií přistupovalo komplexně. V současné době se obvykle řeší dílčí jednotlivosti. Jak již bylo zmíněno v předchozí kapitole, mnohá opatření mohou jít proti sobě a vzájemně se rušit nebo znemožňovat.

2.2.1. Cíle udržitelného rozvoje (Sustainable Development Goals)

V globálním měřítku jsou dnes základním dokumentem Cíle udržitelného rozvoje (Sustainable Development Goals). 17 Cílů udržitelného rozvoje (SDGs, Obrázek 1) představuje program rozvoje na následujících 15 let (2015 – 2030). Jejich cílem je ukončit chudobu, chránit planetu a zlepšit životy a vyhlídky všech a všude. Těchto 17 cílů bylo přijato všemi členskými státy OSN v roce 2015 jako součást Agendy pro udržitelný rozvoj 2030, jež stanovila patnáctiletý plán k jejich dosažení. Cíle udržitelného rozvoje OSN [4] schválila EK v květnu 2019 a mnohé z nich se přímo týkají zemědělství, lesnictví a životního prostředí.

Mnoho z těchto cílů se přímo nebo nepřímo dotýká zemědělství. Zmíňme alespoň dva z nich uvedené v [5].

Cíl č. 2. Vymýtit hlad, dosáhnout potravinové bezpečnosti a zlepšit výživu, prosazovat udržitelné zemědělství.

- Do roku 2030 vymýtit hlad a zajistit přístup všem lidem, zejména chudým a ohroženým, včetně malých dětí, k bezpečné, výživné a dostačující stravě po celý rok.
- Do roku 2030 odstranit všechny formy podvýživy a do roku 2025 dosáhnout mezinárodně dohodnutých cílů ohledně zakrnění a hubnutí dětí mladších pěti let a vyřešit výživové potřeby dospívajících dívek, těhotných a kojících žen a starších osob.
- Do roku 2030 zdvojnásobit zemědělskou produktivitu a příjmy malých zemědělců – zejména žen, původních obyvatel, rodinných farmářů, pastevců a rybářů – pomocí



CÍLE UDRŽITELNÉHO ROZVOJE



Obrázek 1: Cíle udržitelného rozvoje (převzato z [5]).

zajištění bezpečného a rovného přístupu k půdě, dalším výrobním zdrojům (a vstupům), znalostem, finančním službám, trhům a vytvářením příležitostí k vytváření přidané hodnoty a přístupu k zaměstnání v zemědělském sektoru.

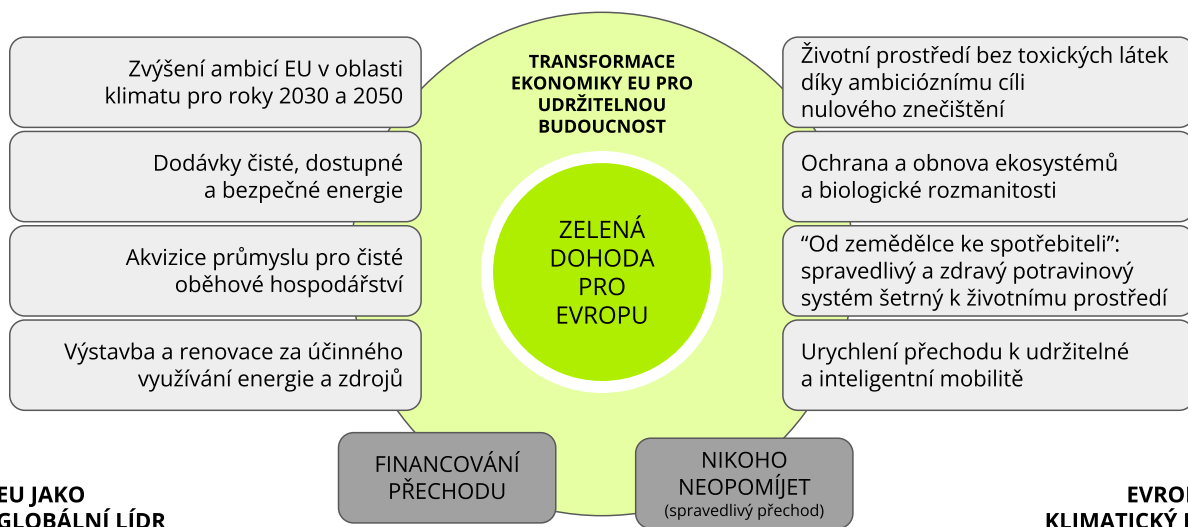
- Do roku 2030 zavést systémy udržitelné výroby potravin a odolné zemědělské postupy, které zvýší produktivitu a výrobu a pomohou zachovat ekosystémy posilující schopnosti půdy přizpůsobit se klimatické změně, extrémnímu počasí, suchu, záplavám a dalším pohromám, a které postupně zlepší kvalitu půdy.
- Do roku 2020 zajistit zachování genetické rozmanitosti osiv, pěstovaných plodin, hospodářských a domácích zvířat a jejich divoce žijících příbuzných druhů, mimo jiné pomocí správně spravovaných a diverzifikovaných semenných a rostlinných bank na národní, regionální i mezinárodní úrovni a zajistit přístup ke spravedlivému sdílení přínosů z využívání genetických zdrojů a tradičních poznatků, v souladu s mezinárodními dohodami.
- Zvýšit investice, mimo jiné i prostřednictvím posílené mezinárodní spolupráce, do venkovské infrastruktury, zemědělského výzkumu a nadstavbových služeb, vývoje technologií a genových bank zvířat a rostlin za účelem zlepšení zemědělských výrobních kapacit v rozvojových zemích, zejména v těch nejméně rozvinutých.
- Napravit a předcházet obchodním omezením a pokřivením na světových zemědělských trzích, mimo jiné prostřednictvím souběžného odstranění všech forem dotací na vývoz a dalších opatření s obdobným účinkem, v souladu s mandátem rozvojového kola v Dauhá.

- Přijmout opatření k zajištění řádného fungování trhu potravinářských komodit a jejich derivátů a usnadnit včasný přístup k informacím o trhu, včetně informací o potravinových rezervách, s cílem přispět k omezení extrémní volatility cen potravin.

Cíl č. 13. Přijmout bezodkladná opatření na boj se změnou klimatu a zvládnání jejich dopadů.

- Ve všech zemích zvýšit odolnost a schopnost adaptace na nebezpečí související s klimatem a přírodními pohromami.
- Začlenit opatření v oblasti změny klimatu do národních politik, strategií a plánování.
- Zlepšit vzdělávání a zvyšování povědomí o klimatické změně, rozšířit lidské i institucionální kapacity pro zmírňování změny klimatu, adaptaci na ni, snižování jejích dopadů a včasné varování.
- Uvést do praxe závazek přijatý vyspělými zeměmi v Rámcové úmluvě OSN o změně klimatu a do roku 2020 společně dát k dispozici ze všech zdrojů 100 miliard dolarů ročně na řešení potřeb rozvojových zemí v souvislosti se smysluplnými opatřeními na zmírňování a transparentností při jejich zavádění a plně zprovoznit Zelený klimatický fond v co nejkratší době.
- Podporovat mechanismy pro zvyšování kapacit pro efektivní plánování a řízení v oblasti změny klimatu v nejméně rozvinutých zemích, se zaměřením na ženy, mládež, místní a přehlížené komunity.
- S vědomím, že Rámcová úmluva OSN o změně klimatu je primárním mezinárodním, mezivládním fórem pro vyjednávání globální reakce na změnu klimatu.

ZELENÁ DOHODA PRO EVROPU



Obrázek 2: Zelená dohoda pro Evropu [6].

2.2.2. Zelená dohoda pro Evropu (Zelený úděl - Green Deal)

Zelená dohoda pro Evropu (též Zelený úděl - Green Deal) je balíček opatření Evropské komise oznámený 11. prosince 2019 [6], který by měl občanům a podnikům v Evropské unii zajistit přechod na udržitelnější a zelenější ekonomiku do roku 2050. Zelená dohoda pro Evropu (Obrázek 2) se skládá ze souboru opatření, která mají zajistit přechod na čisté oběhové hospodářství, zabránit ztrátě biologické rozmanitosti a snižovat znečištění. Tato opatření se týkají následujících oblastí politiky:

- Biologická rozmanitost
- Z pole na vidličku (Farm-to-fork)
- Udržitelné zemědělství
- Čistá energie
- Udržitelný průmysl
- Stavebnictví a renovace
- Udržitelná mobilita
- Eliminace znečištění
- Opatření v oblasti klimatu

Z výše uvedených pilířů strategie Zelené dohody pro Evropu jsou opatření Biologická rozmanitost, Z pole na vidličku, Udržitelné zemědělství a Klimatická opatření přímo spojena s aktuálními diskusemi o budoucnosti Společné zemědělské politiky (SZP), která je definována jako klíčová hnací síla při podpoře evropského zemědělského odvětví a má být důležitým nástrojem pro řízení přechodu k udržitelné produkci potravin a posílení úsilí evropských zemědělců přispět k řešení problémů EU v oblasti klimatu a životního prostředí. [7]

Z pole na vidličku (Farm to fork)

Cílem strategie je zajistit, aby potravinové systémy byly spravedlivé, zdravé a šetrné k životnímu prostředí, a to přijetím opatření:

- do roku 2030 snížit o 50 % používání chemických pesticidů a riziko jejich používání,
- do roku 2030 snížit o 50 % používání nebezpečných pesticidů,
- snížit ztráty živin nejméně o 50 % a zároveň zajistit, aby se nezhorsila úrodnost půdy,
- do roku 2030 snížit používání hnojiv alespoň o 20 %,
- dosáhnout dokončení rychlého širokopásmového přístupu k internetu ve venkovských oblastech,
- zvýšit plochu pro biologickou rozmanitost, včetně zemědělské plochy v rámci krajinných prvků s vysokou rozmanitostí.

Udržitelnost zemědělství v celé EU je dalším budoucím cílem Zelené dohody pro Evropu, který má být podporován SZP s ohledem na ekonomické, sociální a environmentální aspekty. Budoucí kroky na cestě k dosažení udržitelného systému zemědělství v EU zahrnují 9 cílů:

- zajistit zemědělcům spravedlivý příjem,
- zvýšit konkurenceschopnost,
- obnovit rovnováhu sil v potravinovém řetězci,
- opatření v oblasti změny klimatu,
- péče o životní prostředí,
- zachování krajiny a biologické rozmanitosti,
- podpora generační obnovy,
- živé venkovské oblasti,
- ochrana kvality potravin a zdraví.

Opatření v oblasti klimatu

Cílem opatření v oblasti klimatu v rámci zelených dohod je čelit změnám klimatu prostřednictvím klíčových právních předpisů a politik EU. Z oblasti relevantní pro zemědělství to zahrnuje:

- Národní cíle pro obchodování s emisemi v oblastech jako je doprava, budovy a **zemědělství**.
- Zajištění toho, aby **lesy a půda** přispívaly k boji proti změně klimatu.
- Snížení emisí skleníkových plynů z dopravy, např. prostřednictvím emisních norem CO₂ pro vozidla.
- Zvýšení energetické účinnosti, obnovitelných zdrojů energie a řízení energetických a klimatických politik zemí EU.
- Podpora inovativních nízkouhlíkových technologií.
- Přizpůsobení se dopadům změny klimatu.
- Financování opatření v oblasti klimatu.

Biologická rozmanitost

V rámci strategie pro biologickou rozmanitost budou přijata opatření, jejichž cílem je:

- zřídit chráněné oblasti na nejméně 30 % území v Evropě,
- zřídit chráněné oblasti alespoň na 30 % moří v Evropě,
- obnovit degradované ekosystémy na souši i na moři v celé Evropě cestou
 - rozšíření ekologického zemědělství a krajinných prvků bohatých na biologickou rozmanitost na zemědělské půdě,
 - zastavení a zvrácení úbytku opylovačů,
 - snížení používání pesticidů a rizika jejich používání o 50 % do roku 2030,
 - obnovení nejméně 25 000 km řek EU do volně tekoucího stavu,
 - vysazení 3 miliard stromů do roku 2030,
- uvolnit 20 miliard EUR ročně pro biologickou rozmanitost prostřednictvím různých zdrojů, včetně fondů EU a vnitrostátních a soukromých finančních prostředků,
- zajistit EU vedoucí postavení ve světě při řešení globální krize biologické rozmanitosti.

Zde je třeba si uvědomit, že se jedná o strategické cíle, které dosud nejsou ve většině promítnuty do jednotlivých

politik. Při formování těchto politik je třeba vycházet z důkladné analýzy (velkých) dat (viz další text). Zároveň je jasné, že toto bude klást enormní požadavky na prvovýrobce.[8]

2.3. ZNALOSTNÍ PYRAMIDA

Zjednodušíme-li mírně obsah předchozí podkapitoly, můžeme konstatovat:

“V budoucnu je třeba celosvětově zvýšit zemědělskou produkci, která bude kvalitnější a zároveň bude využívat méně půdy a méně vstupů.” [9]

Vzhledem ke složitosti tohoto problému je třeba, aby všichni aktéři v oblasti zemědělství počínaje politiky a konče farmáři lépe porozuměli všem procesům, které se na něm podílejí. Je potřeba i vytvořit znalostní řízení i pro formování národních politik. Řízení znalostí se obvykle popisuje jako převod dat na informace, znalosti a nakonec na moudrost. Pro každou diskusi o řízení znalostí nebo informací je důležité lépe porozumět základním pojmům jako jsou data, informace a znalost a také vztahům mezi těmito kategoriemi. Důležité je také porozumět problematice interoperability a standardizace. Pro lepší vysvětlení budeme používat hierarchii data-informace-znalosti-vědomí (DIKW pyramida, viz [10]) (Obrázek 3).

- Data: jako symboly;
- Informace: data, která jsou zpracována tak, aby byla užitečná; poskytují odpovědi na otázky "kdo", "co", "kde" a "kdy";
- Znalost: aplikace dat a informací; odpovídá na otázku "jak";
- Vědomí (moudrost): vyhodnocené porozumění.

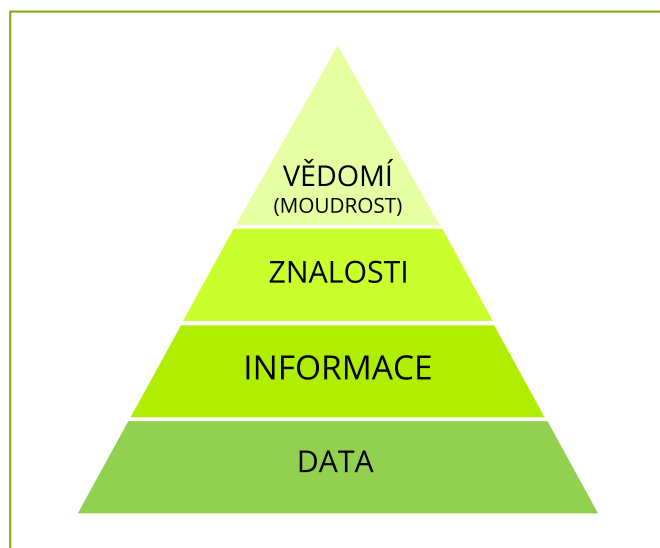
Pro pochopení následujících kapitol jsou důležité první tři pojmy data, informace a znalosti, které definují tři úrovně související s řízením zemědělských podniků, ale i řízením zemědělství jako celku. Moudrost lze pak chápat jako součást rozhodovacích procesů [11].

Definovali jsme tři úrovně řízení:

- správa dat,
- správa informací,
- řízení znalostí.

Správa dat zahrnuje

- datovou architekturu, analýzu a návrh,
- správu databází,
- řízení bezpečnosti dat,
- řízení kvality dat,



Obrázek 3: Hierarchie data-informace-znalosti-vědomí podle Ackoffa (upraveno dle [10]).

- správu metadat,
- správu dokumentů, záznamů a obsahu,
- správu referenčních a kmenových dat.

Správa informací zahrnuje

- snadné vyhledávání informací,
- přístup k informacím.

Řízení znalostí zahrnuje

- shromažďování, šíření, analýzu a využívání informací; toto zahrnuje i využívání modelů, umělé inteligence apod.,
- rámec pro navrhování cílů, struktur a procesů tak, aby přinášely přidanou hodnotu,
- sladění s cíli a strategiemi,
- zaměření na sdílení a tvorbu strategií,
- propojení lidí a organizací.

Pochopení vztahu mezi daty a moudrostí je důležité pro to, aby se zemědělství (nejen české) mohlo přizpůsobit budoucím změnám. Je nutné na jedné straně pochopit, že cílem každého informačního/znalostního systému musí být ne pouze sběr a shromažďování dat, ale také získání vědomí, jak řešit určitý problém, a naopak, že nelze přijímat řešení konkrétních problémů, bez důkladné analýzy dat.

2.4. CHUDÁK FARMÁŘ

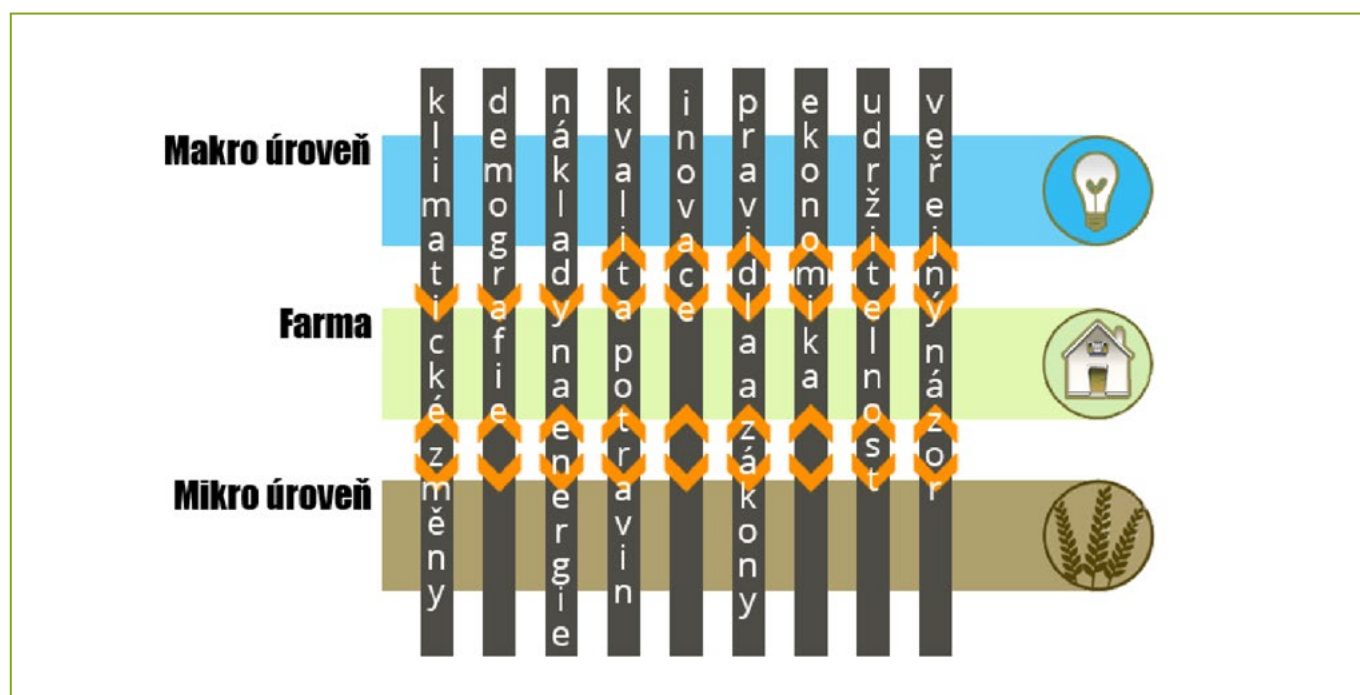
V kapitole 2.1. Externí vlivy jsme identifikovali řadu faktorů, které ovlivňují a budou stále více ovlivňovat zemědělskou výrobu a produkci [12]. Pokud se na tyto faktory díváme z pohledu farmy, pak je třeba data, informace a znalosti rozdělit do tří úrovní (Obrázek 4):

- Makroúroveň, která zahrnuje řízení externích informací, např. trh, dotační systémy, předpověď počasí, globální trh a systémy sledovatelnosti atd.
- Úroveň zemědělského podniku, která zahrnuje např. ekonomické systémy, střídání plodin, systémy na podporu rozhodování.
- Mikro úroveň - neboli úroveň uvnitř honu, která zahrnuje např. precizní zemědělství, shromažďování informací o sledovatelnosti a v budoucnu také robotika.

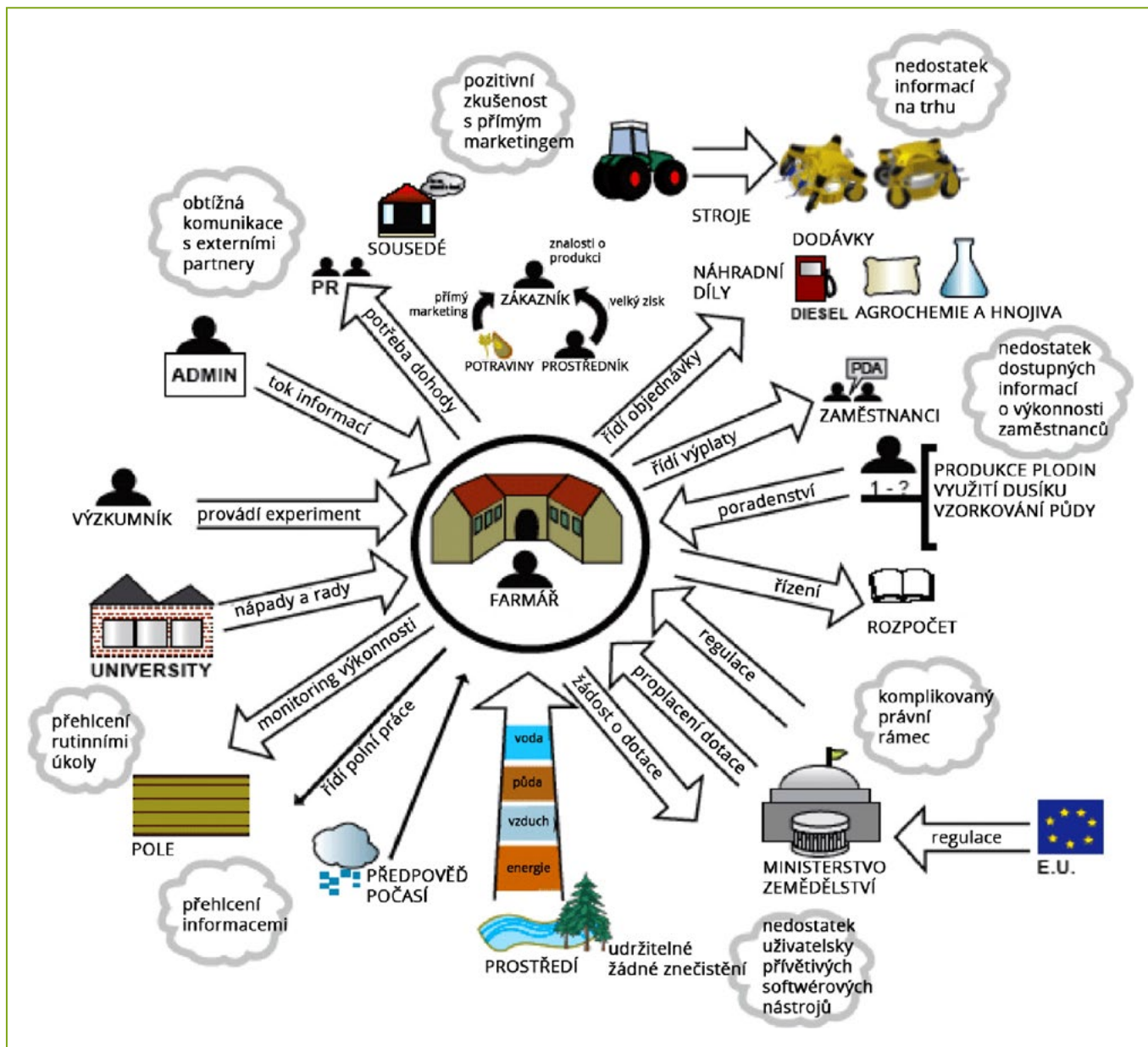
V současné době jsou nejvíce rozvíjeny systémy řízení znalostí na úrovni zemědělského podniku a systémy mikro úrovně. Práce s informacemi a znalostmi na makro úrovni

je prozatím omezená. Aby bylo možné navrhnout budoucí systém řízení znalostí, je třeba vzít v úvahu vzájemné vazby, jednotlivých faktorů, které zemědělství ovlivňují a budou ovlivňovat a jejich prolínání přes všechny tři úrovně.

Budoucí, ale i dnešní zemědělec musí být schopen spravovat velké množství informací, aby mohl činit ekonomicky a ekologicky správná rozhodnutí. Tento proces je velmi náročný na pracovní sílu, protože jednotlivé systémy spolu obvykle nekomunikují a mnoho činností se musí provádět ručně. Běžné řízení provozu farmy zahrnuje úkoly sledování polních operací, správu financí a žádosti o dotace, což je obvykle závislé na různých softwarových aplikacích. Zemědělec musí být schopen používat různé nástroje pro řízení, monitorování a sběr dat on-line na poli, viz Obrázek 5 níže. Potřebuje analyzovat informace týkající se dotací, komunikovat s finančními úřady, prodejci produktů atd. Již před mnoha lety toto charakterizovali např. [13, 14].



Obrázek 4: Úrovně řízení znalostí na farmě a vliv externích faktorů (upraveno dle [12]).



Obrázek 5: Úrovně řízení znalostí na farmě a vliv externích faktorů (upraveno dle [12]).

Jakékoli rozhodnutí na úrovni zemědělského podniku vyžaduje intenzivní tok dat, informací a znalostí. To vyžaduje mnoho různých nástrojů. Při řešení tohoto problému hraje roli tzv. interoperabilita. Interoperabilitu na úrovni farmy nelze omezit na jednoduché datové struktury, protokoly, ale je třeba ji chápat v plném rozsahu komunikace mezi jednotlivými systémy [16]. Více k tomuto problému v kapitole 5. Co je to Interoperabilita.

2.5. DESTINATION EARTH: DIGITÁLNÍ DVOJČE ZEMĚ

Vzhledem k předchozím bodům se jako důležitá jeví iniciativa Destination Earth (DestinationE). Evropská komise společně se zveřejněním Zelené dohody pro Evropu přijala průmyslovou strategii EU, která má řešit dvojitý výzvu, a to zelenou a digitální transformaci. Cílem je využít potenciál digitální transformace, aby se stala klíčovým prostředkem pro dosažení cílů Zelené dohody. Komise plánuje iniciativu "GreenData4All", která se zaměří na revizi směrnice o zřízení infrastruktury pro prostorové informace v EU (INSPIRE viz kapitola 5.2. Standardizační organizace a iniciativy relevantní pro zemědělství) a spojí ji se směrnicí o přístupu k informacím o životním prostředí. Součástí bude iniciativa "Destination Earth".

"Iniciativa 'Destination Earth' má za úkol spojit evropskou vědeckou a průmyslovou excelenci s cílem vytvořit velmi přesný digitální model Země - Digitální dvojče (Digital Twins¹). Tato průkopnická iniciativa bude mít za úkol vyvinout modelovací platformu pro vizualizaci, monitorování a předpovídání přírodní a lidské činnosti na planetě pro podporu udržitelného rozvoje a zároveň podpořit zavádění zelené dohody. Digitální dvojče Země bude budováno postupně, počínaje rokem 2021." Součástí tohoto záměru je i "Společný evropský datový prostor pro zemědělství, který má zvýšit výkonnost, udržitelnost a konkurenceschopnost zemědělského odvětví prostřednictvím zpracování a analýzy produkčních a dalších dat, což umožní přesné a na míru šité uplatňování výrobních přístupů na úrovni zemědělských podniků."

Destination Earth se bude postupně rozvíjet prostřednictvím následujících kroků [8, 17]:

- Do roku 2024: Vývoj otevřené základní digitální platformy a prvních dvou digitálních dvojčat o extrémních přírodních jevech a přizpůsobení se změně klimatu.
- Do roku 2027: Integrace dalších digitálních dvojčat, jako je digitální dvojče oceánu.
- Do roku 2030: "úplná" digitální replika Země prostřednictvím konvergence digitálních dvojčat, která již budou v rámci platformy nabízena.

Tento harmonogram ale časově nepodporuje implementaci strategii Zelené dohody pro Evropu do zemědělských

politik, jelikož velké množství cílů Green Deal je stanoveno do roku 2030 a tudíž není reálné využití tohoto nástroje pro formování politik a ani využití těchto nástrojů pro podporu zemědělské výroby. Jako řešení se jeví využití existujících infrastruktur a zkušeností z minulých a současných výzkumných projektů pro potřebu rozvoje zemědělství u nás. Vybrané možnosti budou nastíněny v následujících kapitolách.

¹ Princip Digitálních Dvojčat byl převzat z průmyslu. Digitální dvojčata vytvářejí virtuální repliku fyzického výrobku, procesu nebo systému. Replika může například na základě analýzy dat předpovědět, kdy dojde k poruše stroje, což umožňuje zvýšit produktivitu díky prediktivní údržbě.



3. CO JSOU BIG DATA (VELKÁ DATA)

3.1. ZÁKLADNÍ DEFINICE VELKÝCH DAT

Pro budování znalostního zemědělství a zemědělských digitálních dvojčat budeme muset pracovat s obrovským množstvím heterogenních dat s vysokou rychlostí, tj. s velkými daty. Důvodem je obrovský a neustále rostoucí tok dat z nejrůznějších senzorů a měřících zařízení, jako jsou kamery na satelitech, letadlech a dronech, a také data z měření ze senzorů ve vzduchu, v půdě a ve vodě. Navíc se exponenciálně zvyšuje rozlišení a četnost sběru dat z těchto senzorů.

Mnoho průmyslových odvětví těžší z velkých dat, která byla označena za "novou ropu" [18]. Termín Big Data se používá od roku 2001, kdy Doug Laney představil charakteristiku **3V**: Volume - objem, Velocity - rychlost a Variety - rozmanitost [19, 20].

Postupem času byla přidána další dvě "V" (value - hodnota a veracity - pravdivost), která datovým vědcům pomáhají efektivněji formulovat a komunikovat důležité charakteristiky velkých dat. Jednotlivá **V** jsou chápána následovně:

Volume - Objem

Základním problémem při sběru, ukládání a správě dat je objem. Jde o velikost dat z dálkového průzkumu Země, velikost dat, která jsou vytvářena senzory, data ze strojů, mapy, fotografie, videa, literatura atd. Místo nezabírají pouze nezpracovaná data. Aby byla data užitečná, musí být indexována, propojena s jinými daty, obohacena a transformována do datových produktů. Pro efektivní využití dat může být nutné uložit několik verzí stejných dat, aby se urychlily dotazy a umožnila vizualizace v reálném čase. To vše přispívá k tomu, že data a velikost úložiště jsou mnohonásobně větší než surová data. Kromě prostoru je úložiště také otázkou nákladů. Úložiště stojí peníze, a přestože se složitost správy úložiště s přechodem na cloudová úložiště snížila, náklady zůstávají stejné. To vede ke značné spotřebě energie. Náklady na ukládání dat, jsou tedy ekologické, technické a ekonomické.

Velocity - Rychlost

Kromě objemu je dalším problémem, že měřena data ze senzorů, strojů a podobně přicházejí velkou rychlostí. Moderní

řešení Internetu věcí (IoT) jsou schopna okamžitě předávat měření uživatelům. Nejdůležitějším aspektem rychlosti je schopnost zpracovat všechna přichodící data rychlostí, jakou přicházejí. Na čistě technologické úrovni to souvisí se zpracováním tisíců souběžných měření.

Rychlost má také vliv na to, kde je možné daná data použít. Pokud se tok dat používá pro řízení procesů, může být vyžadována okamžitá odezva na měření a synchronizace dat je tak pro provoz zařízení velmi důležitá. Totéž platí, pokud jsou data používána pro podporu kritických rozhodnutí.

Dalším rozměrem rychlosti je dostupnost a stabilita sítí. Proudové načítání dat vyžaduje nepřetržitou stabilitu sítě a ta není samozřejmostí. Zejména kvůli heterogenitě senzorových zařízení.

Variety - Rozmanitost

Termín Variety se používá k popisu rozmanitosti typů dat. Datové infrastruktury mohou shromažďovat data z různých zdrojů dat a měřících přístrojů, které se mohou výrazně lišit. To vede k potřebě normalizace, harmonizace, standardizace a distribuce všech shromažďovaných dat. Data mohou mít různou strukturu v závislosti na přístrojích nebo metodách použitých k jejich sběru. Nestrukturovaná a polostrukturovaná data se snadno shromažďují, ale v konečném důsledku se obtížně používají; cílem většiny procesů sběru, správy a sdílení dat je proto dospět ke standardizovaným strukturovaným datovým produktům. Strukturovaná data jsou pak k dispozici pro efektivní zpracování a analýzu dat. Velká datová úložiště mohou přijímat data z libovolného počtu různých standardů a ukládat formáty jako XML a JSON.

Ukládání však není totéž, co používání. Aby byla data užitečná, musí být strukturovaná, význam datových modelů je proto nyní větší než kdykoli předtím. A protože jsou nyní data nabízena prostřednictvím rozhraní API a ukládána v cloudu, přičemž k nim lze přistupovat z mnoha různých koncových bodů, je také nutné zvládnout verzování rozhraní API a datových modelů, aby se závislé aplikace nerozply, když se webové služby a datové modely vyvíjí.

Veracity - pravdivost

Termín Veracity se používá k popisu důvěryhodnosti a kvality dat. Data in situ mohou mít různou kvalitu v závislosti na přístrojích nebo metodách použitých k jejich sběru. Podstatným faktem je věrohodnost, která se týká kvality a přesnosti dat. Shromážděná data mohou být neúplná, mohou být nepřesná nebo nemusí být schopna poskytnout znalosti, které jsou potřeba pro rozhodnutí. Tento termín se také používá k popisu úrovně důvěry ve shromážděné údaje. Vzhledem ke způsobu sběru dat mohou data chybět, mohou být chybná, chybějící nebo zpřeházená v důsledku technických problémů a obtížné synchronizace verzí a aktualizací. Množství dat se může snadno stát prostředkem k zastření faktů stejně jako k odhalení nových poznatků, pokud jsou data neúplná.

Value - hodnota

Posledním, ale neméně důležitým "V", je hodnota. Hodnota se týká environmentální, ekonomické a sociální hodnoty, kterou lze vytvořit díky chytrému využití. Týká se přímo toho, co mohou organizace s těmito shromážděnými daty udělat. Schopnost vytěžit hodnotu = znalost a moudrost z velkých dat je podmínkou. Hodnota velkých dat se výrazně zvyšuje v závislosti na poznacích, které z nich lze získat. Jaký smysl má sdílení dat, pokud neexistují uživatelé znalostí? Vždy však existuje šance, že někdo, někde může vytvořit hodnotu, dokud jsou data k dispozici. V některých oblastech však víme, jaká data jsou potřeba, abychom mohli provádět patřičná rozhodnutí. V těchto případech bychom měli vynaložit úsilí na zvyšování povědomí a podporu řešení, která mohou sloužit jako předstupeň komplexnějšího úsilí s širším dopadem. Různé organizace mohou ke shromažďování a analýze dat používat stejné nástroje pro Big Data, ale způsob, jakým z těchto dat získají hodnotu, by měl být pro ně jedinečný.

V následujících kapitolách se zaměříme na různé aspekty velkých dat a i na nástroje, které pomohou k jejich optimálnímu využití.

3.2. KDO SKUTEČNĚ POTŘEBUJE DATA

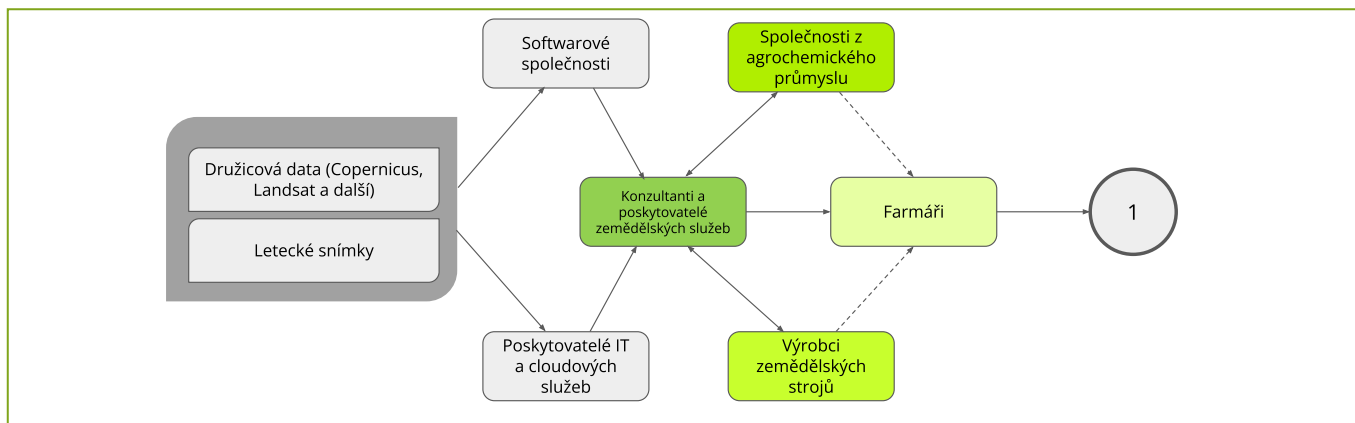
V souladu s tím, co jsme již napsali v kapitole 2.3. Znalostní pyramida, je nutné zdůraznit, že farmáři ale i další ekonomické subjekty, samospráva, obyvatelé venkovských oblastí, turisté a další, nejsou skutečnými uživateli. Cílem těchto skupin, není zpracovávat a analyzovat data, ale využívat znalosti. Zde musí existovat skupina odborníků (mediátorů), kteří pomohou s analýzou dat a jejich transformací do znalostí (Obrázek 6). Je však nutné jistým způsobem zajistit transparentnost procesu generování znalostí a schopnost srozumitelným způsobem dokumentovat, jak k tvorbě těchto znalostí došlo. Zároveň je nutné být schopen doložit (v souladu s předchozí kapitolou) i kvalitu vstupních dat.

Do budoucna se nabízí otázka, do jaké míry bude možné tuto skupinu mediátorů nahradit umělou inteligencí

[AI - artificial intelligence]. Doposud je AI používána především jako efektivní nástroj pro řešení dílčích úkolů (algoritmů), ale bez přítomnosti expertů není schopna řídit celý proces zemědělské produkce. Do budoucna je hlavní otázkou, jak bude AI schopna zajistit transparentnost procesu rozhodování a provádět zdůvodnění jednotlivých rozhodnutí.



Obrázek 6: Kdo pracuje a nepracuje s velkými daty [upraveno dle [15]].



Obrázek 7: Řetězec organizací přidávajících datům hodnotu (upraveno dle [22]).

3.3. ZPĚT KE ZNALOSTNÍ PYRAMIDĚ

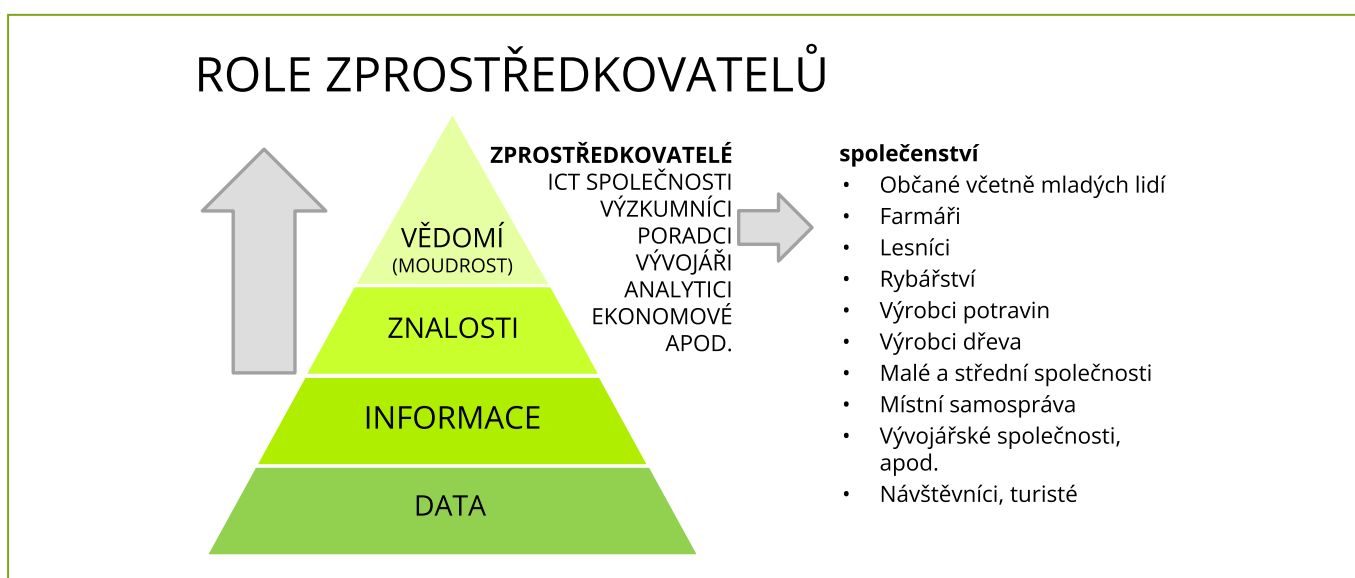
Cesta, jak pomoci farmářům řešit problém se zpracováním velkých dat byla již v minulosti nastíněna projekty Evropské kosmické agentury (ESA²) VOICE a AMI4FOR. Nutné znalosti musí být generovány pomocí organizací předávajících datům hodnotu (Value Added Organization - VAO), které transformují data ve znalosti, a tyto znalosti jsou pak dále užívány koncovými uživateli. V rámci projektu VOICE byla zahájena realizace prototypu, který měl za cíl demonstrovat proveditelnost vytvoření podmnožiny kompletní nabídky organizací s přidanou hodnotou (VAO) až ke koncovému uživateli pro data Dálkového průzkumu Země. Projekt AMI4FOR pak rozšiřuje možnosti využívání VAO v oblasti sběru primárních dat a jejich sdílení. [21]

Tento koncept byl nyní opět analyzován projektem EO4Agri. Uživatele dat lze rozdělit podle toho, zda jsou zpracovatelé dat nebo spotřebitelé výsledků vytvořených společnostmi zpracovávajícími data. V první skupině jsou: poskytovatelé služeb IT, výrobci softwaru a konzultanti a servisní organizace z řad pozorování Země, letecké fotogrammetrie, aplikace dronů, fytopatologie,

agronomie, interpretace vegetačních dat z družicových a leteckých snímků atd. Druhou skupinu tvoří zemědělci, inženýrské a ekologické zemědělské společnosti, agronomové, výrobci strojů a dodavatelé vstupů (hnojiva, chemikálie). Členové obou těchto skupin musí spolupracovat, aby mohli vyvinout všechny požadované aplikace. Z těchto aplikací pak budou mít prospěch všichni: zemědělci mohou činit informovaná rozhodnutí týkající se jejich plodin a poradci mohou prodávat služby zemědělcům a poskytovatelé vstupů mohou využívat mapy počasí a půdy k předvídání poptávky po hnojivech. Dle EO4Agri musí být přesné zemědělství vzhledem ke své komplexnosti podpořeno celým řetězcem přidané hodnoty, jehož konečným bodem je zemědělec, viz Obrázek 7 výše.

Vrátíme-li se tedy ke znalostní pyramidě (Obrázek 8), vidíme, že je potřeba interakce celé řady odborníků k tomu, abychom mohli plně transformovat velká data do znalostí a moudrosti [22].

² ESA = European Space Agency



Obrázek 8: Hierarchie data-informace-znalosti-vědění v životě regionů (upraveno dle [15]).



4. OTEVŘENÁ A SDÍLENÁ DATA

V předchozí kapitole 3.3. Zpět ke znalostní pyramidě jsme si ukázali, že při transformaci dat do znalostí je nezbytné, aby data a informace spolu sdílelo více uživatelů či organizací. Mnoho lidí si dnes neuvědomuje, že otevřená data a sdílená data, jsou dva různé pojmy. Toto je nutné pochopit, aby nedocházelo k omylu, že pokud zemědělec s někým data sdílí, stávají se tato data daty otevřenými. Budoucí znalostní systémy musí být schopny pracovat jak s otevřenými daty, tak i se sdílením uzavřených dat. Otevřená data dnes hrají důležitou roli při tvorbě naší pyramidy a získávání znalostí a moudrostí. Stát spravuje mnoho dat, která mohou napomoci celému sektoru.

Základní myšlenkou za tzv. otevřenými daty je to, že taková data by měla být volně dostupná každému, aby je mohl používat a publikovat podle svého přání, bez omezení vyplývajících z autorských práv, patentů nebo jiných kontrolních mechanismů [23]. V případě uzavřených datových sad je přístup k nim omezen heslem nebo jinými bezpečnostními mechanismy, které neumožňují jejich veřejné volné využití. Data jsou uzavřena z důvodu politiky vydavatelů a poskytovatelů dat. Uzavřená data mohou být soukromá data, obchodní nebo z hlediska bezpečnosti citlivá data, která nejsou zpřístupněna zbytku světa.

Všechna data se nachází na škále mezi uzavřenými a otevřenými daty. Uzavřená data mohou být sdílena s konkrétními osobami v rámci uzavřeného, např. firemního prostředí. Otevřená data obvykle vyžadují uvedení zdroje, který je poskytl a na základě splnění této jednoduché podmínky mohou být zcela k dispozici koncovému uživateli. Ještě jednou připomínáme, že je třeba si uvědomit, že sdílení dat je něco jiného než otevřená data. Sdílená data mohou být na základě určité dohody uzavřená, nicméně poskytovaná data. Sdílená data mohou být přístupná jako otevřená data, ale jejich užití může být omezeno nějakými podmínkami, například podmínkou nekomerčního využití či přístupností pouze určitým skupinám uživatelů. Sdílení dat je klíčové v zemědělství. Pro sdílení dat musí mít zemědělci jistotu, že jejich data jsou chráněna před neoprávněným použitím. Proto je nutné si uvědomit, že sdílení dat se liší od konceptu otevřených dat. Sdílená data mohou být uzavřená data na základě určité dohody mezi

konkrétními stranami, např. v podnikovém prostředí, zatímco otevřená data jsou veřejně dostupná komukoli. Otevřená data mohou vyžadovat uvedení zdroje, který k nim přispěl, ale přesto mohou být zcela dostupná koncovému uživateli. Data jsou neustále sdílěna mezi zaměstnanci, zákazníky a partnery, což vyžaduje strategii, která průběžně zabezpečuje datová úložiště a uživatele. Data se pohybují mezi různými veřejnými a soukromými úložišti, aplikacemi a operačními prostředími a jsou přístupná z různých zařízení a platform. To se může stát v kterékoli fázi životního cyklu zabezpečení dat, a proto je důležité použít správné bezpečnostní kontroly ve správný čas. Důvěra vlastníků dat je klíčovým aspektem pro sdílení dat [20, 24].

Obecně se otevřená data liší od uzavřených dat třemi klíčovými faktory [24]:

- **Dostupnost a přístup:** otevřená data jsou přístupná a dostupná obvykle prostřednictvím internetu ve vhodné a upravitelné formě (strojově čitelném formátu).
- **Opakované použití a další šíření:** data musí být poskytnuta za podmínek, které umožňují opakované použití a další šíření, včetně kombinace s jinými datovými zdroji (viz FAIR principy popsány níže).
- **Všeobecná účast:** data jsou licencována jako open source, což umožňuje komukoli data používat nebo sdílet za účelem nekomerčního nebo komerčního zisku. Nejsou například povolena "nekomerční" omezení, která by bránila "komerčnímu" využití, nebo omezení využití pro určité účely.

Princip FAIR dat a služeb (Findable, Accessibility, Interoperate, Reusable) zajišťuje, že data a služby lze nalézt prostřednictvím katalogů nebo vyhledávačů, jsou přístupná přes otevřená rozhraní, jsou v souladu se standardy pro interoperabilní zpracování těchto dat (viz kapitola 5. Co je to interoperabilita), a lze je tedy snadno opakovaně používat. Opakované použití dat je klíčovou složkou FAIR dat. Zajišťuje, aby data mohla být opakovaně použita k jiným účelům, než pro které byla původně vytvořena. Toto opakované použití zlepšuje poměr nákladů na původní tvorbu dat a umožňuje vzájemné obohacování napříč komunitami [25].

FAIR principy však záměrně nspecifikují technické požadavky, jsou souborem hlavních zásad, které zajišťují kontinuální zvyšování opakované použitelnosti prostřednictvím mnoha různých implementací. Nejde však o standard pro poskytování dat, umožňuje celou řadu přístupů a postupů, jak data a služby vytvořit v duchu těchto principů. Navíc splnění těchto principů nedefinuje otevřená data. FAIR principy vyžadují jasnost a transparentnost podmínek přístupu a opakovaného použití. Ačkoli tedy data FAIR nemusí být otevřená, aby byla splněna podmínka opakovaného použití, musí mít data FAIR jasnou, nejlépe strojově čitelnou licenci [26].

Principů FAIR je s úspěchem využíváno i v rámci životního cyklu správy dat, která mají být shromážděna, zpracována a/nebo vytvořena v rámci projektů financovaných z fondů Evropské unie [27]. V této publikaci lze najít několik příkladů projektů, které tyto principy uplatnily.

Pro otevřená data jsou definovány požadavky, tzv. stupně otevřenosti dat, jejichž splnění určuje jejich míru otevřenosti a také propojitelnosti a dle ní jsou daná data označena hvězdičkami od jedné do pěti. Seznam těchto požadavků je následující [28]:

- otevřenost z pohledu licence (*),
- dostupnost ve strukturovaném formátu dat umožňující strojové čtení dat (**),
- dostupnost v otevřeném neproprietárním³ formátu dat (***),
- jednoznačné identifikátory prvků pomocí URI (****),
- standardizované propojení na jiná data poskytujícím datům kontext (například na objekty v databázích Dbpedia, LinkedGeoData.org nebo GeoNames.org) (*****).

Uzavřená data omezují přístup k informacím několika možnými způsoby:

1. Jsou dostupná pouze určitým osobám v rámci organizace.
2. Data jsou patentovaná nebo chráněná.
3. Data jsou částečně omezena na určité skupiny.
4. Data, která jsou přístupná veřejnosti prostřednictvím licenčního poplatku nebo jiné předběžné podmínky.

5. Údaje, které jsou obtížně přístupné, například papírové záznamy, které nebyly digitalizovány.

Příkladem uzavřených údajů jsou informace, které vyžadují bezpečnostní prověrku; informace týkající se zdravotního stavu shromažďované nemocnicí nebo pojišťovnou; nebo v menším měřítku osobní daňová přiznání [20].

³ Neproprietární = neprivátní, neuzavřený, tj. s otevřeným kódem (open source).



5. CO JE TO INTEROPERABILITA

V předchozích kapitolách jsme několikrát zmínili pojem Interoperabilita. **Interoperabilita** je v praxi důležitá, protože umožňuje spolupráci různých komponent (systémů, organizací, dat atp.). Tato schopnost komponent, jejich propojování a vzájemná kompatibilita je pro budování velkých a složitých systémů zásadní. Bez interoperability je to téměř nemožné – jak dokládá nejznámější mýtus o babylonské věži, kde (ne)schopnost komunikovat (spolupracovat) vedla k úplnému krachu úsilí o stavbu věže. Datová interoperabilita tedy výrazně zlepšuje schopnost kombinovat různé datové soubory, a tím vyvíjet další a lepší produkty a služby. Pokud budete užívat systémy, které nejsou interoperabilní, budete navždy přivázáni k jednomu dodavateli služeb a nebudete schopni integrovat další data a další služby. Jen zavedením standardizačních modelů je možné umožnit komunikaci a fungování více méně nezávislých systémů používaných v zemědělství a nastolit tak jejich interoperabilitu. Pod pojmem "standards" zde rozumíme protokoly, které popisují způsob definování dat a jejich výměny a umožňují tak digitální výměnu dat mezi zařízeními. Standards snižují transakční náklady na sdílení dat a často podporují hospodářskou soutěž, protože uživatelé mohou snadno změnit dodavatele. Uživatelé nejsou "uzamčeni" v uzavřeném systému. Normy často podporují inovace nebo poskytují základní vrstvu, na níž jsou postaveny nové inovace. Nicméně odvrácenou stranou standardizace může být i to, že organizací poskytujících standardy je celá řada a že implementace standardů v zemědělské praxi jsou stále bohužel zřídka k vidění. Tyto standardy umožňují interoperabilitu mezi všemi zúčastněnými subjekty, zajišťují kompatibilitu a jsou základním nástrojem pro integraci různých komponent do jednoho funkčního systému [24, 29].

5.1. DATOVÉ TYPY A DATOVÉ FORMÁTY

Hovoříme-li o datech, můžeme je chápat jako „Reprezentace hodnot vhodně formalizované pro komunikaci, interpretaci a zpracování lidmi a počítači. Data mohou být reprezentována libovolnými řetězci znaků (čísel, příkazů, vět) uloženými na informačním nosiči. Data nemají zpravidla význam sama o sobě, ale teprve jsou-li pochopena, interpretována, komunikována a využita člověkem nebo počítačem, stávají se smysluplnými informacemi“ [30]. Právě pro pochopení těchto dat slouží tzv. metadata, která doplňují k datům jejich význam či kontext. Metadata se detailně zabývá kapitola 8. Co jsou metadata a opravdu je potřebujeme?.

Pro potřeby lidského pochopení dat je vhodnější hovořit o datech na vyšší úrovni abstrakce, např. že daný soubor je obrázek, který je zaznamenaný (uložený) v určitém datovém formátu. Příkladem takového datového formátu může být značkový jazyk HTML nebo formáty dokumentů (formát textového procesoru Microsoft Word, OpenDocument, Office Open XML), grafické formáty pro výše uvedený příklad s obrázkem atp.

Aby počítače a programy různých výrobců mohly s daty pracovat, je datový formát obvykle standardizován. Neveřejné (proprietární) formáty nutí uživatele používat pro práci s daty programy určitého výrobce naopak otevřené formáty, např. pro texty či obrázky, bývají často podporovány různými výrobci a softwary a jejich použití je jednodušší. Znalost formátu dat je totiž nezbytná k většině manipulací s daty, ať už se jedná o ukládání, předávání nebo konverzi dat. Z pohledu standardizace datových formátů hovoříme o třech typech těchto dat [31]:

1. Strukturovaná data – jsou taková, která mají předem známý a definovaný jejich datový model, tj. jejich podobu: jaká data a jak budou uložena a jakých hodnot mohou nabývat. V takto strukturovaných datech se lépe vyhledává a také se s nimi snáze pracuje. Strukturovaných dat se s úspěchem využívá u webových stránek, u kterých strukturovaná data umožní internetovým vyhledávačům lépe pochopit informace obsažené na webu, protože je v nich ve strojově čitelné podobě popsáno, co je náplní webové stránky. To je základem

pro tvorbu tzv. sémantického webu na základě propojených dat (více o propojených datech viz kapitola 4. Otevřená a sdílená data).

2. Částečně strukturovaná data – jsou na rozhraní strukturovaných a nestrukturovaných dat. Používají k identifikaci určitých prvků v datech předem definované značky, ale data v nich nemají pevnou strukturu. Typickými příklady těchto dat jsou např. emaily či obrázky. Emaily mají odesílatele, příjemce, datum, čas a další pevné údaje, k nim jsou přidány části s různým obsahem a strukturou, což je samotná e-mailová zpráva či případné přílohy. Fotografie nebo jiné grafické prvky mohou být označeny klíčovými slovy, jako je autor, datum, místo a dalšími klíčovými slovy, což umožňuje tyto grafické prvky uspořádat a vyhledat.

3. Nestrukturovaná data – jsou taková, která nemají předem daný datový model, nejsou nijak organizovaná a nelze je jednoduše klasifikovat. To vede k nejednoznačnosti ve výkladu a pochopení takovýchto dat. Jde o veškerá data bez metadatových záznamů, tj. kontextu a popisu obsahu, kromě výše uvedených možností jde např. i o soubory PDF, PowerPointové prezentace, nestrukturované dokumenty atp. Pokud je to však možné, lze v se v nich orientovat pouze pomocí fulltextového vyhledávání.

Se zřetelem na již zmíněná velká data (viz kapitola 3. Co jsou to Big Data (Velká Data) bylo vytipováno šest kategorií velkých dat z pohledu způsobu využití a různých technik jejich analýzy a uložení [32]:

1. strukturovaná data,
2. časosběrná data,
3. geoprostorová data,
4. média, obrazová, video a zvuková data,
5. textová data, včetně dat zpracování přirozeného jazyka a genomických reprezentací,
6. grafová data, síťová/webová data a metadata.

5.2. STANDARDIZAČNÍ ORGANIZACE A INICIATIVY RELEVANTNÍ PRO ZEMĚDĚLSTVÍ

V úvodu této kapitoly byly zmíněny standardy jako základní kámen umožňující interoperabilitu nejen dat. Tyto standardy však musí být produkovány buď organizací, která má normotvorný statut, nebo se daný standard ujme pro svoji použitelnost v praxi na základě jeho častého používání uživateli. Dle [29] jsou zásadními standardizačními organizacemi a iniciativami takové, které jsou uvedeny v následujících odstavcích.

ISO

ISO je Mezinárodní organizace pro normalizaci, která vytváří a vydává mezinárodní normy. Normy ISO zajišťují, aby výrobky a služby byly bezpečné, spolehlivé a kvalitní. Pro podniky představují strategické nástroje, které snižují náklady minimalizací plýtvání a chyb a zvyšují produktivitu. Pomáhají společností přistupovat na nové trhy, vyrovnávají podmínky pro rozvojové země a usnadňují volný a spravedlivý globální obchod [33]. "ISO normy pro zemědělství pokrývají všechny aspekty zemědělství, od zavlažování až po globální zemědělství, od systémů pro určování polohy (GNSS) až po zemědělské stroje, dobré životní podmínky zvířat, udržitelný rozvoj a řízení zemědělských podniků. Pomáhají prosazovat efektivní zemědělské metody a zároveň zajišťují, aby vše v dodavatelském řetězci – z pole na vidličku – splňovalo odpovídající úroveň bezpečnosti a kvality. Stanovením mezinárodně dohodnutých řešení globálních výzev, normy ISO pro zemědělství také podporují udržitelnost a zdravé životní prostředí, které přispívají k lepší budoucnosti" [34].

W3C

"World Wide Web Consortium (W3C) je mezinárodní konsorcium, které se zabývá vývojem webových standardů. Posláním W3C je vést World Wide Web k jeho plnému potenciálu tím, že vyvíjí protokoly a pokyny, které zajišťují dlouhodobý rozvoj webu [35]. Podle W3C je původním posláním Komunitní skupiny pro zemědělství (Agriculture Community Group⁴) shromažďovat a kategorizovat existující uživatelské scénáře z celého světa, které využívají webové rozhraní API a služby v zemědělském průmyslu a sloužit jako portál, který pomáhá vývojářům webu i uživatelům v oblasti zemědělství vytvářet chytřejší zařízení, webové aplikace a služby" [36].

⁴ <https://www.w3.org/community/agri/>

OASIS

Organization for the Advancement of Structured Information Standards⁵ (OASIS) je neziskové konsorcium, které řídí vývoj a přijímání otevřených standardů pro globální informační společnost. OASIS podporuje průmyslový konsensus a vytváří celosvětové bezpečnostní standardy pro cloud computing, SOA, webové služby, inteligentní sítě, elektronické publikování, nouzové systémy a další oblasti. Otevřené standardy OASIS nabízejí potenciál snížit náklady, stimulovat inovace, rozvíjet globální trhy a chránit právo na svobodný přístup k informacím [37].

IEEE

IEEE je největší světová profesní asociace, jejímž cílem je rozvoj technologických inovací a dokonalosti ve prospěch lidstva. IEEE a její členové inspirují globální komunitu prostřednictvím vysoce citovaných publikací, konferencí, technologických standardů a odborných a vzdělávacích aktivit IEEE. IEEE, vyslovováno "Eye-triple-E", je zkratka pro Institute of Electrical and Electronics Engineers (Institut elektrotechnických a elektronických inženýrů). Pod tímto názvem je asociace registrována a jedná se o plný právní název [38].

ISOBUS

Obsah normy ISOBUS spravovala skupina ISOBUS v rámci VDMA (Verband Deutscher Maschinen und Anlagenbau – Německý strojírenský svaz), což je síť přibližně 3 000 strojírenských podniků v Evropě a 400 odborníků z oboru. Norma ISOBUS specifikuje sériovou datovou síť pro řízení a komunikaci v lesnických nebo zemědělských traktorech. Skládá se z několika částí: obecná norma pro mobilní datovou komunikaci, fyzická vrstva, vrstva datového spoje, síťová vrstva, řízení sítě, virtuální terminál, aplikační vrstva implementačních zpráv, zprávy o pohonné jednotce, výměna dat řídicí jednotky traktoru, výměna dat řídicí jednotky a řídicího informačního systému, slovník mobilních datových prvků, diagnostika, souborový server. Práce na dalších částech stále probíhají. V současné době se jedná o normu ISO 11783 [39].

agroXML

agroXML je značkový jazyk pro zemědělskou problematiku, který poskytuje prvky a datové typy XML pro reprezentaci dat o pracovních procesech na farmě včetně doprovodných provozních prostředků, jako jsou hnojiva, pesticidy, plodiny apod. Je definován pomocí XML Schémat W3C. agroRDF je doprovodný sémantický model, který je v současné době stále ve fázi intenzivního vývoje. Je vytvořen pomocí RDF specifikace.

Zatímco existují další standardy pokrývající určité oblasti zemědělství, jako je např. datový slovník ISOBUS pro výměnu dat mezi traktorem a nářadím nebo ISOagriNet pro komunikaci mezi zařízeními pro chov hospodářských zvířat, účely agroXML a agroRDF jsou následující [40]:

- výměna informací mezi systémy v zemědělských podnicích a externími zúčastněnými stranami,
- dokumentace zemědělských procesů na vysoké úrovni,
- integrace dat mezi různými odvětvími zemědělské výroby,
- sémantická integrace mezi různými standardy a slovníky,
- prostředek pro standardizované poskytování údajů o provozních zásobách,

Z pohledu praktického využití standardů v zemědělství je tento jazyk velmi hojně používán pro popis jednotlivých komponent integrovaných systémů, především z pohledu výrobců zemědělské techniky.

OGC

Open Geospatial Consortium (OGC) je mezinárodní konsorcium více než 500 podniků, vládních agentur, výzkumných organizací a univerzit, které usiluje o to, aby geoprostorové informace a služby odpovídaly principu FAIR - Findable, Accessible, Interoperable a Reusable (vyhledatelné, dostupné, interoperabilní a pro opakované použití). Konsensuální proces OGC řízený členy vytváří veřejně dostupné geoprostorové standardy bez poplatků. OGC je na špičkové úrovni, aktivně analyzuje a předvídá vznikající technologické trendy a provozuje agilní, spolupracující laboratoř pro výzkum a vývoj (R&D), která vytváří a testuje inovativní prototypy řešení pro případy použití členů. Členové OGC společně tvoří

⁵ <https://www.oasis-open.org>

globální fórum odborníků a komunit, které využívají polohu k propojení lidí s technologiemi a ke zlepšení rozhodování na všech úrovních. OGC se zavazuje vytvořit udržitelnou budoucnost pro nás, naše děti a budoucí generace [41].

Pracovní skupina pro zemědělství (Agriculture Domain Working Group⁶) se zabývá otázkami technologií a technologické politiky se zaměřením na prostorové informace v zemědělství, a také na prostředky, kterými lze tyto otázky vhodně zohlednit v procesu vývoje norem OGC. Úkolem pracovní skupiny pro zemědělství je identifikovat problémy a výzvy v oblasti geoprostorové interoperability v zemědělství a poté zkoumat způsoby, jakými lze tyto výzvy řešit prostřednictvím použití stávajících standardů OGC nebo prostřednictvím vývoje nových standardů geoprostorové interoperability pod záštitou OGC. Úkolem pracovní skupiny pro zemědělství je sloužit jako fórum v rámci OGC pro zemědělskou geoinformatiku; předkládat, upřesňovat a zaměřovat otázky interoperability související se zemědělstvím technickému výboru; a případně sloužit jako styčný bod pro další průmyslové, vládní, nezávislé, výzkumné a normalizační organizace působící v oblasti zemědělství. Pro zajímavost uvádíme, že činnost této skupiny je koordinována z Čech [42].

INSPIRE

V Evropě je od roku 2007 v platnosti Směrnice Evropského Parlamentu a Rady 2007/2/ES ze dne 14. března 2007 o zřízení Infrastruktury pro prostorové informace v Evropském společenství (INSPIRE), která zřizuje infrastrukturu pro prostorové informace v Evropě na podporu politik Společenství v oblasti životního prostředí a politik nebo činností, které mohou mít dopad na životní prostředí. INSPIRE vychází z infrastruktur pro prostorové informace zřízených a provozovaných všemi členskými státy Evropské unie. Směrnice se zabývá 34 tématy prostorových dat potřebných pro environmentální aplikace, přičemž klíčové složky jsou specifikovány prostřednictvím technických prováděcích pravidel. Díky tomu je směrnice INSPIRE jedinečným příkladem legislativního "regionálního" přístupu [43].

5.3. STANDARDIZACE GEOPROSTOROVÝCH INFORMACÍ

Velké množství informací v zemědělství má prostorový charakter. Proto rozebereme standardizační aktivity v Geoprostorových informacích podrobněji. Jak je uvedeno v [44], otázky standardizace geoprostorových informací začaly s rostoucím úspěchem geografických informačních systémů (GIS) v 80. letech v rámci standardů a softwaru uzamčených dodavatelem. Mnoho standardů, které jsou implementovány v celém geoprostorovém průmyslu, bylo vyvinuto a publikováno OGC (viz výše). Standard OGC je dokument vytvořený na základě konsensu a schválený členy OGC, který stanoví pravidla a pokyny, jejichž cílem je optimalizovat stupeň interoperability v daném kontextu. Pravidla a pokyny jsou specifikovány s ohledem na požadavky komunity a členů, jakož i na tržní a technologické trendy. Normy OGC usilují o zlepšení interoperability geoprostorových dat, procesů, pracovních postupů, vizualizace, produktů, reprezentace a dalších aspektů správy geoprostorových informací.

OGC udržuje úzké formální vztahy s technickým výborem TC211 Mezinárodní organizace pro normalizaci (ISO), viz výše. Tento vztah umožnil, že řada norem OGC byla přijata jako normy ISO, například norma OGC Geography Markup Language (GML) byla přijata ISO jako ISO 19136:2007 Geographic information; Geography Markup Language (GML) a norma OGC Web Feature Service (WFS) byla přijata ISO jako ISO 19142:2010 Geographic information - Web Feature Service. OGC rovněž udržuje úzké formální vztahy s konsorciem World Wide Web Consortium (W3C). Tento vztah vedl k vývoji řady specifikací, například časové ontologie v OWL, sémantické ontologie sítě senzorů a specifikace osvědčených postupů pro prostorová data na webu. Tyto specifikace spravuje zájmová skupina Spatial Data on the Web Interest.

Geoprostorová data se používají v mnoha oblastech, aby se zlepšilo porozumění charakteristikám místa a jejich dopadu na různé aspekty společnosti. Kromě geoprostorových technologií pro získávání dat umožňují geografické informační systémy (GIS) shromažďovat data a v rámci plánované analýzy propojovat informace o jevech reálného světa prostřednictvím jejich umístění na Zemi. Od sběru dat až po rozhodování vyplývající z analýzy v rámci GIS existují standardy, které umožňují bezproblémovou a efektivní práci v rámci pracovního postupu představujícího plánovanou analýzu.

⁶ <https://www.ogc.org/projects/groups/agriculturedwg>

Pokud jsou v rámci určité domény široce přijaty zásady FAIR (popsané výše), je tento pracovní postup efektivní, stejně jako pracovní postupy napříč doménami.

Snaha o interoperabilitu dat, která stojí za vývojem norem, může být v závislosti na oblasti velmi náročná. Strukturální a syntaktická shoda na interoperabilitě může záviset na úspěšnosti stávajících formátů a norem pro výměnu dat v rámci oboru, komunity nebo přímo na metodách sběru dat, např. na schopnostech zařízení pro sběr dat. Překonat sémantickou překážku za účelem dosažení doménové interoperability je však obtížnější úkol. Jinými slovy, starší formát dat a osvědčené postupy v rámci oboru se často přiklání k určitému standardnímu rámci, ale různé země nebo mírně odlišné komunity využívající stejná tematická data mohou používat konkurenční sémantické nebo ontologické rámce, které je třeba sblížit, aby bylo dosaženo interoperability. Všimněte si, že tato konvergence pro interoperabilitu dat znamená pouze dosažení kompatibilních standardů, tj. interoperability, nikoli nutně jejich sjednocení. Vývoj evropských specifikací tematických dat INSPIRE byl možný díky rámci a metodice INSPIRE, např. díky pracovním skupinám, expertům atd. ale s různým pokrokem v závislosti na tematických datech. Dosažení harmonizace dat a FAIRness je podmíněno mírou přijetí standardů a datových specifikací, čímž se měří jeho úspěšnost.

Výměna dat o půdě je dobrým příkladem tohoto někdy zdoluhavého a mnohostranného procesu. V roce 2010 byl navržen značkovací jazyk pro půdní data nazvaný SoilML [45]. Ve [46] byl integrován různý přístup k půdním databázím Soil and Terrain (SOTER) Databases, klasifikace Organizace pro výživu a zemědělství (FAO) z roku 1988 a World Reference Base of Soil Resources (WRB) do značkovacího jazyka SOTER (SoTerML) umožňujícího flexibilní charakterizaci, který navrhoval určité mapování mezi různými klasifikacemi. SoTerML byl navržen tak, aby byl kompatibilní se standardem Geography Markup Language (GML), standardem OGC, na kterém již byl postaven GeoSciML pro geologická data. V roce 2013 byla vytvořena specifikace půdních dat INSPIRE a v témže roce byl vyvinut značkovací jazyk pro globální půdní mapy

(GSMML). V roce 2013 také Mezinárodní organizace pro normalizaci (ISO) zveřejnila normu ISO 28258:2013 pro digitální výměnu dat týkajících se půdy. Cílem normy ISO 28258:2013 je usnadnit výměnu dat týkajících se půdy prostřednictvím digitálních systémů mezi subjekty, jako jsou jednotlivci a organizace. Norma ISO 28258:2013 rovněž využívá normu GML k reprezentaci prvků a k explicitnímu georeferencování půdních dat, čímž usnadňuje používání půdních dat v rámci GIS. V roce 2016 byl proveden OGC Soil Data Interoperability Experiment, který měl ilustrovat řadu různých používaných datových modelů, včetně přístupu SoTerML a INSPIRE [47]. Tento OGC Soil Data Interoperability Experiment vyhodnotil stávající modely a navrhl společný základní model včetně schématu založený na GML, který byl testován prostřednictvím nasazení webových služeb OGC a demonstračních klientů.

Zdroj [48] poskytuje přehled různých standardizačních aktivit týkajících se půdy a zemědělství. Jejich studie zdůrazňuje, že i přes existenci standardů na podporu půdoznalství a zemědělství nadále existují rozdíly mezi modely přijatými různými zájmovými komunitami. Pro zlepšení interoperability mezi různými komunitami je potřeba společných aktivit mezi [49, 50]:

- Pracovní skupinou OGC pro oblast zemědělství,
- FAO (Globální partnerství pro půdu, tj. pilíř 4 s GLOSIIS jako elektronickou infrastrukturou)⁷,
- pilíř 5 FAO pro harmonizační aspekty včetně standardu pro výměnu dat o půdních profilech,
- pracovní skupina ISO Mezinárodní unie věd o půdě - Půdní informační standardy (IUSS WG-SIS).

Všimněte si zde častého použití jazyka XML jako podpory modelování dat zapouzdřujícího specifikaci dat, tedy větší sémantické interoperability. To platí pro data o půdě a pro řadu environmentálních dat používaných v zemědělství. Pro distribuci geoprostorových informací po internetu se však ve velké míře používají i jiné serializační syntaktické formáty, např. formát RDF, JSON v podobě propojených dat. I když teoreticky lze přejít z jedné serializace na druhou,

⁷ <http://www.fao.org/global-soil-partnership/pillars-action/4-information-data/glosis/en/>

různé vývojářské komunity dávají přednost té či oné, což více či méně usnadňuje jejich přijetí a vývoj.

Zdroj [44] nadále hovoří o tom, že v dílčích oblastech zemědělství, jako je zemědělská ekonomika, zemědělství na orné půdě, chov skotu, protierozní ochrana, smíšené zemědělství, řízení úrodnosti půdy a zlepšování půdy, hraje geoprostorová složka klíčovou roli při přiřazování atributových informací k místu na Zemi. Takové atributové informace mohou zahrnovat například genetické údaje, genomické údaje a údaje o půdě. Zatímco atributové informace mohou být shromážděny z pole nebo farmy, v jiných případech mohou být výsledkem vědeckého modelování a simulace, například v případě modelování růstu rostlin. Vývoj datových služeb a e-infrastruktury umožňující každému partnerovi multidisciplinárního projektu vizualizovat zdroje dat ostatních je často hnacím motorem pro integraci dat vytvářející motivaci v rámci týmů [51]. Výsledkem je, že strategie a analytika pro vizuální zkoumání komplexních environmentálních dat jsou průvodci mezioborových inovací. Geoprostorová doména strukturující požadované atributové informace spolu s volbou vhodných časoprostorových měřítek umožňuje partnerům předvídat nepředvídatelnou provázanost takových multidisciplinárních informací. Tyto role také usnadňují učení a výuku v oblasti zemědělství a souvisejících environmentálních studií [52]. Mezi standardy vyvinuté a zveřejněné OGC, které se ukázaly jako účinné nástroje umožňující mezioborovou výměnu zemědělských informací, patří dle [44] například:

- **OGC Sensor Observation Service (SOS):** Tento standard specifikuje rozhraní webové služby pro vyžádání, filtrování a vyhledávání pozorování a informací o sensorových systémech. Jedná se o zprostředkovatele mezi klientem a úložištěm pozorování nebo kanálem senzorů pracujícím téměř v reálném čase.
- **OGC SensorThings API (STA):** Tento standard poskytuje otevřený, RESTful, geoprostorový, na JSON založený a jednotný způsob propojování a zadávání úloh zařízením, datům a aplikacím internetu věcí prostřednictvím webu.
- **OGC Sensor Model Language (SensorML):** Tento standard definuje modely a schémata pro popis sensorových systémů a procesů; poskytuje informace potřebné pro

zjišťování senzorů, umístění pozorování senzorů, zpracování nízkourovňových pozorování senzorů a výpis vlastností, které lze zadat.

- **Schéma OGC pro pozorování a měření (O&M):** Tato norma definuje koncepční modely pro kódování pozorování a měření ze senzorů, a to jak archivovaných, tak v reálném čase. Přijata Mezinárodní organizací pro normalizaci (ISO) jako ISO 19156:2011.
- **OGC Sensor Planning Service (SPS):** Tato norma specifikuje rozhraní webové služby pro zadávání požadavků na akvizice a pozorování řízené uživatelem. Tato norma umožňuje zadávat úkoly senzorům.
- **OGC Geography Markup Language (GML):** Tato norma specifikuje gramatiku rozšiřitelného značkovacího jazyka (XML) pro vyjádření geografických prvků. GML slouží jako modelovací jazyk pro geografické systémy a také jako otevřený výměnný formát pro geografické transakce na internetu.
- **OGC Web Map Service (WMS):** specifikuje rozhraní, které poskytuje jednoduché rozhraní HTTP pro vyžádání geolokalizovaných mapových obrazů (rastrů) z jedné nebo více distribuovaných geoprostorových databází.
- **OGC Web Coverage Service (WCS):** specifikuje rozhraní, které nabízí vícerozměrná data pokrytí pro přístup přes internet.
- **OGC Web Feature Service (WFS):** specifikuje rozhraní, které nabízí přímý přístup ke geografickým informacím na úrovni prvků a vlastností prvků prostřednictvím operací vyhledávání, operací dotazování, operací zamykání, transakčních operací a operací pro správu uložených parametrizovaných dotazovacích výrazů.
- **OGC Web Processing Service (WPS):** specifikuje rozhraní, které poskytuje pravidla pro standardizaci způsobu vstupů a výstupů (požadavků a odpovědí) pro služby geoprostorového zpracování, jako je například překrývání polygonů.
- **Katalogové služby OGC pro web (CSW):** specifikuje rozhraní, které podporuje možnost publikovat a prohledávat sbírky popisných informací (metadat) pro data, služby a související informační objekty.
- **OGC WaterML:** umožňuje reprezentaci dat z pozorování vody se záměrem umožnit výměnu takových datových souborů napříč informačními systémy.

5.4. PROPOJENÍ S NÁRODNÍMI NORMAMI

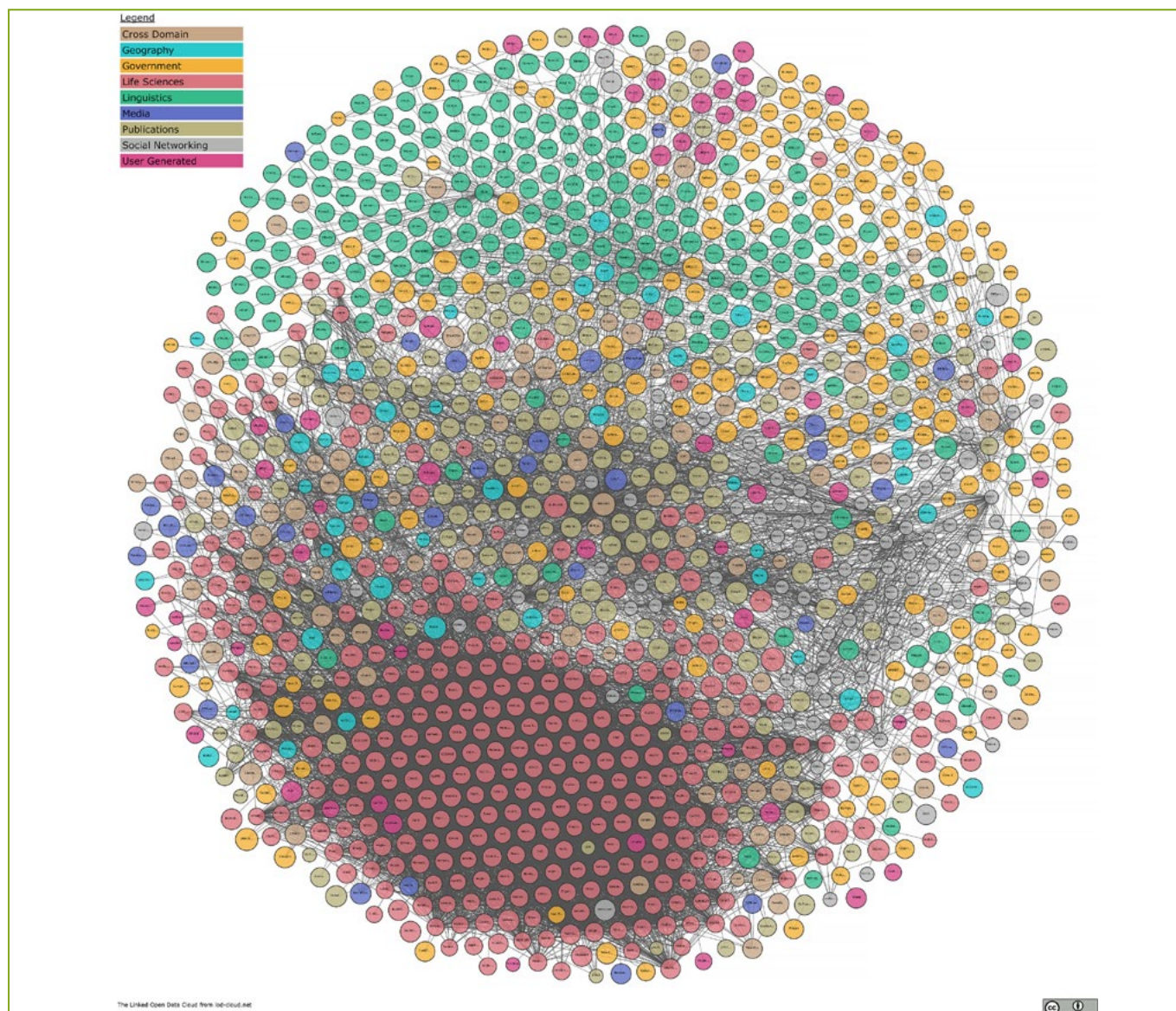
Přítomnost a využívání cloud computingu (na internetu založený model vývoje a používání počítačových technologií) vedly k vytvoření neuvěřitelného množství strukturovaných i nestrukturovaných geografických dat, která nyní představují přínos nebo výzvu, v závislosti na způsobu jejich využití. Dle [44] taková data vyžadují přijetí používaných standardů, aby se maximalizoval jejich inovační potenciál jako podklad pro rozhodování na národní, regionální a kontinentální úrovni. Pochopení vývoje norem je pak zásadní, a to zejména pro výrobní odvětví, kterým zemědělství je.

ISO/TC 211 je běžně přijímanou národní normou, která se snaží přijmout a přispívat do rodiny mezinárodních norem, které podporují porozumění a využívání geografických informací s cílem zvýšit dostupnost, přístup, integraci a sdílení geografických informací. Interoperabilita počítačových systémů s geoprostorovými informacemi umožňuje jednotný přístup k řešení globálních ekologických a humanitárních problémů prostřednictvím geoprostorových e-infrastruktur, které přispívají k udržitelnému rozvoji, čímž se plní závazky zemí k cílům udržitelného rozvoje.

Koordinovaný rozvoj, který je zde naznačen, vyžaduje porozumění organizacím, které využívají geopolitické rozhodování na národní, regionální a kontinentální úrovni. To pomáhá podporovat neomezený tok digitálních prostorových informací, které využívají otevřené systémy pro poskytování dat v rámci poměrně rychle se vyvíjejících odvětví. Po vytvoření organizačního rámce se v rámci kooperativního paradigmatu výzkumu a vývoje rozvíjejí způsoby sdílení sektorových informací, které podporují administrativní a technickou podporu v rámci technologicky sladěných paradigmat. To podporuje přenos výhod plynoucích soukromému sektoru pro usnadnění přijetí a doprovodných investic ze strany soukromého sektoru. Realizace tohoto cíle však není možná, pokud vlády nepovedou ke spotřebě řešení GIS. Demonstrace využívání inovací je nejlepší, když je vedena vládami, zejména v oblastech komunikace v reálném čase odvozené od datové infrastruktury. To pak podporuje vytváření tržní poptávky, kterou soukromý sektor pouze posiluje a přebírá vedoucí

úlohu, když jsou stanoveny komerční přínosy s využitím vládní poptávky po službách a souvisejících standardech. Je také důležité uznat, že je třeba rozšířit rozvoj kapacit (zejména v rozvojových zemích). To pomáhá při integraci norem OGC/ISO s regionálními a národními normami. Směrnice INSPIRE identifikuje řadu témat, která tvoří rámec pro prostorové informace v Evropě. Jedním z těchto témat je "Zemědělská a akvakulturní zařízení". Aby členské země mohly vyhovět směrnici INSPIRE, vedla Evropská komise iniciativy k vypracování technických pokynů pro specifikace dat na základě jednotlivých témat.⁸

⁸ Specifikace údajů INSPIRE pro zemědělská a akvakulturní zařízení – technické pokyny naleznete na adrese <http://inspire.ec.europa.eu/Themes/137/2892>.



Obrázek 9: Linked Open Data cloud diagram [56].

5.5. OTEVŘENÁ PROPOJENÁ DATA – LINKED OPEN DATA

Přístup Linked data (označovaná jako propojená data) vznikl před více než deseti lety. Jejich princip spočívá ve využívání technologií, standardů (organizací W3C, OGC a dalších) a postupů známých z prostředí webu pro publikování strukturovaných dat, nikoli pouze webových stránek ve formě víceméně volného textu [28]. Jedná se o přístup, kdy jsou informace uložené na webu strukturovány a uloženy podle standardizovaných pravidel, což usnadňuje jejich vyhledání a zpracování. Tohoto přístupu s úspěchem využívají tzv. otevřená data, přičemž pak hovoříme o Linked Open Data, tedy o otevřených propojených datech.

Ani přes poměrně důležitý podíl prostorových dat v Linked Open Data cloud diagramu (tyrkysově podbarvené položky v Obrázku 9, aktuální ke květnu 2020) není možné najít velké množství skutečných implementací propojených dat v oblasti geoinformačních technologií, potažmo v zemědělství. A to i přes nesporné výhody propojených dat, kterými jsou například zavedení jednoznačných

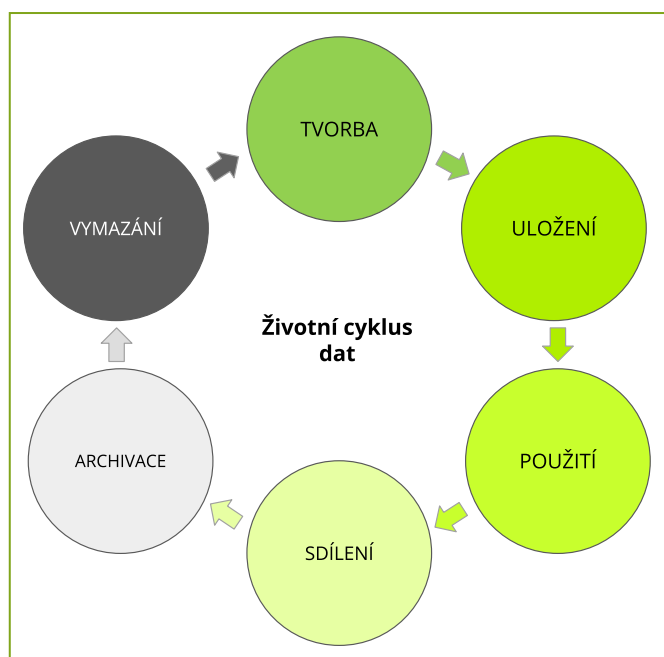
identifikátorů pro jednotlivé objekty prostorových dat (jeden objekt či jev je definován jednotně a při jeho různých použitích je chápán vždy stejně), požadavek na jasné deklarování licence, nutnost publikování dat na internetu, využívání neproprietárních formátů, doplnění explicitní sémantiky nebo propojení na externí datové zdroje. Princip otevřených propojených dat navíc vede k provazování různých (databázových) systémů mezi sebou, čímž zásadně přispívá k vysoké míře interoperability [53].

Splnění všech pěti stupňů otevřenosti dat, viz kapitola 4. Otevřená a sdílená data definují tzv. pětihvězdičková Linked Open Data. Ukázky implementace principu otevřených propojených dat a jejich aplikace a využití v zemědělství najdete v kapitolách 16.1. DataBio a nástroje velkých dat, 9.1. FOODIE data model, 9.2. Otevřená data využití půdy – Open Land Use, 9.3. Chytré body zájmu – Smart Point of Interest uvedené v této knize dále. Více o technologickém pozadí Linked data v publikacích [28, 54, 55].

6. BEZPEČNOST DAT A JEJICH OCHRANA

Ochrana a anonymizace dat je v oblasti velkých dat a datové analytiky velkým tématem. Každým rokem se shromažďuje stále větší množství citlivých informací o konkrétních osobách z různých zdrojů dat, jako jsou sociální sítě, aplikace mobilních telefonů a systémy elektronických lékařských záznamů atd. Analýza takového množství a objemu dat nabízí vlastníkům údajů pozoruhodné možnosti, ale zároveň vyžaduje použití nejmodernějších řešení pro ochranu osobních údajů a uplatňování právních předpisů o ochraně osobních údajů, aby byla zaručena důvěrnost údajů o osobách, které jsou v údajích zastoupeny. Ochrana údajů je sice důležitá při vývoji každého moderního informačního systému, ale v souvislosti s rozsáhlým zpracováním citlivých údajů se stává klíčovou.

Vedle záruk ochrany soukromí se další důležitý aspekt týká schopnosti přístupů k ochraně dat a zachování užitečnosti datových souborů, na které jsou aplikovány s cílem podpořit různé typy datových analýz. Řešení ochrany soukromí, která nabízejí záruky a zároveň



Obrázek 10: Životní cyklus dat (upraveno dle [31]).

zachovávají vysokou užitečnost dat činí z technologie ochrany soukromí klíčový faktor umožňující aplikaci analýz na vlastní a potenciálně citlivá data. [32]

Abychom se mohli zabývat zabezpečením dat, je třeba identifikovat různé fáze životního cyklu dat a jejich ochranu a bezpečnost v rámci dané fáze. Bezpečnost dat se tedy odráží od jejich vytvoření, přes jejich používání, sdílení, archivaci až po vymazání. Bezpečné nakládání s daty v průběhu jejich životního cyklu vytváří základy strategie ochrany především citlivých dat. Tyto fáze životního cyklu dat jsou zobrazeny na Obrázku 10 níže a následně jsou jednotlivé fáze popsány.

Životní cyklus dat obsahuje tyto jednotlivé fáze [31]:

1. Tvorba

Tato první fáze zahrnuje vytvoření jak strukturovaných, tak i nestrukturovaných či surových (nezpracovaných) dat. Z pohledu ochrany citlivých dat je možné hovořit o následujících kategoriích:

- podniková data – komerčně citlivá data,
- osobní data – osobní citlivá data podléhající právním předpisům o ochraně osobních údajů,
- ostatní data – taková, která nespadají do žádné z předchozích kategorií.

Zejména u podnikových dat dochází již ve fázi vytváření k jejich bezpečnostní klasifikaci na základě jejich bezpečnostní politiky.

2. Uložení

Jakmile jsou data vytvořena a zahrnuta do souborové struktury, jsou někde uložena. Je třeba zajistit, aby uložena data byla chráněna, aby byly zavedeny potřebné kontroly zabezpečení dat a aby bylo zajištěno a minimalizováno riziko úniku informací při zajištění účinné ochrany dat. V této fázi by měla být řešena i možnost obnovy dat, při jejich neúmyslné ztrátě.

3. Použití

V této fázi, kdy jsou data různě prohlížena, zpracovávána, upravována a ukládána v rámci procesu jejich užití, tak se na data přímo aplikují bezpečnostní kontroly se

zaměřením na sledování činnosti uživatelů a použití bezpečnostních kontrol k zajištění prevence úniku dat. V této fázi se také uplatňuje stupeň otevřenosti používaných dat, popsany v kapitole 4. Otevřená a sdílená data.

4. Sdílení

Data jsou neustále sdílena mezi zaměstnanci, zákazníky a partnery, což vyžaduje strategii, která nepřetržitě monitoruje datová úložiště a uživatele. Data se pohybují mezi různými veřejnými a soukromými úložišti, aplikacemi a operačními prostředími a přistupují k nim různí vlastníci dat z různých zařízení a platforem. To se může stát v kterékoli fázi životního cyklu zabezpečení dat, a proto je důležité použít správné bezpečnostní kontroly ve správný čas.

5. Archivace

V případě dat, která již nejsou aktivně používána, ale stále musí být k dispozici, by měla být bezpečně uložena ve vhodných úložištích, obvykle s nízkými náklady na provoz, někdy i off-line. To se může týkat i verzování (nejen) dat, kdy jsou jejich starší verze také uchovávány v archivních úložištích jako zálohy jejich nejaktuálnějších verzí. Tyto zálohy jsou pak uloženy a mohou být znovu uvedeny do provozu v přiměřeném časovém rámci, který zajistí, že nedojde ke škodlivému dopadu ze ztráty nebo poškození dat.

6. Vymazání

V případě dat, která již nejsou potřeba, by mělo dojít k jejich bezpečnému vymazání, aby se předešlo jakýmkoliv problémům spojeným s únikem těchto dat.

6.1. BLOCKCHAIN

Pod pojmem blockchain si v první řadě představíme kryptoměny, ale možností aplikací této technologie je mnohem více. Jedinečnou funkcionalitou blockchainu je dohledatelnost a autenticita veškerých provedených operací. Každá transakce, která projde blockchainem, dostane unikátní otisk (kontrolní součet), který je součástí řetězce pro výpočet otisku následné transakce. Každý záznam tak v sobě obsahuje svým způsobem i otisky veškeré historie transakcí provedených v rámci blockchainu. Kvůli praktické využitelnosti se transakce řetězí do bloků a pro zajištění autenticity transakcí jsou počítány i kontrolní součty bloků, u historických transakcí není potřeba je ověřovat jednu po druhé, stačí ověřit autenticitu bloku, kterého je součástí. Lze si to představit jako jednotlivé stránky účetní knihy, kdy po zaplnění všech řádků pro transakce, se zůstatek z předchozí stránky přenese na novou stránku. Tímto mechanismem je zaručeno, že uskutečněné transakce je nemožné zpětně zfalšovat, při jejich verifikaci by se přišlo na to, že kontrolní součty neodpovídají. [32]

Praktických aplikací výše popsaného principu je nepřeberné množství, v zemědělství např. při zajištění autenticity zemědělské produkce, kdy producent zadá do blockchainu množství své produkce a logika blockchainu již nedovolí, aby mohlo dodavatelsko-odběratelským řetězcem protéci větší množství produktu, než do něj na začátku výrobce zadal. Jsou tím chráněny všechny strany – producent si může být jistý, že nikdo nezneužije jeho značku, obchodník, že prodává autentický produkt od jeho výrobce, popř. autorizovaného distributora a zákazník, že dostává produkt, za který si zaplatil, žádný padělek.



7. NÁVRATNOST NÁKLADŮ (ROI NEBOLI RETURN OF INVESTMENT)

Jak bylo již ukázáno, znalostní řízení a nové technologie budou nezbytné pro budoucí udržitelné zemědělství. Znalostní řízení bude třeba implementovat nejen na farmách, ale i v potravinářském sektoru a ve veřejné správě.

Technologie se ukazují být dokonalým spojencem, který přispívá k tomu, aby bylo možné těmto výzvám adaptovat zemědělství na všechny externí faktory a zmírnit jejich vliv na efektivní fungování zemědělského odvětví, zejména pokud jde o rozhodování prostřednictvím využívání informací s přidanou hodnotou. Obecně lze říci, že technologickými výzvami jsou transformace odvětví na zemědělství založené na datech; dopad interoperability a analýzy velkých dat; vývoj standardů pro snadné a rychlé nasazení technologií; vývoj technologií velkých dat a strojového učení pro otevřená data; a zapojení zemědělců a dalších subjektů do řetězce přidané hodnoty dat. Dále je nutné vytvoření scénáře spolupráce, který umožní služby bohaté na data skutečně prakticky využívat.

Implementace a adaptace technologií pro udržitelné zemědělství je finančně náročnou otázkou. Vznikají náklady spojené s používanými technickými prostředky, jako jsou drony, senzory, data EO, roboti a další. Dále se jedná o náklady, které se objevují při vývoji technologií pro přesné zemědělství. Je nutno zahrnout finanční možnosti zemědělců, náklady na seznámení s nástroji digitalizace a na druhé straně obecně dosažení hmatatelných výsledků týkajících se přínosů používání těchto nástrojů.

Jak již bylo zmíněno, zemědělci využívají mnoho různých zdrojů dat a informací, které vyžadují individuální, ale vzájemně propojené procesy, jako je shromažďování, ukládání, sdílení a analýza, aby poskytly přidanou hodnotu z velkých množství nezpracovaných dat, která jsou denně generována. Za tímto účelem vzniklo precizní zemědělství, které má pomocí dat podpořit rozvoj, růst a udržitelnost tohoto odvětví.

Velká data ovlivňují zemědělství výrazným způsobem a jsou významným trendem z hlediska investic a přidané hodnoty v zemědělsko-potravinářském odvětví, protože mohou zajistit efektivitu celého hodnotového a dodavatelského řetězce tím, že umožňují předpovědi v zemědělství, provozní rozhodnutí v reálném čase, inovativní obchodní procesy a definici nových

obchodních modelů. V tomto smyslu můžeme definovat technologický tlak (technology push) a požadavky (pull faktory), které pohánějí zavádění Big Data pro zemědělský sektor:

1. Technologický tlak:

- Technologický rozvoj (internet věcí, nástroje založené na datech, precizní zemědělství, agrotechnické společnosti, sledovatelnost na bázi RFID, pokročilý dálkový průzkum).
- Sofistikovanost technologií (EO, globální navigační satelitní systém, roboti, GPS jednotky na traktorech, bezpilotní letadla, senzory).
- Tvorba a ukládání dat (zpracování strojově a lidmi generovaných dat, interpretace nestrukturovaných dat, pokročilá datová analytika).
- Digitální konektivita (zvýšená dostupnost, zvýšení výpočetního výkonu, Edge Computing).
- Možnosti inovací (otevřená správa zemědělských podniků prostřednictvím aplikací, návrhy nebo rozhodnutí na dálku/pomocí počítače, regionální otevřená data pro výzkum, elektronické obchodování (B2B/B2C)).

2. Požadavky:

- Podnikatelské faktory (zvýšení efektivity, nižší náklady na jednotku, lepší tržní cena, zlepšení kontroly řízení a rozhodování, posílení podpory místního řízení, digitalizace, čelit výzvám souvisejícím s počasím kvůli nestálosti podmínek).
- Veřejné hnací síly (bezpečnost a zabezpečení potravin a výživy, udržitelnost).
- Potřeba více a lepších informací.

Celkově lze říci, že technologie velkých objemů dat (Big Data Technology, BDT) je nové technologické paradigma, které je hnacím motorem celé ekonomiky, včetně odvětví s nízkou technologickou náročností, jako je zemědělství, kde je implementováno pod hlavičkou precizního zemědělství. Jak již bylo řečeno, zemědělci se primárně nezaměřují na (velká) data, ale na znalosti a rozhodnutí z těchto dat generované [57].

V diskusi o návratnosti investic je třeba porovnat právě oba aspekty - technologický tlak a reálné požadavky trhu. Často se stává, že nové technologie jsou spíše tlakem technologických poskytovatelů než přání zákazníků. Z toho také někdy vyplývají přehnaně optimistická očekávání technologických poskytovatelů.

Ukažme si případ realistického odhadu potenciálu trhu v ČR. Je třeba zmínit, že skladba českých zemědělských podniků je pro zavádění precizního zemědělství poměrně dobrá. V České republice je více než 50 % obhospodařované půdy využívané zemědělskými podniky s celkovou výměrou nad 1000 ha. Přibližně 68 % obhospodařované půdy je využíváno zemědělskými podniky s celkovou výměrou nad 500 ha. Z hlediska velikosti polí pozemky nad 5 ha orné půdy tvoří přes 90 % orné půdy. Z tohoto údaje je zřejmé, že Česká republika má optimální nebo jedny z nejlepších podmínek v Evropě pro využití precizního zemědělství. Z tohoto důvodu je důležité, aby naše analýza nebyla příliš optimistická v odhadu potenciálu trhu pro určité služby.

Dominantní plodinou v České republice je pšenice ozimá. V roce 2019 se pšenice ozimá pěstovala na 814 tisících hektarech, což je 33 % české orné půdy (orná půda je podmnožinou zemědělské půdy). V pilotním testování bylo prokázáno, že hnojení na základě doporučení založeného na zpracování dat DPZ může zlepšit účinnost hnojení o 8 %. Ta spočívá jak ve snížení spotřeby hnojiv, tak ve zvýšení výnosu. Snížení spotřeby hnojiv výrazně převažuje. Budeme jednoduše analyzovat variantu, kdy nedošlo ke zvýšení výnosu, ale pouze ke snížení množství hnojiva. Z dostupných statistik vyplývá, že náklady na hnojiva činí přibližně 7 070 Kč (278 €) na hektar. Úspora 8 % hnojiv v případě ozimé pšenice znamená úsporu přibližně 565 Kč na hektar (22,1 EUR). To znamená, že rozhodovací prostor pro cenu poradenské služby, která spočívá ve snížení spotřeby hnojiv o 8 % je teoreticky až 22,1 EUR na hektar, tedy obecně 2,76 EUR na procento ušetřených hnojiv. Vyšší cena by byla ekonomickým nesmyslem, protože náklady na službu by byly vyšší než úspora nákladů na hnojiva. Pokud by maximální cena poradenské služby, kterou je zemědělský podnik ochoten akceptovat, činila 25 % ušetřených nákladů, v případě 8% snížení spotřeby hnojiv je to asi 5,5 € na hektar. Tato hodnota je optimistickým odhadem. Vezmeme-li v úvahu pesimistický odhad, že zemědělec je schopen akceptovat 10 % ušetřených nákladů na hnojiva jako cenu služby, byla by maximální cena služby přibližně 2,2 € na hektar. Hnojení podle výnosového potenciálu (na základě dat z dálkového průzkumu Země z Landsat 8 a Sentinel 2) je teoreticky využitelné pro pozemky nad 5 ha, ale skutečně výhodné je pro pozemky nad 10 ha. Při zjed-

nodušeném předpokladu, že struktura výměry pozemků s ozimou pšenicí je stejná jako celková struktura zemědělské půdy v ČR, je tento typ poradenství teoreticky vhodný pro 80 % ploch s ozimou pšenicí, což je přibližně 651 tis. ha, s celkovou cenou za služby ve výši 3 580 000 € (optimistický odhad), pokud by jedna poradenská firma teoreticky získala 100 % trhu. Při velmi zjednodušeném předpokladu, že náklady na hnojení ostatních plodin jsou stejné jako u pšenice a že metoda je použitelná na 80 % orné půdy, je celkový objem tohoto typu poradenských služeb na českém trhu přibližně 10 850 000 €, pokud by zemědělec byl ochoten akceptovat 25 % úspory jako cenu, a 4 340 000 €, pokud by akceptoval 10 %. Tato jednoduchá čísla ukazují, že mnohá očekávání organizací poskytujících služby v této oblasti jsou velice optimistická. [58]

Jaká je tedy cesta k řešení tohoto problému? Je nutné podporovat spolupráci všech aktérů z veřejného i soukromého sektoru. K budování společného prostředí přizve také potravinářský průmysl, strojírenství, chemický průmysl, IT průmysl, finanční organizace. Složitost tohoto problému bude vyžadovat kombinaci veřejného a soukromého financování. Je důležité vzít v úvahu, že úlohou zemědělství není pouze zajišťovat potraviny, ale také produkovat veřejné statky. Jako příklady veřejných statků lze uvést utváření krajiny, kvalitu a dostupnost vody, biologickou rozmanitost zemědělské půdy, funkčnost půdy a stabilitu klimatu (zvyšování ukládání uhlíku a snižování emisí skleníkových plynů). Zemědělství má ve venkovských regionech také sociální roli. To znamená, že některé potenciální náklady je třeba pokrýt z veřejných zdrojů. Produkce kvalitních zemědělských produktů je i zájmem potravinářského průmyslu a ve světě je známá řada příkladů, kdy potravinářské firmy financují zavádění precizního zemědělství. Výsledky analýz jsou relevantní i pro finanční sektor, pro banky mohou poskytovat údaje o situaci na trhu, a pojišťovněm informace o škodách a pomáhat potenciální škody eliminovat. Takovéto sdružování požadavků bude pro budoucnost pravděpodobně jedinou cestou, jak plně využít potenciálu velkých dat pro potřeby zemědělství. Zde je ale třeba i připomenout, že nutným předpokladem je vybudovat bezpečné mechanismy pro kontrolu a sdílení dat a i interoperabilitu, jak již bylo uvedeno v předchozích kapitolách. [9, 59]



8. CO JSOU METADATA A OPRAVDU JE POTŘEBUJEME?

Metadata jsou, jak už název napovídá, data o datech. Popisují vlastnosti datové sady nebo zdroje. Metadata mohou zahrnovat různé typy informací, které lze podle [NISO04] zhruba rozdělit do tří kategorií [60]:

- popisná metadata zahrnují prvky, jako je název, abstrakt, autor a klíčová slova, a většinou se používají k objevení a identifikaci datové sady nebo jiného zdroje;
- strukturální metadata, která udávají, jak jsou složeny objekty sestaveny (logické nebo fyzické vztahy mezi objekty a jejich částmi);
- administrativní metadata s prvky, jako je licence, práva duševního vlastnictví, kdy a jak byla datová sada vytvořena, kdo k ní může přistupovat atd.

Častou otázkou je, zda zemědělec metadata potřebuje a k čemu mu vlastně mohou být. V následujícím textu bude stručně vysvětlen význam metadat.

Metadata jsou obvykle považována za něco nudného, nezáživného, nepotřebného. A přitom se s metadaty setkáváme v každodenním životě. Pokud si třeba jdeme do supermarketu koupit nějakou potravinu, jako např. uzeninu nebo mléčný výrobek, zajímá nás původ, složení datum výroby, datum spotřeby a podobně. To jsou metadata v každodenním životě.

Zkusme se zamyslet nad tím, jak používat metadata při správě velkých dat v zemědělství. Jak již bylo řečeno, zemědělec sám obvykle velká data zpracovává a nevyhledává. Metadata ale potřebuje určitě zpracovatel, aby věděl, s jakými údaji pracuje, kdy data vznikla, kdo je připravoval a podobně. Metadata jsou jedním ze základních nástrojů pro zavádění FAIR principů. Pro uživatele pak mají význam zjednodušená metadata věnovaná zpracovanému výsledku, která mu mohou pomoci pochopit, jak potřebná znalost vznikla. Vlastní technologie pro metadata jsou skutečně dosti technické a pro většinu uživatelů i nudné. Naopak poskytovatelé služeb a software by měli této problematice rozumět. Proto doporučuji koncovým uživatelům přeskočit následující podkapitulu 8.1. Technologie pro metadata, která je poměrně dost technická a zaměřit se na závěrečnou podkapitulu 8.2. (Ne)viditelná metadata.

8.1. TECHNOLOGIE PRO METADATA

Datové sady v zemědělství jsou buď přidávány lokálně, uživatelem, získávány z existujících datových portálů, nebo získávány z provozních systémů či ekosystémů internetu věcí. Definice souboru prvků metadat je nezbytná, aby bylo možné identifikovat obrovské množství spravovaných informačních zdrojů, pro které jsou metadata vytvářena, jejich klasifikaci a identifikaci jejich geografické polohy a časové reference, kvality a platnosti, souladu s prováděcími pravidly pro interoperabilitu sad prostorových dat a služeb, omezení souvisejících s přístupem a používáním a organizace odpovědné za daný zdroj.

Kromě toho jsou prvky metadat týkající se samotného metadatového záznamu nezbytné také pro sledování toho, zda jsou vytvořená metadata aktualizována, a pro identifikaci organizace odpovědné za vytváření a údržbu metadat. Takový minimální soubor prvků metadat je rovněž nezbytný pro dosažení souladu se směrnici 2007/2/ES a nebrání organizacím v rozsáhlejší dokumentování informačních zdrojů pomocí dalších prvků odvozených z mezinárodních norem nebo pracovních postupů v jejich zájmové komunitě.

Metadata datových sad a časových řad datových sad (pro zemědělství jsou zvláště důležité produkty DPZ odvozené ze satelitních snímků) by měla být v souladu s nařízením o metadatech INSPIRE, v případě potřeby s přidávanými tematickými metadatovými prvky pro oblast zemědělství, lesnictví a rybolovu. Tento přístup zajistí, že metadata vytvořená pro datové sady, datové řady a služby budou v souladu s požadavky INSPIRE i s mezinárodními standardy. Kromě toho lze metadata odpovídající směrnici INSPIRE vyjádřit také prostřednictvím aplikačního profilu DCAT, který definuje minimální soubor prvků metadat pro zajištění mezidoménové a přeshraniční interoperability mezi schématy metadat používanými na evropských datových portálech. Takové mapování by mohlo podpořit začlenění metadat INSPIRE do celoevropského portálu otevřených dat za účelem širšího zjišťování údajů napříč odvětvími mimo geoprostorovou oblast.

Distribuce představuje způsob, jakým jsou data zpřístupněna. DCAT je poměrně malý slovník, ale záměrně ponechává mnoho detailů otevřených. Vítá "aplikační profily": konkrétnější

specifikace postavené nad DCAT, resp. nad GeoDCAT-AP jako geoprostorové rozšíření.

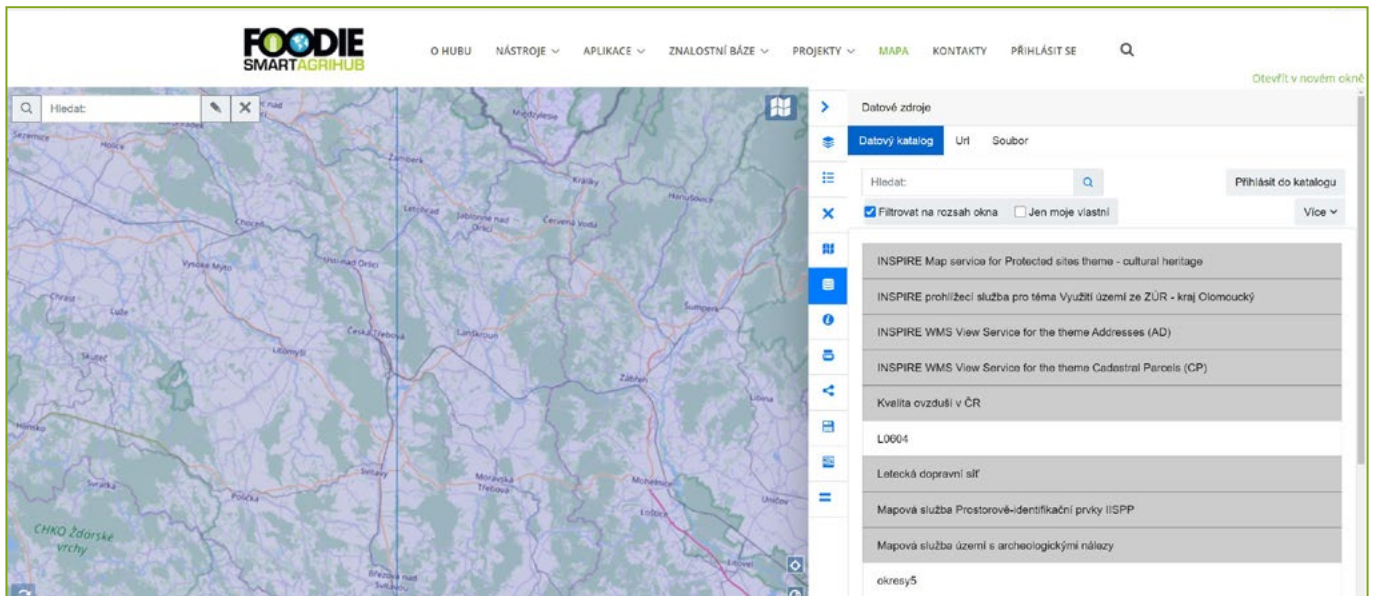
Pro senzory existuje také SensorML, standard, který lze použít k popisu široké škály senzorů, včetně dynamických i stacionárních platforem a in-situ i vzdálených senzorů. Další možností je Sémantická ontologie senzorové sítě, která popisuje senzory a pozorování a související koncepty. Nepopisuje doménové koncepty, čas, místa atd.; ty mají být zahrnuty z jiných ontologií prostřednictvím importu OWL. Tuto ontologii vyvinula skupina W3C Semantic Sensor Networks Incubator Group (SSN-XG).

Existuje potřeba harmonizace metadat prostorových a ne-prostorových datových sad a služeb. GeoDCAT-AP je jasnou volbou vzhledem k silnému zaměření na geografické datové sady. Jeho hlavní výhodou je, že umožňuje uživatelům jednoduše se dotazovat na všechny sady geoprostorových dat. [60]

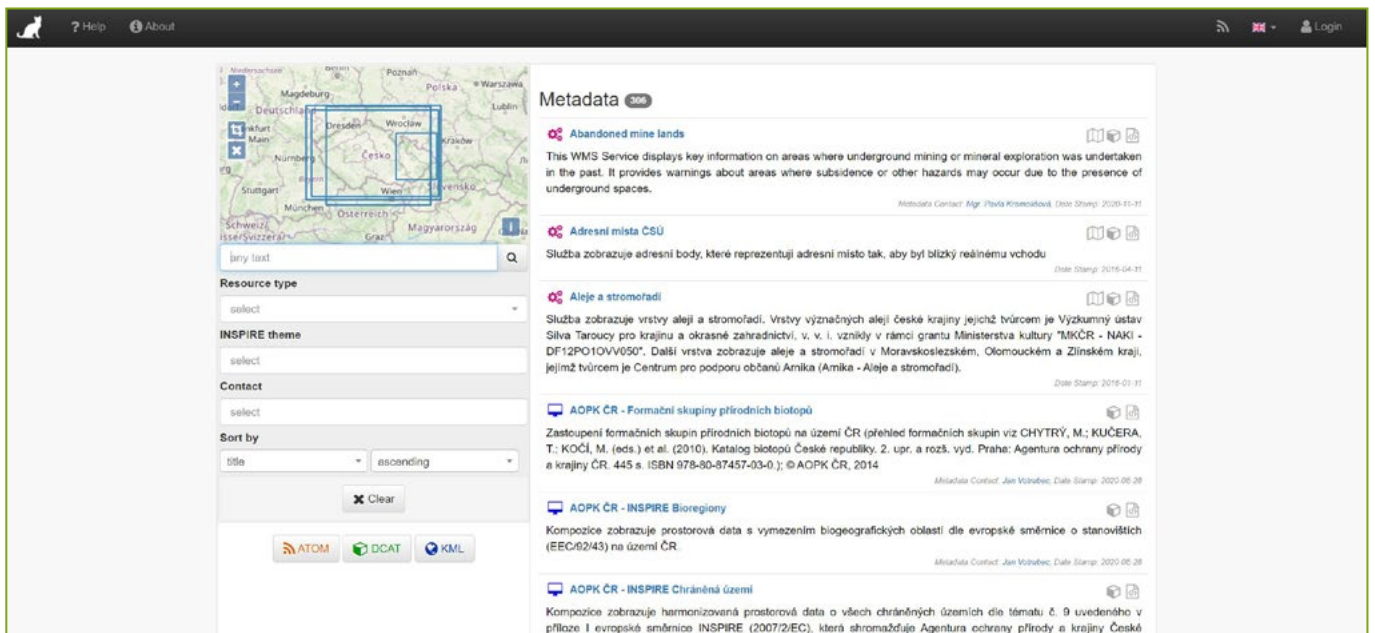
8.2. (NE)VIDITELNÁ METADATA

Vraťme se na chvíli k příkladu s nákupem uzeniny v supermarketu. Pokud chodíme do stejného supermarketu nakupovat každý týden, obvykle přesně víte, kde daný produkt hledat (mě osobně silně rozčílí, pokud změní systém uložení zboží). Někdy i využijeme reklamní leták. Pokud ale jste v některém supermarketu poprvé, tak vám určitě pomohou informační tabule - katalogy, které vám řeknou, kde dané zboží hledat. Obdobě je to i s daty. U dat, která používáme každý den, obvykle nepotřebujeme nástroj na jejich vyhledání. K tomu, abychom mohli použít data nová, potřebujeme tato data vyhledat. A samotné vyhledávání by mělo být co nejjednodušší, bez ohledu na to, jak sofistikovaný metadatový systém je za tímto vyhledáváním nasazen. Můžeme i zde využít reklamní letáky nebo třeba vyhledávat data přímo v katalogu integrovaného do mapového klienta (Obrázek 11 a 12). [61]

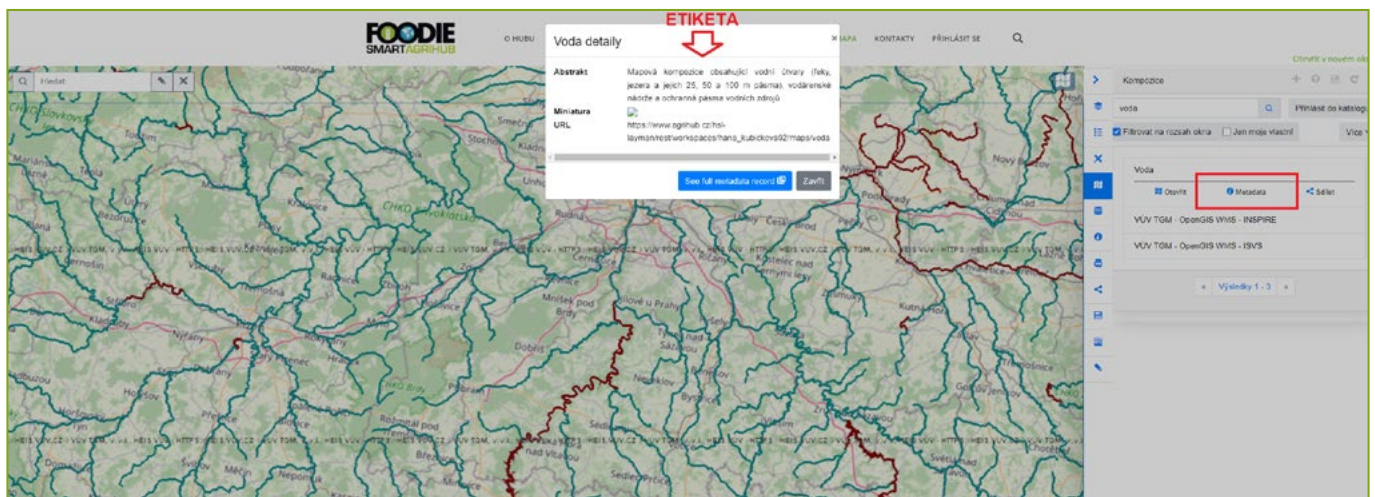
Obrázek 11: Ukázka nabídky připravených mapových kompozic na portále AgriHub.cz.



Obrázek 12: Ukázka přidávání vrstev do mapové aplikace přímo z datového katalogu.



Obrázek 13: Ukázka několika datových sad v prostředí metadatového katalogu Micka.



Obrázek 14: Ukázka základního metadatového záznamu vrstvy integrované do mapového okna.

← / Basic metadata / Full Metadata

Voda

Abstrakt

Mapová kompozice obsahující vodní útvary (řeky, jezera a jejich 25, 50 a 100 m pásma), vodárenské nádrže a ochranná pásma vodních zdrojů. Zdrojem následujících prostorových dat je portál www.eagri.cz, resortní portál Ministerstva zemědělství ČR. Datové služby jsou v souladu se standardizovanými protokoly Open GIS konsorcia (OGC) WMS verze 1.1.1. a 1.3.0. Odkaz na OGC specifikaci: <http://www.opengeospatial.org/standards/wms>

Polygonové vrstvy: Vodní útvary - buffer 25 m - Data obsahují zóny 25 metrů v okolí vodních toků na území ČR. Data jsou viditelná do měřítka 1:30 000. Vodní útvary - buffer 50 m - Data obsahují zóny 50 metrů v okolí vodních toků na území ČR. Data jsou viditelná do měřítka 1:30 000. Vodní útvary - buffer 100 m - Data obsahují zóny 100 metrů v okolí vodních toků na území ČR. Data jsou viditelná do měřítka 1:30 000. OPVZ - vodárenské nádrže - Data obsahují polygony vodárenských nádrží. Data jsou viditelná do měřítka 1:30 000. OPVZ - podzemní nebo povrchové - Data obsahují polygony vodárenských nádrží – podzemní nebo povrchové. Data jsou viditelná do měřítka 1:30 000. Zdrojem následujících prostorových dat je Výzkumný ústav vodohospodářský T. G. Masaryka, veřejná výzkumná instituce, nebo zkráceně VÚV TGM, v.v.i. Liniové vrstvy: Vodní útvary povrchových vod kategorie "řeka" včetně silně ovlivněných vodních útvarů a umělých vodních útvarů: 2. plánovací cyklus Polygonové vrstvy: Vodní útvary povrchových vod kategorie "jezero" včetně silně ovlivněných vodních útvarů a umělých vodních útvarů: 2. plánovací cyklus Vodní útvary povrchových vod byly vymezeny na základě vybraných přírodních charakteristik vodních toků a nádrží se zohledněním administrativních hranic (oblastí povodí ČR, státní hranice) nad Základní vodohospodářskou mapou 1 : 50 000. Ve verzi 3.0, platné od dubna 2006, byly vodní útvary povrchových vod identifikovány nad vstřívami vodních toků a vodních nádrží Digitální báze vodohospodářských dat (DIBAVOD) v měřítku 1 : 10 000. Identifikace silně ovlivněných a umělých útvarů byla převzata z Plánů národních částí mezinárodních povodí. Verze 4.1 je aktualizací verze 4.0 (únor 2010). Doplněn byl údaj o příslušnosti útvaru k dílčímu povodí a upravena byla struktura dat. Pro 2. cyklus plánování (2010-2021) bylo vymezení útvarů povrchových vod revidováno a změněna byla typologie útvarů (Vyhláška č. 49/2011 Sb. o vymezení útvarů povrchových vod). Pro přípravu plánů povodí byla publikována verze 5.0 vymezení (4. listopadu 2014). V průběhu přípravy plánů povodí byly v této verzi opraveny některé nekonzistence v identifikátorech útvarů a jejich názvech a doplněny údaje o typu útvaru. Tyto úpravy byly zahrnuty do aktuálně platné verze 5.1, (25. září 2015). Popisné údaje evidence blíže specifikuje Katalog evidencí ISVS VODA vedených VÚV TGM, v.v.i. (Praha: VÚV TGM, v.v.i., 2015). Vytvoření 28.2.2010 Revize: 31.3.2011 4.11.2014 25.9.2015

Browse Graphic

HTTP://HBSW.VV.CZ/VU/TGM/1/2/HELE/VV/

Obrázek 15: Ukázka detailního metadatového popisu vrstvy.

Obvykle pouze experti nebo nadšenci pracují přímo s metadatovým katalogem (Obrázek 13).

Již jsme našli náš produkt, uzeninu nebo mléčný výrobek a jelikož nás zajímá kvalita toho, co jíme, prohlédneme si etiketu, eventuálně někde budeme hledat ještě podrobnější popis na letáku. A tak je tomu s daty. Pro prvotní orientaci nás zajímá etiketa (Obrázek 14).

Můžeme si však rozbalit příbalový leták (Obrázek 15).

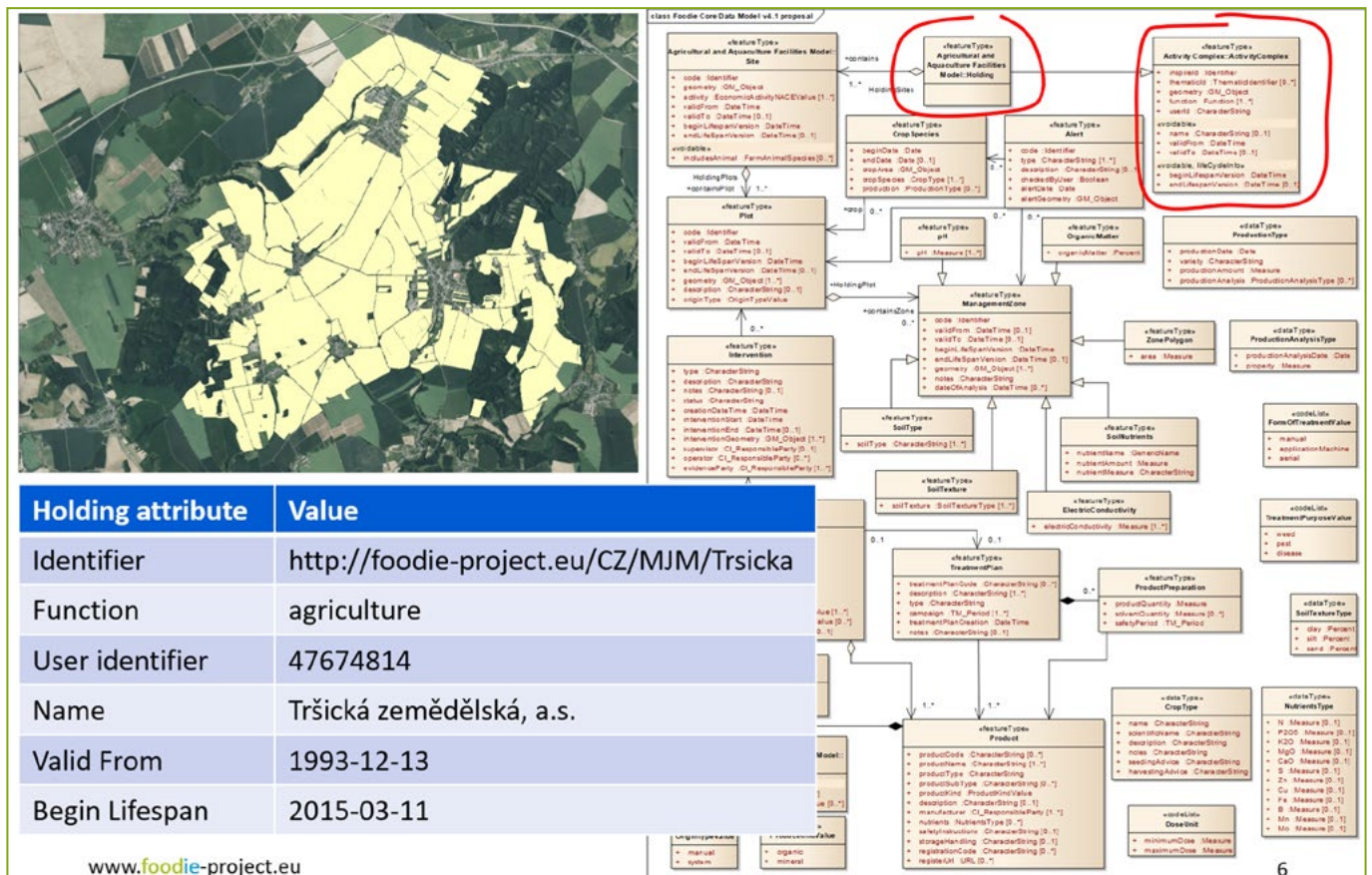
Takže se dá říci, že metadata jsou něco, co denně používáme a dobře známe. Jenom se s nimi musíme naučit pracovat. A to jak uživatelé, tak samozřejmě i poskytovatelé dat, aby bylo možné data snadno vyhledávat.

9. DATOVÉ ZDROJE A JEJICH HARMONIZACE

V předchozích kapitolách jsme si představili základní principy nutné pro správu velkých dat. V kapitole 2.5. Destination Earth: digitální dvojče Země jsme si řekli, co jsou to digitální dvojčata a zmínili jsme se o tom, že to je potenciální cesta, jak získat znalosti a moudrost z dat, aby naše zemědělství a venkovské regiony byly udržitelné. V kapitole 2.3. Znalostní pyramida jsme vysvětlili, že základem každého rozhodování musí být data. Také jsme se zmínili, že tato data musí být interoperabilní, aby s nimi mohli pracovat různí uživatelé a různé aplikace. Nyní si ukážeme na třech různých případech, jak mohou taková data vypadat, jak se mohou konstruovat a co je důležité, jak se mohou takovéto datové zdroje provázet.

9.1. FOODIE DATA MODEL

Datový model FOODIE vychází z obecných datových modelů evropské direktivy INSPIRE, zejména z datových modelů pro zemědělská a akvakulturní zařízení a informační systém o pozemcích. Klíčovou motivací bylo reprezentovat souvislou plochu zemědělské půdy s jedním druhem plodin, kterou spravuje jeden uživatel v jednom režimu hospodaření (konvenční vs. přechodné vs. ekologické zemědělství). Datový model FOODIE navíc obsahuje koncepty pro údaje o plodinách a půdě, ošetření, zásahy, zemědělské stroje a další. V neposlední řadě model znovu využívá datové typy definované v normách ISO (ISO 19101, ISO/TS 19103, ISO 8601 a ISO 19115) a také standardizační úsilí zveřejněné



Obrázek 16: FOODIE Datový model - Farma (převzato z [65]).

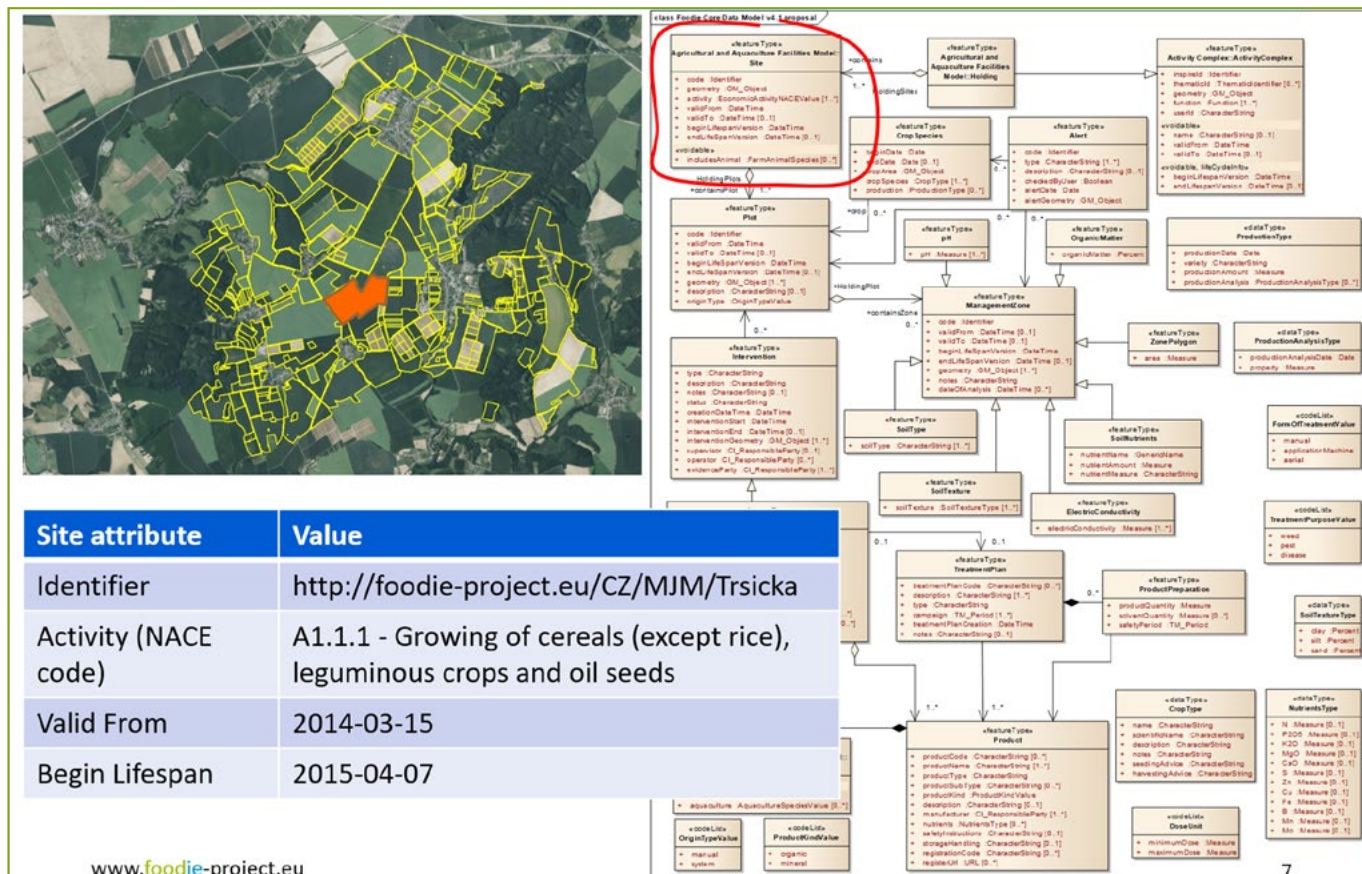
v rámci směrnice INSPIRE (například struktura jedinečných identifikátorů). Datový model FOODIE byl specifikován v jazyce UML (stejně jako modely INSPIRE), ale lze jej transformovat do ontologie OWL, aby bylo možné publikovat propojená data odpovídající datovému modelu FOODIE [62]. FOODIE datový model je komplexní, pro jeho lepší pochopení se pokusíme popis modelu doplnit ilustrativními obrázky, kde na jedné straně bude užít UML diagram (pro laika asi ne příliš srozumitelný) a na druhé straně mapa s grafickým vyjádřením pojmu. FOODIE data model je otevřený a přístupný každému. Pro technicky zaměřené odborníky lze najít UML schéma datového modelu v [63] a FOODIE

Ontologii v [64]. Model je stále živý a užívá se v řadě projektů a i některých zemích.

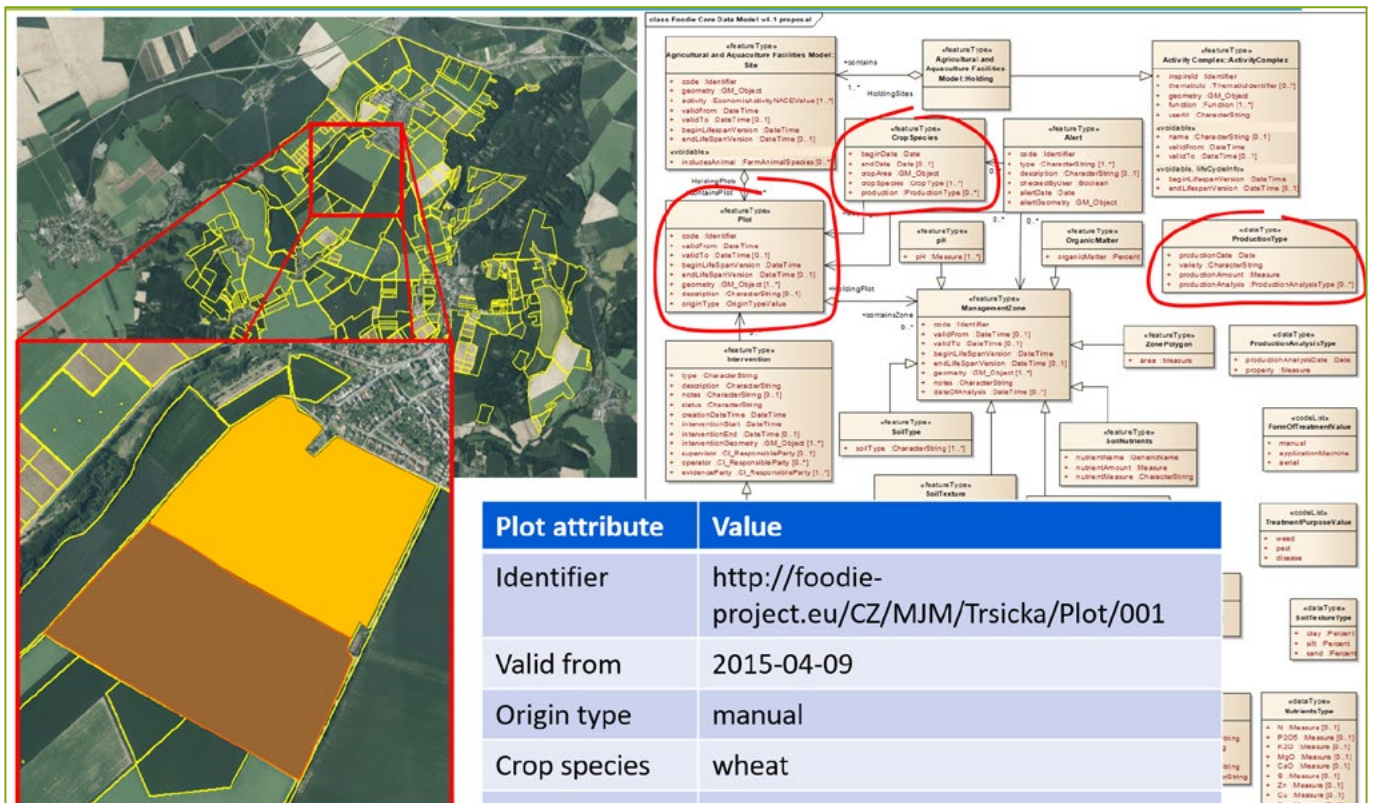
Základním prvkem FOODIE datového modelu je farma (Obrázek 16). Prvek farma zahrnuje veškerou zemědělskou půdu náležící k jedné farmě.

Nižší úroveň daného modelu je Díl půdního bloku (Obrázek 17) tak, jak ho známe z LPISu.

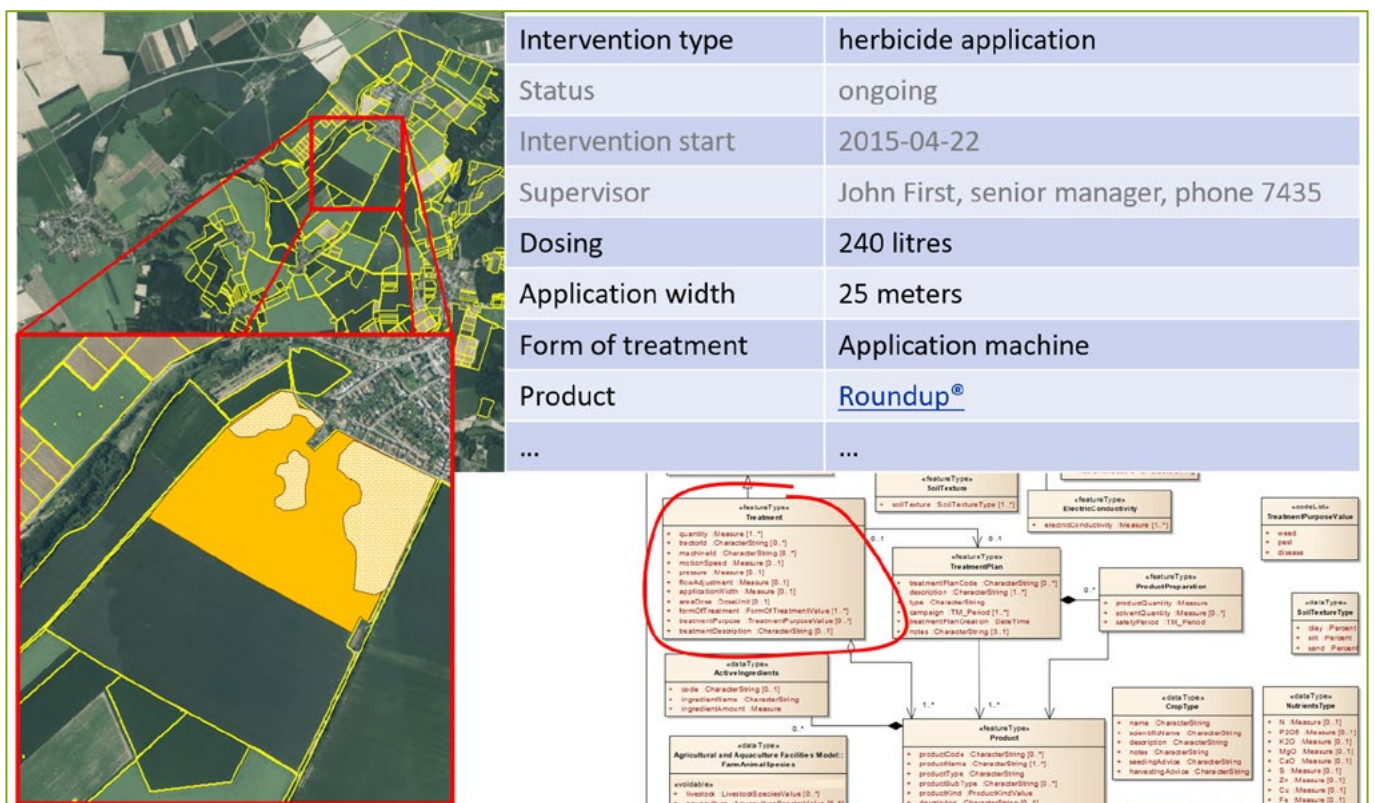
Pro účely FOODIE by měl existovat prvek na podrobnější úrovni než "Díl půdního bloku", který je již součástí datového modelu INSPIRE. Hlavní motivací je reprezentovat



Obrázek 17: FOODIE Datový model - Díl půdního bloku (převzato z [65]).



Obrázek 18: FOODIE Datový model - Zemědělská parcela [převzato z [65]].



Obrázek 19: FOODIE Datový model - Management zona [převzato z [65]].

souvislou plochu zemědělské půdy s jedním druhem plodin, kterou obhospodařuje jeden uživatel v jednom režimu hospodaření (konvenční vs. přechodné vs. ekologické zemědělství). Takový prvek se ve specifikaci základního datového modelu FOODIE nazývá "Zemědělská parcela"

(Obrázek 18) a je jeho základní referenční položkou v datovém modelu.

Datový model FOODIE obsahuje také typ prvku "Management Zona" (Obrázek 19), který je o jednu úroveň nižší

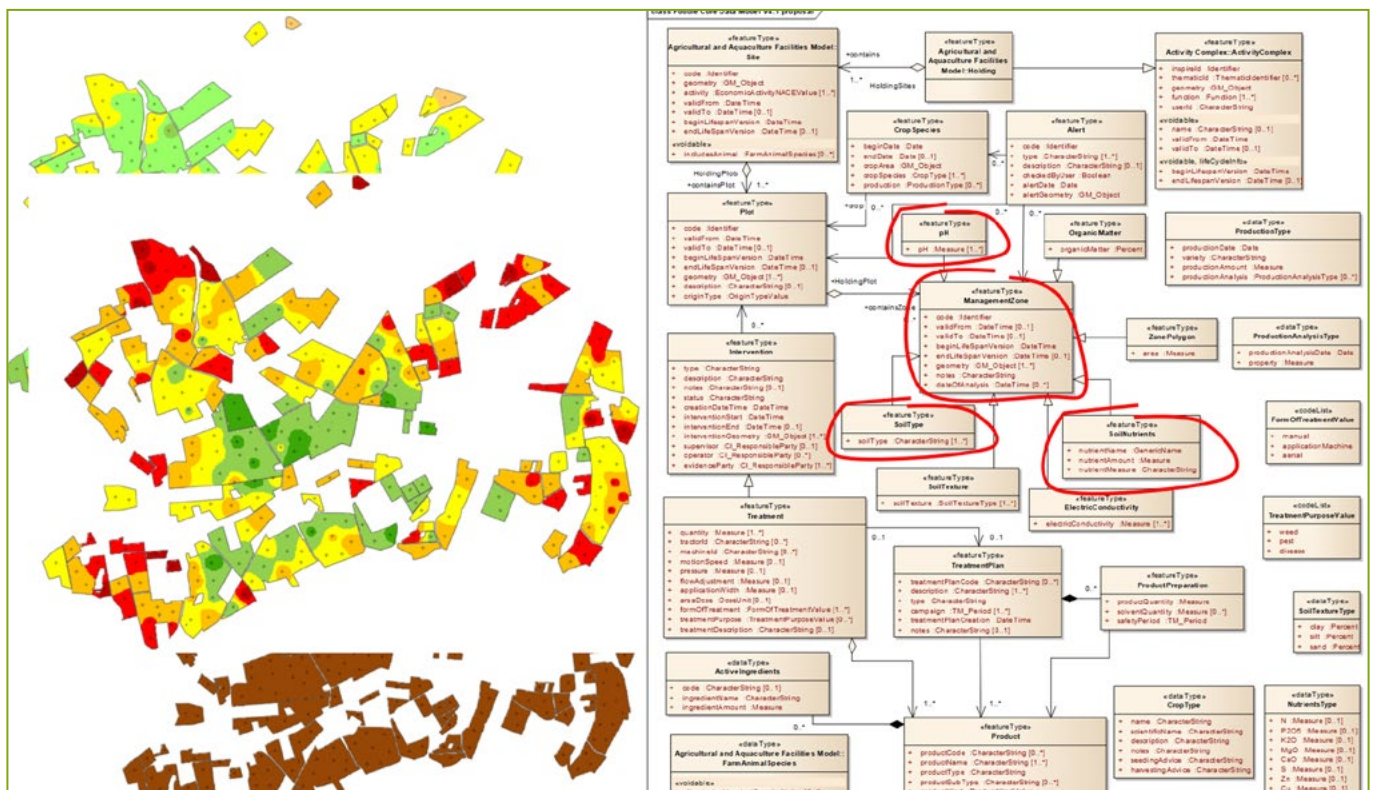
9.2. OTEVŘENÁ DATA VYUŽITÍ PŮDY - OPEN LAND USE

než "Zemědělská parcela". Jinými slovy, "Management Zona" je prostorová podmnožina prvku "Zemědělská parcela", která má specifické vlastnosti, jako je elektrická vodivost, organická hmota, pH, struktura půdy, půdní typ nebo živiny v půdě.

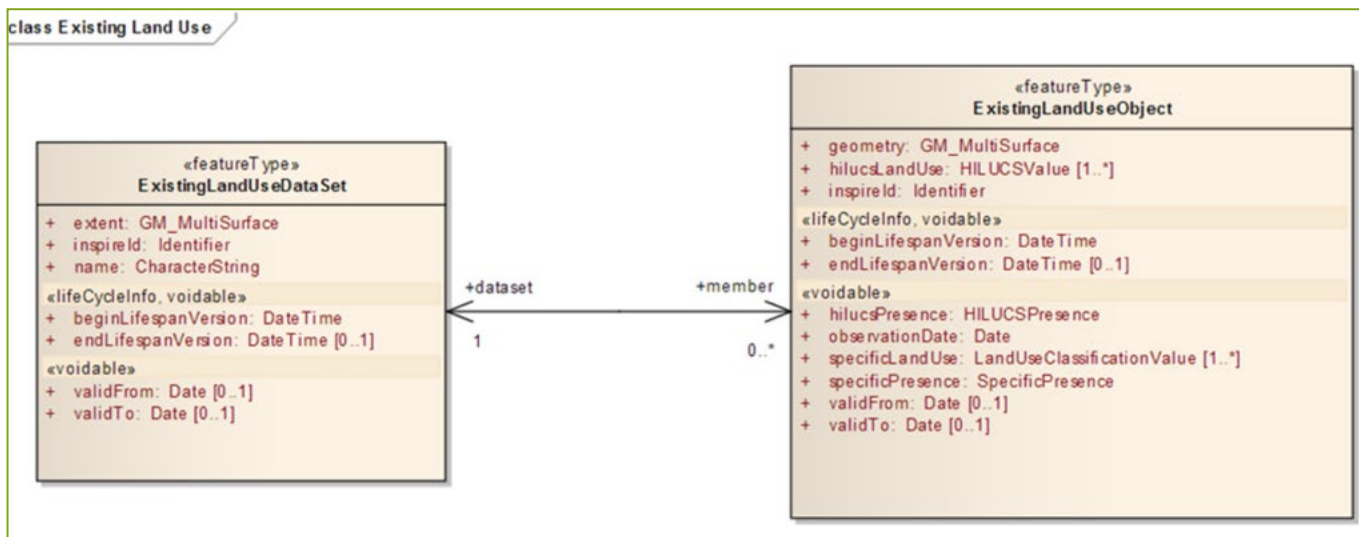
Model může zahrnovat i mapy vzorku a dalších měření, viz. Obrázek 20.

Specifikace základního datového modelu FOODIE byla navržena jako otevřený datový model, což umožňuje jeho rozšíření prostřednictvím asociací a/nebo atributů, které dále specializují funkci "blok". Cílem tohoto rozšíření je zajistit modularitu a umožnit každému zemědělci/externímu poskytovateli služeb využívající platformu FOODIE rozšířit datový model podle svých potřeb. [66]

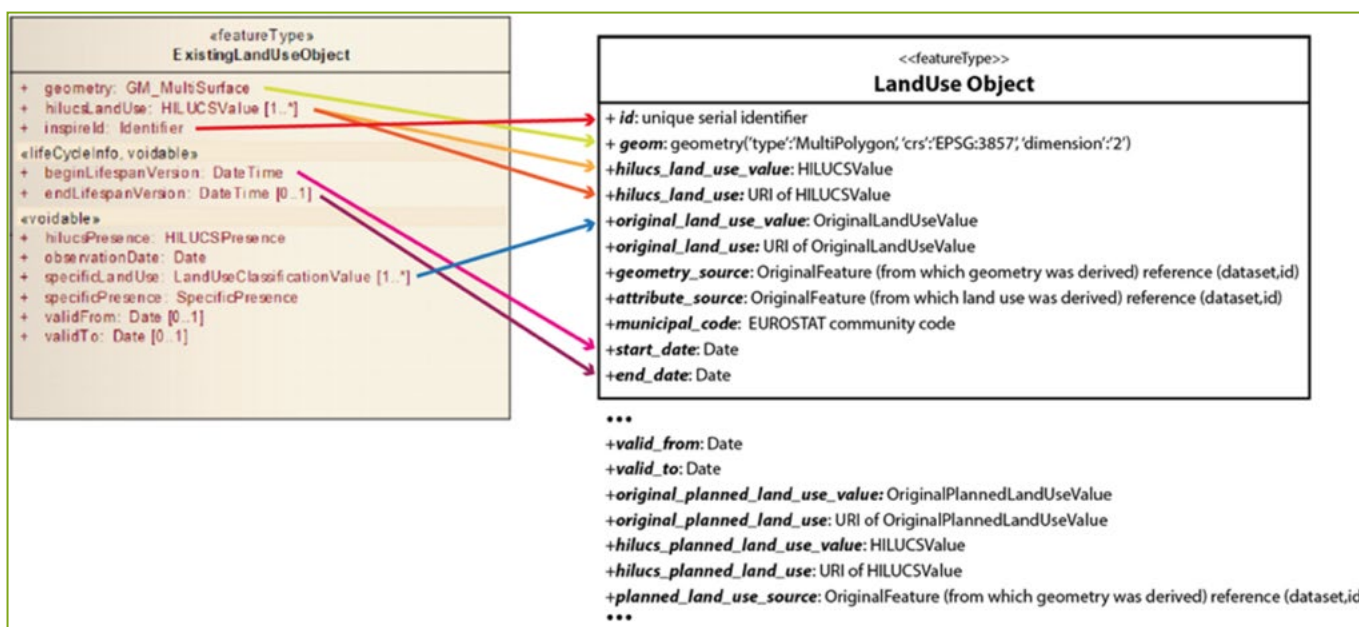
FOODIE data model nám ze schématu na Obrázku 4 pokrývá mikro úroveň dat a informací a částečně i úroveň farmy. Následující model a datová sada se zaměřuje na makro úroveň dat a informací. Digitální produkt nazvaný Open Land Use (OLU) map neboli Otevřená mapa využití půdy si klade za cíl vytvořit co nejvíce podrobnou mapu využití krajiny na základě dostupných dat. Samotná mapa může být využita pro různé úrovně rozhodování a pro různou škálu účelů. Například, mapu lze použít pro účely územního plánování, přizpůsobení se změnám klimatu, ochrany biologické rozmanitosti, řízení vody v krajině, zalesňování a další funkce, které jsou spojeny s územním plánováním. OLU mapa má širší uplatnění, ale pro řízení zemědělství a lesnictví v regionálním měřítku má nezastupitelnou roli.



Obrázek 20: FOODIE Datový model - Mapa vzorků [převzato z [65]].



Obrázek 21: Datový model prvku využití půdy.



Obrázek 22: Přejechod od INSPIRE modelu prvku využití půdy k modelu OLU.

Aby taková mapa mohla vzniknout, je třeba udělat jisté přípravné kroky: vytvořit datový model, udělat rešerši dostupných datových zdrojů, vybrat nástroje na harmonizaci a agregaci dat, prezentovat a publikovat výslednou datovou sadu.

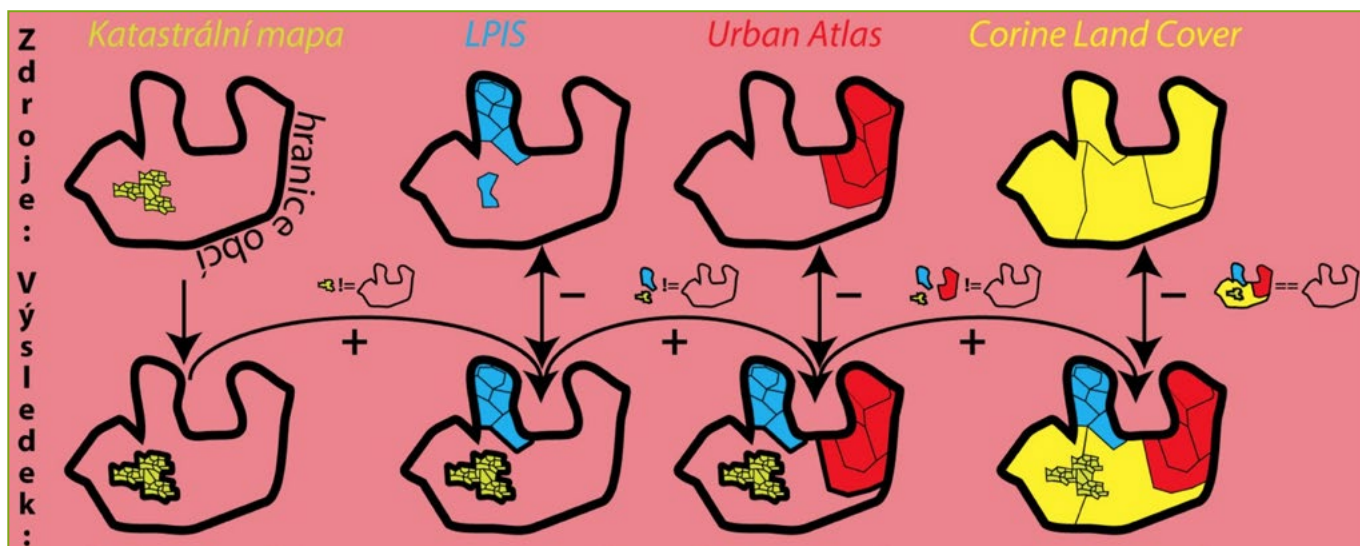
Při tvorbě datového modelu se bralo v potaz, že využití půdy je jedno z témat INSPIRE (směrnice k vytvoření celoevropské infrastruktury prostorových informací), takže datový model prvku využití půdy byl již dostupný, viz Obrázek 21. Ten byl převzat jako základ modelu existujícího prvku v OLU a rozšířen o několik zájmových atributů. Tento model pro prvek v OLU a mapování atributů mezi ním a originálním datovým modelem prvku využití půdy dle INSPIRE je k vidění na Obrázku 22.

Po stanovení datového modelu proběhl import zdrojových datových sad do otevřené databáze PostgreSQL. Násled-

ně po importu dat byl spuštěn skript na generování OLU. Obrázek 23 zobrazuje algoritmus tvorby OLU České republiky. Kvůli tomu, že v době tvorby OLU pro ČR nebyl na 100% dostupný digitální katastr a mapa OLU má být souvislou (tj. bezešvou mapou), nebylo možné jednoduše převzít data o parcelách katastru nemovitostí a udělat převod mezi třídami využití půdy, jak jsou vedeny v katastru a třídami využití půdy z klasifikaci HILUCS z INSPIRE [67].⁹

Tím pádem bylo třeba také přidat další datové sady, z nichž bylo možné určit třídu využití půdy. Použitím dat datové sady Land Parcel Identification System (LPIS) neboli půdní bloky [68], Urban Atlas [69] a Corine Land Cover [70] bylo zajištěno souvislé pokrytí celého území České republiky mapou OLU.

⁹ HILUCS = Hierarchical INSPIRE Land Use Classification System



Obrázek 23: Algoritmus tvorby OLU České republiky.



Obrázek 24: Odvození způsobu využití parcely z datové sady nižšího rozlišení.

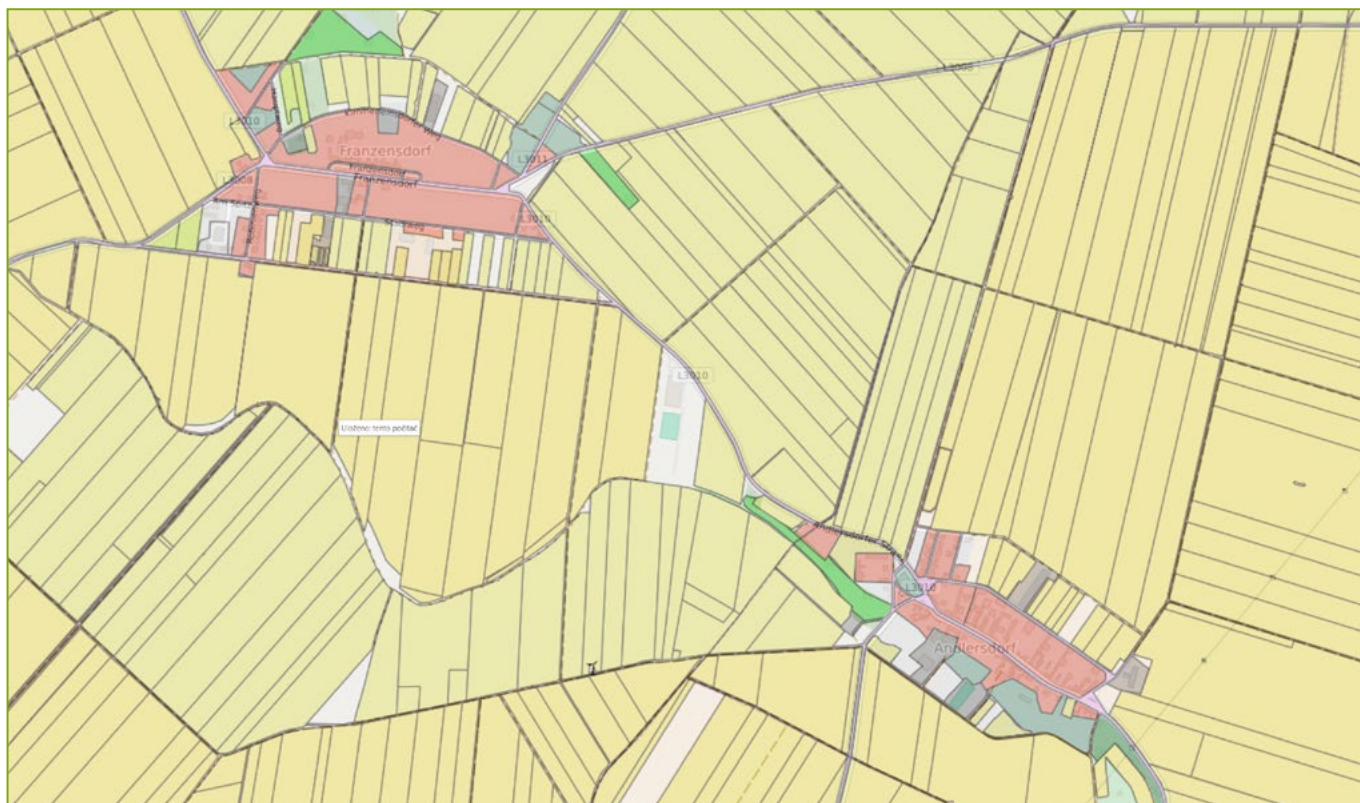
Zjednodušený postup výpočtu geometrie prvku OLU byl následující:

1. Hranice obce, pro niž se generuje OLU, se vezme jako maska.
2. Průnik všech parcel z dostupné DKM s maskou obce se přidá do OLU mapy obce.
3. Zkontroluje se, zda plocha pokrytá parcelami se rovná celkové ploše obce. Pokud je to tak, znamená to, že už celé území obce je pokryto mapou OLU, a tak se celý algoritmus ukončí. Pokud ne, tak to znamená, že ještě není celé území pokryto mapou OLU, a algoritmus bude pokračovat.
4. Následně se maska zmenší o území pokryté parcelami.
5. Pak se spočte průnik farmářských bloků z LPISu s maskou, upravenou v minulém kroku.
6. Průnik se zapíše do výsledné OLU obce.
7. Opět se provede kontrola, zda již není pokryto celé území obce. Pokud není pokryto, tak se od masky odečte území pokryté plochami LPISu.
8. Stejný postup se provedl s datovou sadou Urban Atlas a Corine Land Cover, viz Obrázek 23.

Algoritmus poběží až do té doby, než se celé území obcí pokryje mapou OLU. To, že algoritmus nakonec

doběhne, je jisté skutečnosti, že datová sada Corine Land Cover pokrývá celou Evropu. Je však ze všech zmíněných datových sad nejméně přesná.

Před a nebo po tomto geometrickém sloučení je třeba provést převod atributů, podle kterého se určuje typ využití půdy dané oblasti dle klasifikace HILUCS. K němu se použijí převodní tabulky, při jejichž návrhu je třeba být co nejvíce objektivní. Případy, kdy třída využití půdy ve zdrojových datech je shodná s třídou využití půdy podle HILUCS, nejsou tak časté. V případě, že ze zdrojových dat, které tvoří jednotlivé plochy OLU, není možné odvodit třídu využití půdy, je třeba po vytvoření mapy určit třídu s využitím jiných pomocných datových sad. Toho se dá docílit například pomocí ověření, zda se plocha v OLU kříží s plochou z jiné datové sady, ve které je uvedena informace, z níž se dá určit využití pozemku, i když jde o datovou sadu s výrazně menším detailem. Například, pokud by parcely katastru z minulého obrázku neměly informaci o způsobu využití, dala by se odvodit například z datové sady Corine Land Cover následujícím způsobem na základě převládajícího typu využití půdy v překrývajících se plochách, viz Obrázek 24.



Obrázek 25: Ukázka vizualizace OLU mapy.

Bezesporu je toto velice hrubé určení typu využití půdy, protože prostorové rozlišení datové sady Corine Land Cover je mnohem menší než prostorové rozlišení katastrálních dat. Tímto krokem se ale zajistí bezešvé pokrytí zájmového území, a v budoucnu se tyto prvky v OLU mohou obohatit o informace z podrobnějších zdrojů informací o využití půdy v dané lokalitě (Obrázek 25).

Aby bylo možné provést analýzu a rozšířit mapu OLU o další datové soubory, bylo nutné změnit datový model první verze mapy OLU představený v předchozím textu. Tato kapitola podrobně popisuje druhou verzi datového modelu OLU, po níž následují příklady prvních pokusů o integraci různých atributů souvisejících s využitím půdy/půdním pokryvem a půdou do druhé verze mapy OLU.

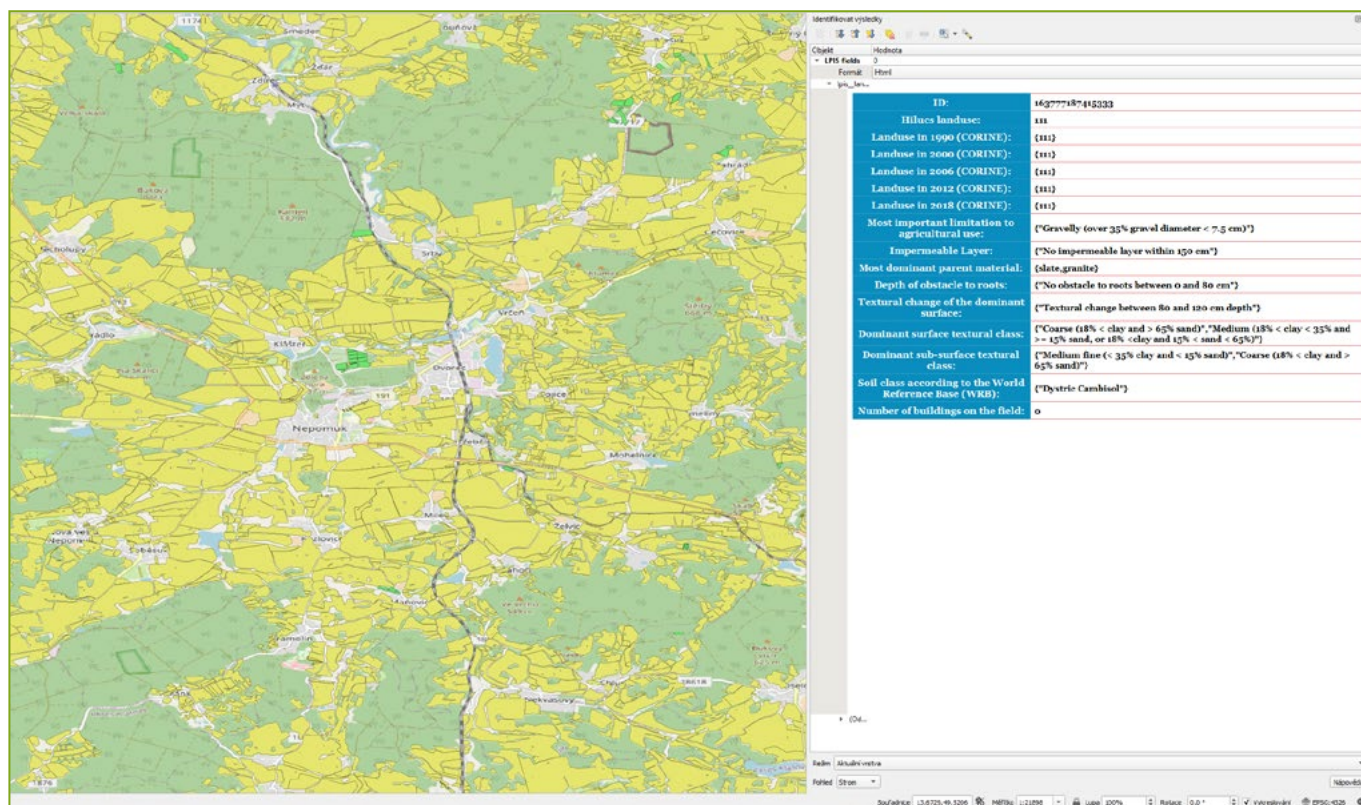
Nová verze datového modelu OLU byla navržena na základě požadavků definovaných potřebami různých evropských projektů. Hlavním důvodem návrhu nové verze datového modelu bylo snazší začlenění tematických datových sad do databáze OLU, která je zdrojem pro sestavení samotné mapy OLU. Po vytvoření hlavní vrstvy datové sady OLU je stále možné přidávat tematické atributy do datové sady OLU. Což bylo v předchozí verzi datového modelu OLU velmi komplikované. Hlavní vrstva OLU se vytváří z datových sad, které poskytují referenční geometrii – nejpodrobnější a nejpřesnější geometrii, jaká je v daných oblastech k dispozici – např. údaje katastru nemovitostí, data o půdních blocích atd. Referenční vrstva OLU obsahuje geometrie z datových sad menšího měřítka, kde není k dispozici nic podrobnějšího.

Hlavní myšlenkou nové verze datového modelu OLU je oddělení geometrie a atributů prvků. Tento přístup zajišťuje, že geometrické prvky mohou získat další tematické atributy v každém bodě zpracování datové sady OLU. Negeometrické atributy jsou uloženy jako soubor atributů a mohou být vztaženy k jakémukoli geometrickému prvku na základě určitého prostorového vztahu.

Ačkoli dostupné a vhodné datové sady pro takto popsaný nástroj mohou být v několika formátech a geometriích, je nutné některé z nich před použitím pro druhou verzi OLU předem zpracovat. Například datová sada Land use and land cover survey (LUCAS) [71] je bodovou sadou, ale mapa OLU je založena na polygonech, takže před začleněním do OLU je třeba provést interpolaci dat LUCAS. Ve skutečnosti jsou vlastnosti reprezentované v bodové datové sadě LUCAS postupně dostupné ve formátu rastrové mapy, což umožňuje její jednodušší budoucí začlenění.

Na území České republiky probíhají experimenty s integrací různých atributů souvisejících s využitím půdy/půdním pokryvem a půdou. V první verzi OLU byly k dispozici čtyři hlavní zdroje dat, z nichž bylo odvozeno OLU, a to: parcely z Registru územní identifikace, adres a nemovitostí (RÚIAN) [72], půdní bloky z Land Parcels Identification System (LPIS), Urban Atlas (UA) a Corine Land Cover (CLC).

Ve druhé verzi mapy OLU, která může sloužit např. jako nástroj pro korelaci údajů o využití půdy a kvalitě půdy, byly tyto čtyři zdroje dat překryty hranicemi obcí v pořadí podle prostorového rozlišení zdrojů dat (RÚIAN, LPIS,



Obrázek 26: Příklad výběru konkrétního pole a jeho atributů v Plzeňském kraji zobrazený v softwaru QGIS.

UA, CLC), aby bylo dosaženo bezešvého pokrytí území celé obce. Kromě toho byla pro integraci do datové sady OLU shromážděna řada tematických datových sad týkajících se půdy a jejího využití. Je třeba zmínit, že samotný LPIS je datová sada sloužící k uchování údajů o polích zemědělců, takže již obsahuje některé údaje našeho zájmu (týkající se půd, využití půdy/půdního pokryvu). Jmenujme je:

- Půdně ekologická jednotka (BPEJ) je pětimístný kód, který popisuje čtyři různé aspekty půdy (klimatický region, půdní typ, hloubku půdy a strukturu půdy),
- úroveň eroze půdy (tři úrovně: pozemek bez eroze půdy, pozemek mírně ohrožený erozí půdy, pozemek silně ohrožený erozí půdy),
- druh rostlin pěstovaných na poli.

Kromě těchto informací byly shromážděny další relevantní zdroje dat, které byly začleněny do druhé verze mapy OLU, např.:

- Údaje o větrné erozi půdy z Evropského centra pro data o půdě (ESDAC¹⁰) [73].
- Údaje o vodní erozi půdy z ESDAC [74].
- Údaje o půdě z Evropské databáze půdy (ESDB¹¹) [75].
- Obsah organického uhlíku v půdě z Mezinárodního referenčního a informačního střediska pro půdu (ISRIC¹²) [76].

Tematické zdroje dat budou do OLU integrovány jednoduchým způsobem. Jako geometrická reference budou vzata pole LPIS a tematický obsah (výše uvedené zdroje dat) bude přidán do polí LPIS na základě geoprostorového

průniku. Nakonec půjde o jednoduchou webovou mapu, kde uživatel klikne na zemědělské pole a uvidí seznam tematických atributů na tomto poli. To umožní získat informace na základě korelace údajů o využití půdy a údajů souvisejících s půdou pro konkrétní pole a dále s těmito informacemi pracovat jejich analýzou, spojováním s dalšími údaji a rozhodováním na jejich základě.

Samotný příklad druhé verze mapy využití volné půdy obsahuje vybrané informace o polích zemědělců, tj. údaje o využití půdy (současné i historické) a údaje o kvalitě půdy pro území Plzeňského kraje, který je regionem NUTS3¹³ v České republice. Příklad je znázorněn na Obrázku 26 (vizualizace v desktopovém softwaru QGIS) a na Obrázku 27 (vizualizace ve webové mapě SIEUSOIL Hub dostupný na webové stránce <https://tinyurl.com/y637skea>). Uvedený příklad zobrazuje území České republiky, nicméně v současné době jsou tendence standardizace a interoperability mezi informacemi o půdách mezi jednotlivými zeměmi. Například mezinárodní projekt SIEUSOIL¹⁴ se snaží standardizovat půdní data nejen napříč Evropou, ale také pro území Čínské lidové republiky.

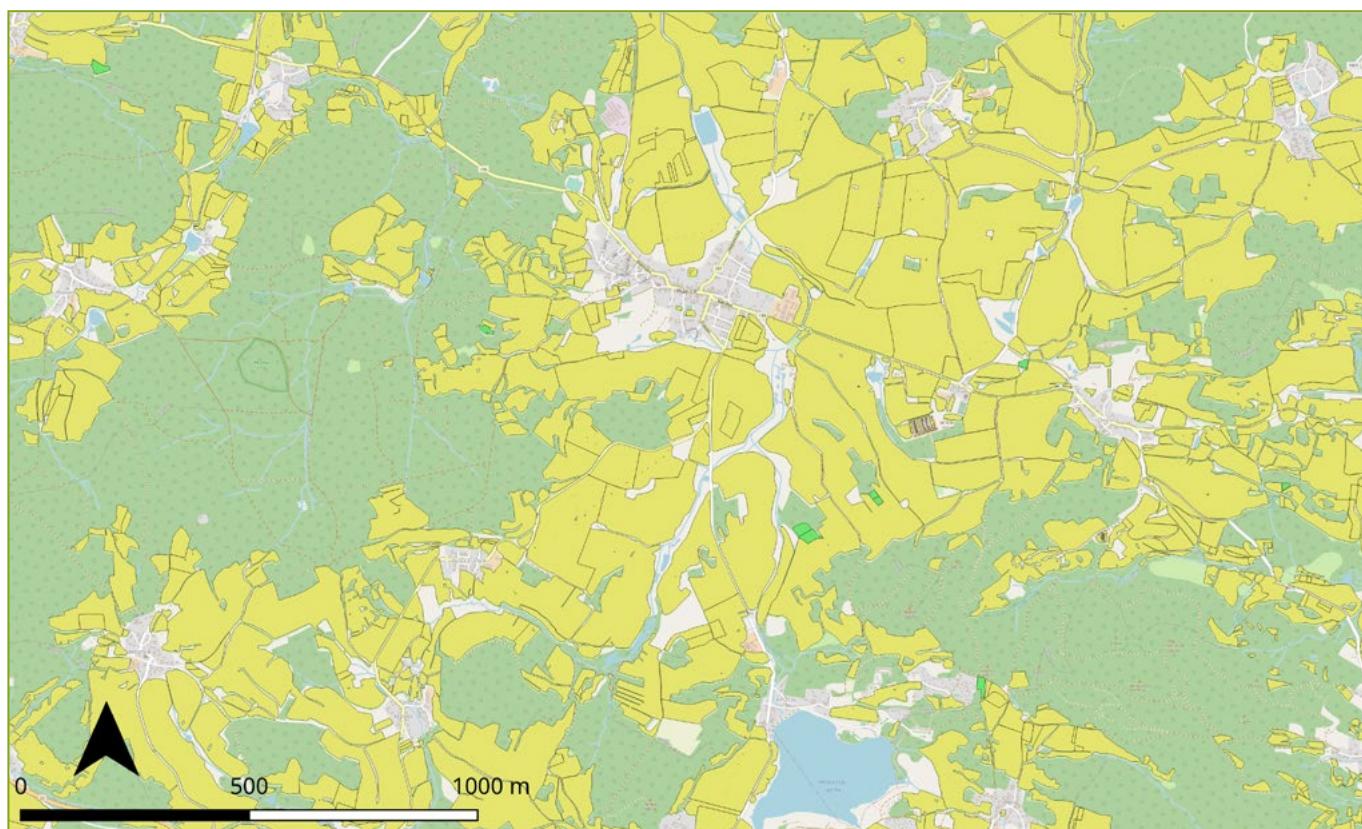
¹⁰ ESDAC = European Soil Data Centre

¹¹ ESDB = European Soil Database

¹² ISRIC = International Soil Reference and Information Centre

¹³ Dle klasifikace NUTS (Nomenclature of territorial units for statistics) poskytovány EUROSTAT, <https://ec.europa.eu/eurostat/web/nuts/background>

¹⁴ www.sieusoil.eu



Obrázek 27: Příklad polí v oblasti Plzeňského kraje vizualizovaných ve webové mapě SIEUSOIL Hubu.

9.2.1. Postup tvorby OLU v2:

Hlavním cílem práce bylo obohatit referenční geometrie o další tematické atributy. Mezi nimi byly nad rámec informací o využití půdy podle klasifikace HILUCS (původní cíl předchozí verze OLU) například topografické charakteristiky pozemků, vybrané půdní vlastnosti z Evropské databáze půd a datových sad ISRIC, typ krajinného pokryvu podle klasifikace CORINE land cover a také historický typ krajinného pokryvu podle časové řady datových sad CLC (1990, 2000, 2006, 2012, 2018). Tímto byl iniciován přechod od datové sady ke komplexní prostorové databázi.

Proces tvorby databáze OLU lze rozdělit do následujících kroků:

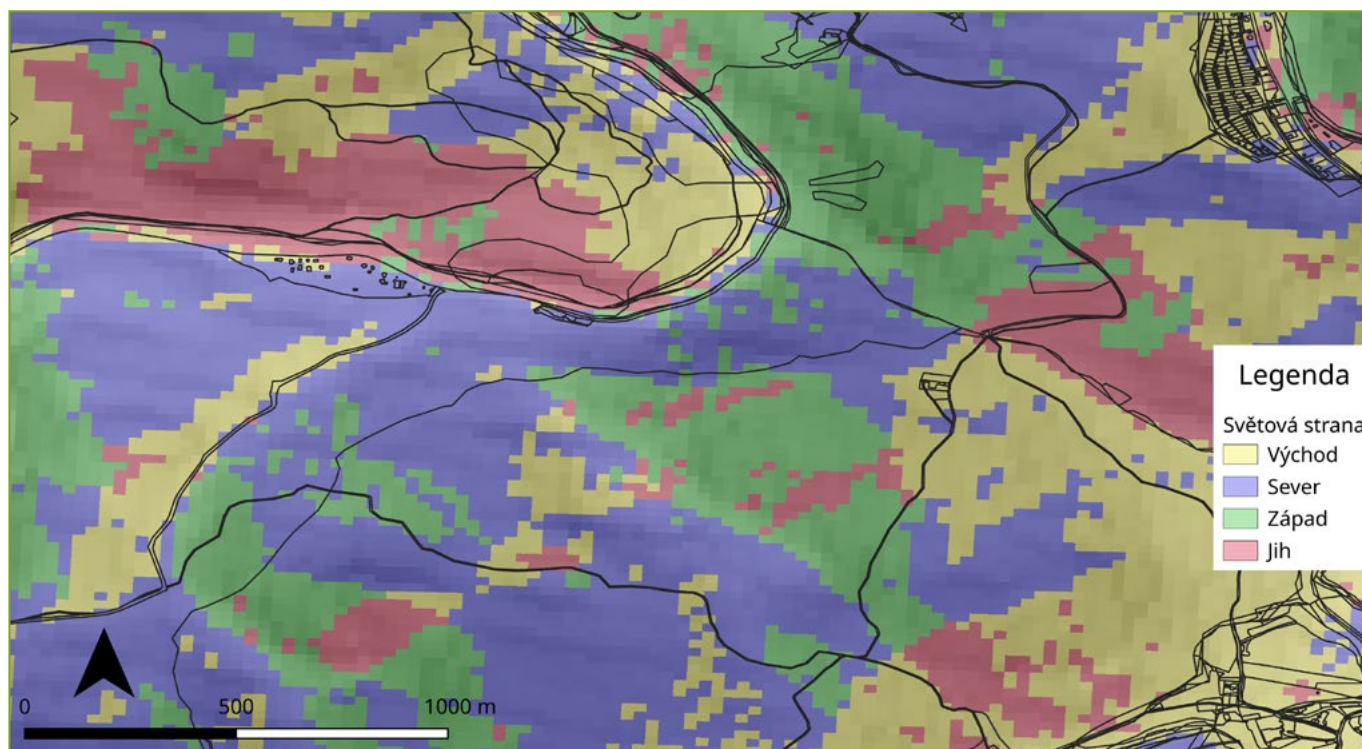
1. Sběr metadatových záznamů referenčních a tematických datových sad.
2. Na základě metadat proběhl sběr (stahování) dat. V procesu sběru dat bylo třeba parametrizovat část metadatových url odkazů datových zdrojů. Zejména mnoho datových zdrojů má ve svých url odkazech referenci na územní celek (číslo územní jednotky nebo někdy jednoduše krajní souřadnice ohraničujícího obdélníka) a časovou referenci (první den aktuálního roku, poslední den předchozího měsíce atd.). Po provedení parametrizace url odkazů byly datové sady automaticky staženy.
3. Nahrání vektorových zdrojů dat do open source databáze PostgreSQL [77]. Vzhledem k tomu, že referenční datové zdroje jsou poměrně velké, byla pro snazší správu dat použita dědičnost tabulek (podle příslušných územních jednotek, protože obvykle jsou data

podle těchto jednotek strukturována již v původní datové sadě).

4. Spolu s referenčními a tematickými údaji byly staženy údaje o administrativních jednotkách. Na základě těchto údajů byl sestaven strom administrativního členění. To bylo provedeno za účelem kombinování údajů pocházejících z jednotek různé úrovně. Například jedna datová sada může být strukturována podle obcí, jiná podle katastrálních území a k nalezení odpovídajících katastrálních území pro obec se použije právě administrativní strom. Tento strom je navíc velice užitečný, protože poskytuje snadný způsob strukturování informací a také by mohl být použit jako prostorový filtr při různých prostorových analýzách (vyhledávání průsečíků atd.).
5. Po shromáždění dat - bylo třeba transformovat jednotlivé prvky na sebe a rozdělit je na geometrické a atributové části - jak je to v datovém modelu OLU verze 2.0.

5a. Transformace atributů:

Nejprve byla provedena harmonizace tříd využití půdy a půdního pokryvu podle klasifikací HILUCS a CORINE. Kromě těchto transformací atributů bylo také nutné definovat časový rozsah platnosti prvku a jeho příslušnost k určité administrativní jednotce. Rozsah časové platnosti prvku je v některých datových sadách k dispozici (například v RUIAN), ale u některých datových sad chybí. V případě chybějícího rozsahu časové platnosti - byl vzat rozsah platnosti celé datové sady (například CORINE je aktualizován jednou za 6 let - časová platnost prvku CORINE je tedy stanovena 6 let od data zveřejnění datové sady CORINE, ze které pochází).



Obrázek 28: Mapa orientace svahů lokality v Karlovarském kraji.

Příslušnost k určité administrativní jednotce je prioritně definována atributem nebo metadaty datové sady, v opačném případě, kdy tam informace nejsou, je možné příslušnost prvku k určité administrativní jednotce definovat průnikem s datovou sadou administrativní mapy.

5b. Transformace geometrie:

Podle požadavků nového modelu byly všechny geometrie transformovány do souřadnicového systému World Geodetic System 1984 (WGS 84)¹⁵. Typ geometrie byl převeden na typ MultiPolygon.

6. Po provedených transformacích byly provedeny následující kroky. Geometrie byla uložena do tabulky s geometriemi, atributy pak do tabulky, spojení mezi nimi je vytvořeno spojovací tabulkou. Dále příslušnost prvků k určitým administrativním jednotkám/jednotkám z tabulky administrativních jednotek byla uložena do spojovací tabulky.
7. Následně atributy prvků různých referenčních vrstev byly navzájem propojeny na základě průsečíku s vrstvami, které mají nižší geometrickou přesnost. Například tedy vrstva s nejpřesnější referenční geometrií - katastrální parcely - přebírá atributy prvků LPIS, Urban atlasu a CORINE, se kterými se protíná.
8. V této fázi se přidávají tematické vrstvy. Těmi mohou být například další vektorové datové sady reprezentující tematické informace (přírodní parky, kulturní památky nebo jiné zájmové body apod.), tematické rastrové datové sady (například půdní typ, meteorologické údaje, automaticky zjišťovaný půdní pokryv) nebo může jít o výsledek různých analýz, jako je tomu u topografických charakteristik (nadmořská výška, sklonitost, azimutální orientace, index topografické vlhkosti), které jsou odvozeny z digitálního modelu reliéfu. Velmi často je v tomto kroku také užitečné harmonizovat atributy, aby byly sro-

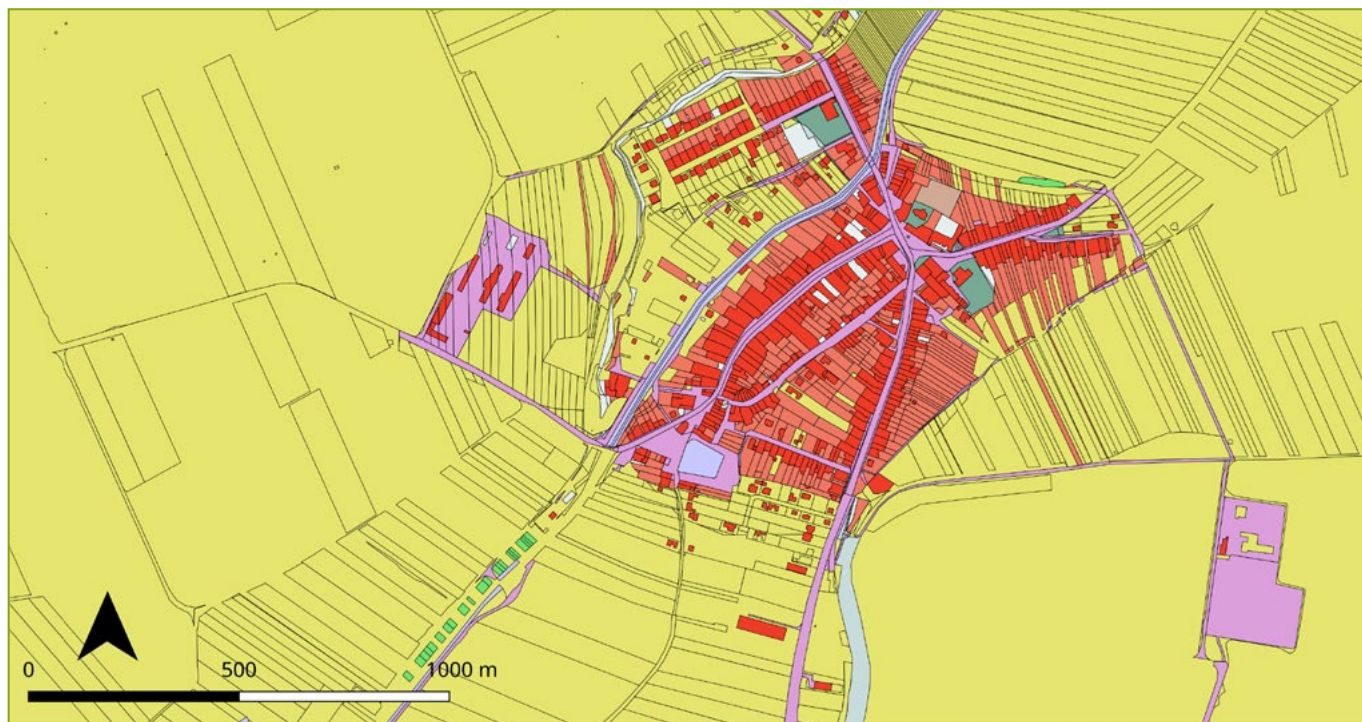
zumitelné pro všechny uživatele. Zde je příklad transformace atributu "zemědělská omezení" z Evropské databáze půd. V tomto konkrétním případě jsou data v 8bitové rastrové reprezentaci: {0: "Bez informací", 1: "Bez omezení pro zemědělské využití", 2: "Štěrkovitá (více než 35 % štěrku o průměru < 7,5 cm)", 3: "Kamenitá (přítomnost kamenů o průměru > 7,5 cm)"..., 18: "Permafrost", 255: "Není známo"}.

Důležitými ukazateli při plánování hospodaření na pozemcích jsou také jejich geomorfologické charakteristiky (nadmořská výška, svažitost, orientace svahu, index topografické vlhkosti (TWI¹⁶), který ukazuje lokality, náchylné k akumulaci vody). Kromě toho je velmi často, jako v případě topografických charakteristik odvozených z digitálního výškového modelu nazvaném EU-DEM [78], rozlišení rastrových dat poměrně jemné a existuje několik jejich kategorií, které spadají do jednoho prvku OLU. Jako příklad takový případ ukazuje Obrázek 28, kde se jeden prvek OLU protíná s mnoha pixely, které mají různé hodnoty:

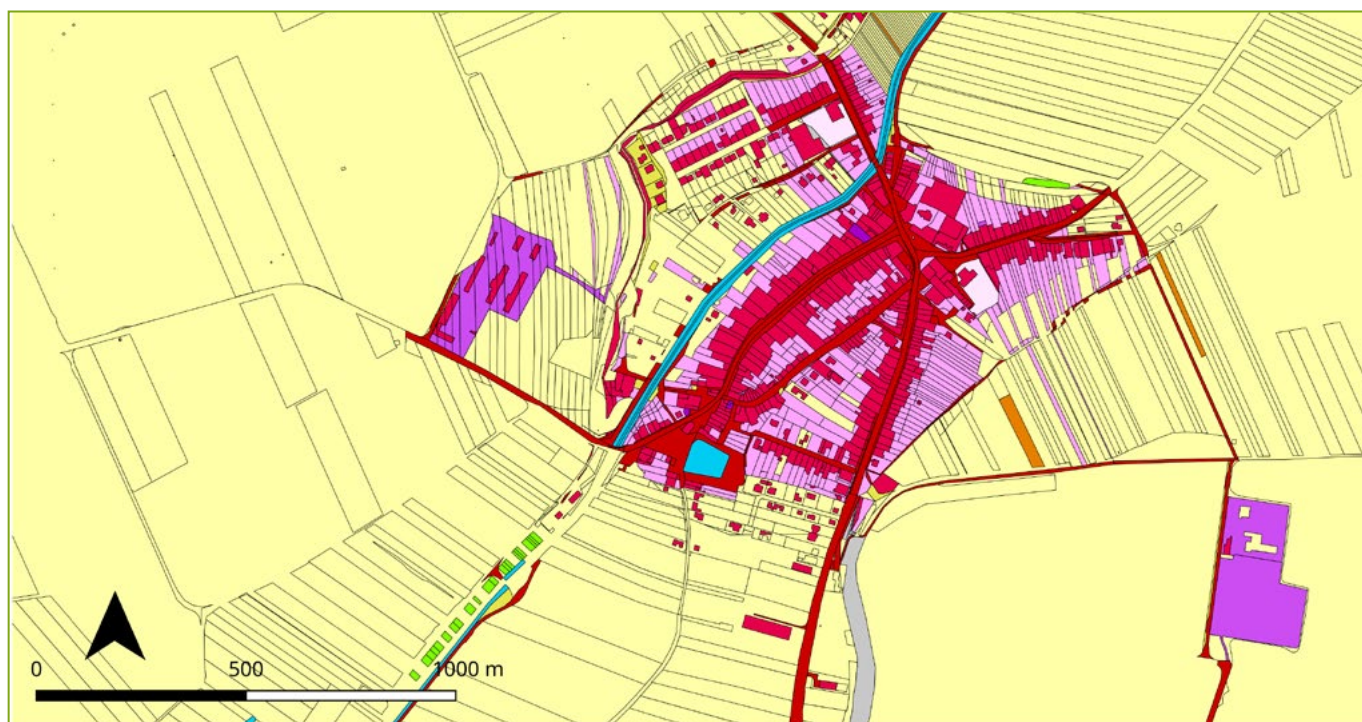
Operátor, který vytváří tematický pohled, by pak na základě těchto statistik mohl rozhodnout, do které tematické kategorie ten či onen objekt OLU patří. Například logika v případě svažitosti by mohla vypadat tak, že v případě, kdy směrodatná odchylka je menší než 3 stupně, přiřadí k objektu OLU průměrnou hodnotu, nebo u kategoriální proměnné jako orientace svahu – v případě, že některá hodnota má počet četností větší, než že činí 50 % hodnot – přiřadí ji k prvku.

¹⁵ <https://epsg.io/4326>

¹⁶ TWI = Topographic Wetness Index



Obrázek 29: Zobrazení dle klasifikace využití území HILUCS.



Obrázek 30: Zobrazení tříd krajinného pokryvu dle CORINE.

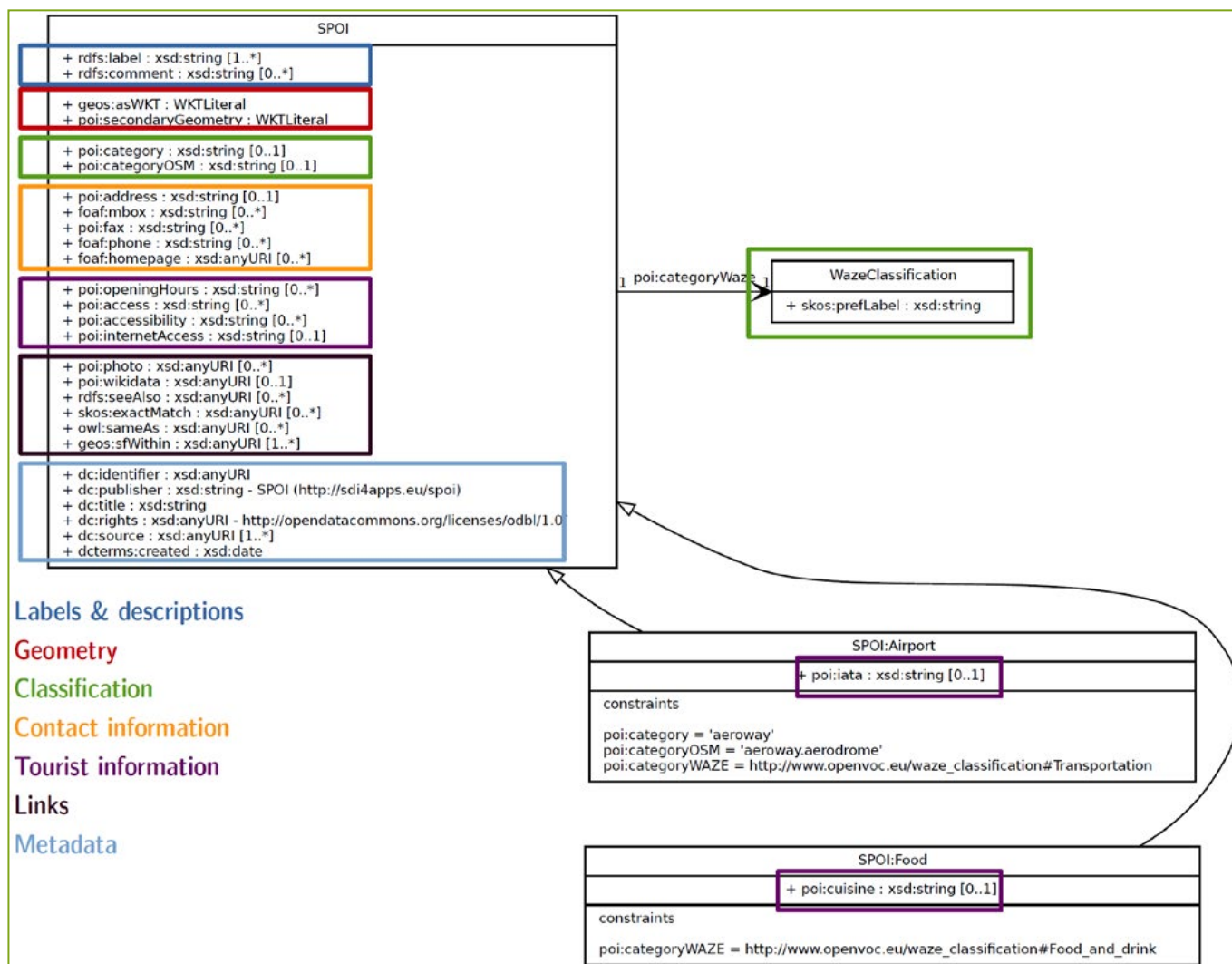
Algoritmus je v podstatě velmi podobný tomu, který byl použit při generování první verze OLU, s tím rozdílem, že nyní se geometrie prvků neořezávají o zájmové území, a tak je možné definovat pořadí vrstev podle priority tématu, takže při řešení například zemědělského nebo půdního tématu by mělo smysl přiřadit půdním blokům LPIS vyšší prioritu než pozemkům RÚIAN. Pro většinu ostatních aplikací v ČR budou mít data RÚIAN nejvyšší prioritu. Také záleží na tom, pro jaký účel analýz nebo vizualizace jsou data připravována, používají se totiž různé tematické atributy. Právě v tomto kroku by bylo možné

použít dříve popsanou logiku – vybrat pro téma právě homogenní objekt.

Následuje několik ukávek využití výstupů OLU v různých oblastech.

První mapa ukazuje třídu využití krajiny na vybraných pozemcích dle klasifikací HILUCS, viz Obrázek 29.

Druhá mapa ukazuje třídu krajinného pokryvu na vybraných pozemcích dle klasifikace CORINE, viz Obrázek 30.



Obrázek 31: Zobrazení dle klasifikace využití území HILUCS.

9.3. CHYTRÉ BODY ZÁJMU - SMART POINT OF INTEREST

V následujícím textu se zaměříme na další velkou datovou sadu a to jsou body zájmu. Body zájmu mohou být využity pro propagaci lokálních produktů, zkracování obytných řetězců, venkovské turistiky a podobně.

Otevřená a bežešvá datová sada, založená na principech otevřených propojených dat, viz kapitola 5.5. Otevřená propojená data – Linked open data, nazvaná Smart Points of Interest (SPOI) vznikla v rámci projektu SDI4apps – Uptake Of Open Geographic Information Through Innovative Services Based On Linked Data¹⁷ ve spolupráci Katedry geomatiky na Západočeské univerzitě, Českého centra pro vědu a společnost a Baltic Open Solution Centre (Lotyšsko). V době jejího vzniku šlo o první větší datovou sadu primárně zaměřenou na otevřená propojená data.

V současné době je v databázi k dispozici více než 30 milionů bodů, které může uživatel využívat při dodržení podmínek licence ODbL (Open Database License). Data je možné prohlížet v mapovém prohlížeči využívajícím knihovny HS Layers nebo stahovat pomocí SPARQL (SPARQL Protocol and RDF Query Language) endpoint (obě aplikace jsou krátce popsány níže). [53]

Veškeré informace o datech SPOI jsou k dispozici na [79]. Zájemci zde mohou nalézt nejen odkazy na vizualizaci dat a možnost stažení dat, ale také aktuální verzi datového modelu, schéma datové harmonizace, seznam realizovaných změn, prezentaci představující SPOI a metadata ve formě DOAP (Description of a Project) a VoID (Vocabulary of Interlinked Datasets). Další informace o SPOI, především týkající se harmonizace a integrace dat z různých zdrojů a popisu tvorby datového modelu jsou k dispozici například v [80].

9.3.1. Datový model SPOI

Data SPOI nejsou uložena v tradiční relační databázi, ale jako grafová struktura ve formě RDF (Resource Description Framework) trojice (viz ukázka kódu níže). Tento formát je v souladu s požadavky na tzv. pětihvězdičková Linked Open Data (popsané v kapitole 4. Otevřená a sdílená data), kde předpokladem k získání čtvrté hvězdy je identifikace datových položek pomocí URI (Uniform Resource Identifier) a páté pak zápis dat v univerzálním RDF formátu. [53]

¹⁷ sdi4apps.eu

Při tvorbě datové sady byla využita převážně existující data z volně dostupných zdrojů. Jedná se například o OpenStreetMap, Natural Earth nebo GeoNames.org. Kromě těchto globálních datových sad jsme zpracovali také lokální datové sady od partnerů projektů Citadel on the Move (cca 30 lokálních datových sad), SDI4apps (Zemgale, Pošumaví), FOODIE (Belluno), OpenTransportNet (Antverpy, Issy) a Peregrinus Silva Bohemica (Pošumaví, Bavorsko).

Velkou výhodou řešení SPOI oproti konkurenčním datovým sadám je kromě faktu, že se jedná o otevřená a bezesová data, především existence jednotného datového modelu (Obrázek 31). Ten maximálně využívá existující slovníky Linked Data a umožňuje snadné propojení s dalšími objekty.

SPOI data jsou uložena v prostředí Virtuoso (virtuoso.openlinksw.com), které představuje moderní (tzv. hybridní) serverové řešení umožňující kromě tradičních relačních databází ukládat také grafové struktury ve formátu RDF. Z tohoto důvodu je tento nástroj populární ve velkém množství datových sad využívajících přístup Linked data (například DBpedia nebo LinkedGeoData). V následujících podkapitolách jsou uvedeny příklady různých aplikací využívající datovou sadu SPOI.

9.3.2. SPARQL endpoint

Pro uživatele SPOI není ani zdaleka tak důležitý způsob uložení dat, ale především jejich dostupnost. Ta je zajištěna pomocí tzv. SPARQL endpoint. Jedná se o klienta dotazovacího jazyka SPARQL [82], který na základě uživatelského dotazu vygeneruje požadovaná data. Dotazovací jazyk SPARQL je podobný tradičnímu SQL (Structured Query Language) nejen posledními písmeny zkratky, ale také strukturou dotazu [53]. Následující ukázka představuje grafické uživatelské rozhraní Virtuoso SPARQL Query Editor¹⁸ pro dotazování pomocí jazyka SPARQL (Obrázek 32).

SPARQL endpoint poskytuje data v celé řadě formátů jako například CSV, různé syntaxe RDF, JSON, HTML (HyperText Markup Language), XML (Extensible Mar-



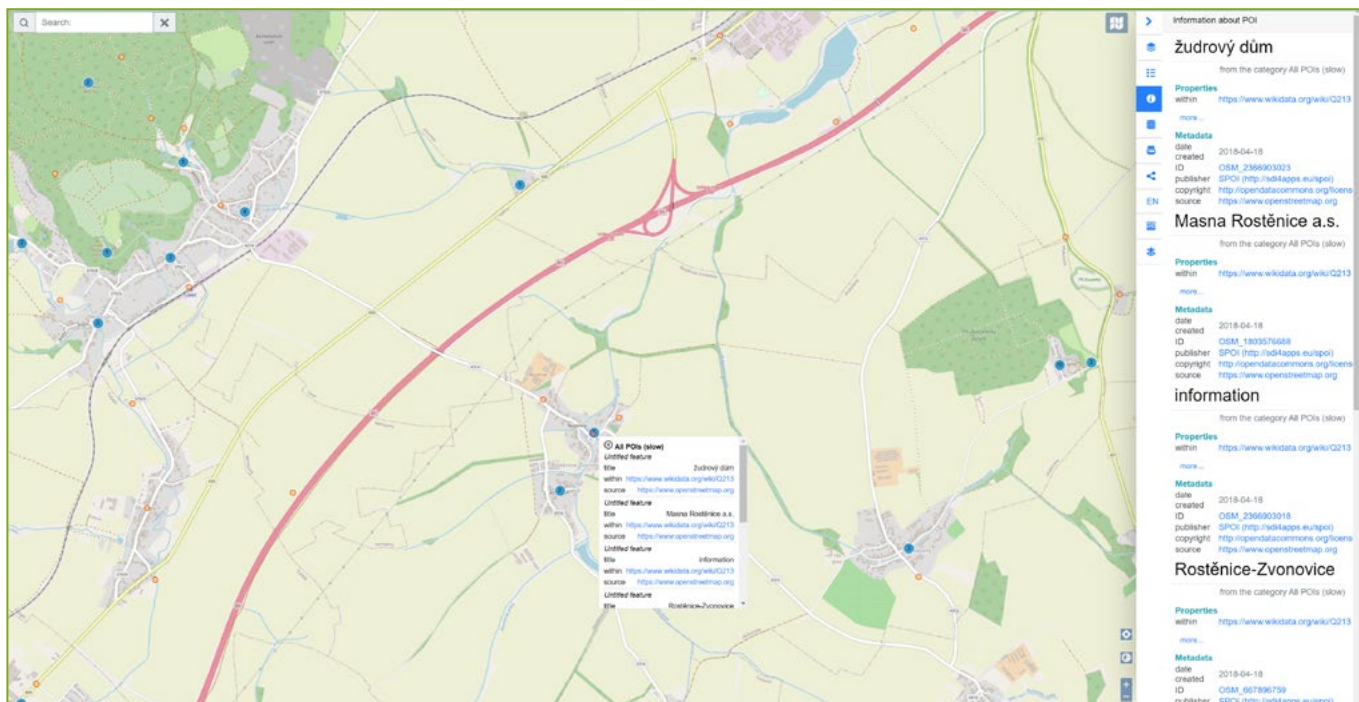
Obrázek 32: Grafické uživatelské rozhraní Virtuoso SPARQL Query Editor pro dotazování pomocí jazyka SPARQL (převzato z [81]).

kup Language), TSV (Tab-Separated Values) a podobně. Vzhledem k tomu, že se jedná o strukturovaná a v mnoha případech i kvalitně sémanticky popsána data, není problém jejich transformace do některého z textových formátů (například KML – Keyhole Markup Language nebo GML – Geography Markup Language) určených pro geografické informační systémy. Navíc většina současných GIS nástrojů dokáže zpracovat i formát CSV. [82]

9.3.3. Mapový klient

“Podobně jako SPARQL endpoint je z webové stránky SPOI dostupný i mapový klient (Obrázek 33). Klient je založený na technologii HS Layers, která představuje prostředí pro tvorbu mapových aplikací využívající především javascriptové knihovny OpenLayers a ExtJS, a je šířena jako open source pod GNU/GPL licenci.

¹⁸ Více informací na: <https://data.europa.eu/sparql?help=intro>



Obrázek 33: Mapový klient SPOI (data z října 2021).

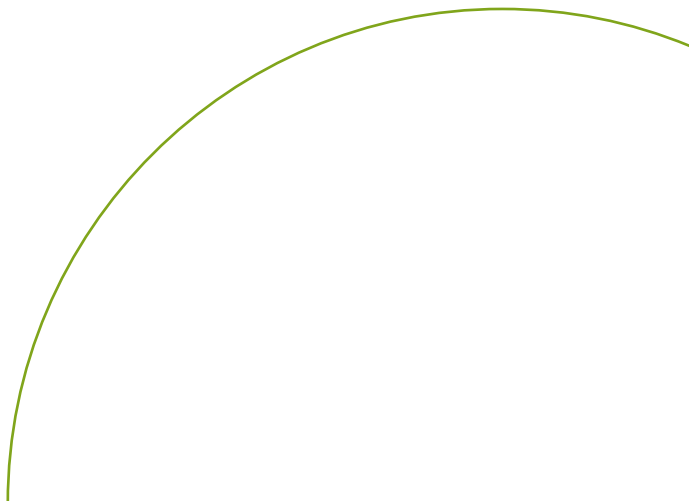


Obrázek 34: Propojení databáze OLU, SPOI a 3D digitálního modelu terénu do jedné webové aplikace [83].

9.4. PROPOJENÍ DATOVÝCH SAD

SPOI klient není pouze tradičním vizualizačním nástrojem, který umožňuje pohyb v mapě, změnu měřítka nebo zapínání jednotlivých vrstev. Uživatel má možnost měnit grafický vzhled jednotlivých zobrazených dat a také data editovat či přidávat nové body. Veškeré změny jsou prozatím uložené v nové datové sadě a do SPOI budou exportovány během pravidelných aktualizací databáze, která probíhá čtyřikrát ročně”. [53]

Výhoda propojených dat a jejich interoperability je v možnosti jejich kombinování a propojování za vzniku nových produktů a služeb. Jedním z takových příkladů je propojení datového modelu OLU a SPOI, respektive propojení samotných datových sad spolu s dalšími datovými zdroji (konkrétně digitálního modelu terénu) za vzniku aplikace, která může napomoci plánování odbytu místní zemědělské produkce do restaurací. Ukázkou takovéto aplikace najdete na Obrázku 34, či na [83].



10. DÁLKOVÝ PRŮZKUM ZEMĚ

Dálkový průzkum Země (DPZ), také označovaný jako Earth Observation (EO), je jednou z nejčastěji používaných metod vyhodnocování jevů na zemském povrchu a je i často užíván právě pro potřeby zemědělství. Výrazným rysem je, že získávání údajů probíhá bez přímé interakce se sledovaným objektem na základě elektromagnetického záření [84]. Termín "dálkový průzkum Země" byl poprvé použit na počátku 60. let 20. století k označení jakéhokoliv způsobu pozorování Země z dálky, zejména v souvislosti s leteckou fotografií, která byla v té době hlavním používaným senzorem. V současné době zahrnuje několik technologií bezpilotní letadla (UAV¹⁹), letadla a družice. [85]

Vlastní snímání dat DPZ může být prováděno různými technologiemi. Výběr technologií je obvykle dán účelem pozorování. To může být tedy realizováno prostřednictvím aktivních (radar se syntetickou aperturou, LiDAR) i pasivních (optické a termální dálkoměry, multispektrální a hyperspektrální) senzorů. Tato data pak mohou být použita k získání široké škály znalostí.

Obecně pod pojmem DPZ dnes nechápeme jenom vlastní monitoring Země, ale i výzkum v řadě technologií, jako jsou fyzikální základy měření, přes družicové technologie, přenos dat až po ukládání a interpretaci dat. Data DPZ jsou pak převedena na relevantní znalosti, které jsou poskytovány širokému spektru potenciálních koncových uživatelů. Data DPZ pak mohou být i kombinována s dalšími daty jakou jsou meteorologická data, data ze senzorů, strojů, půdních a dalších dat. [86]

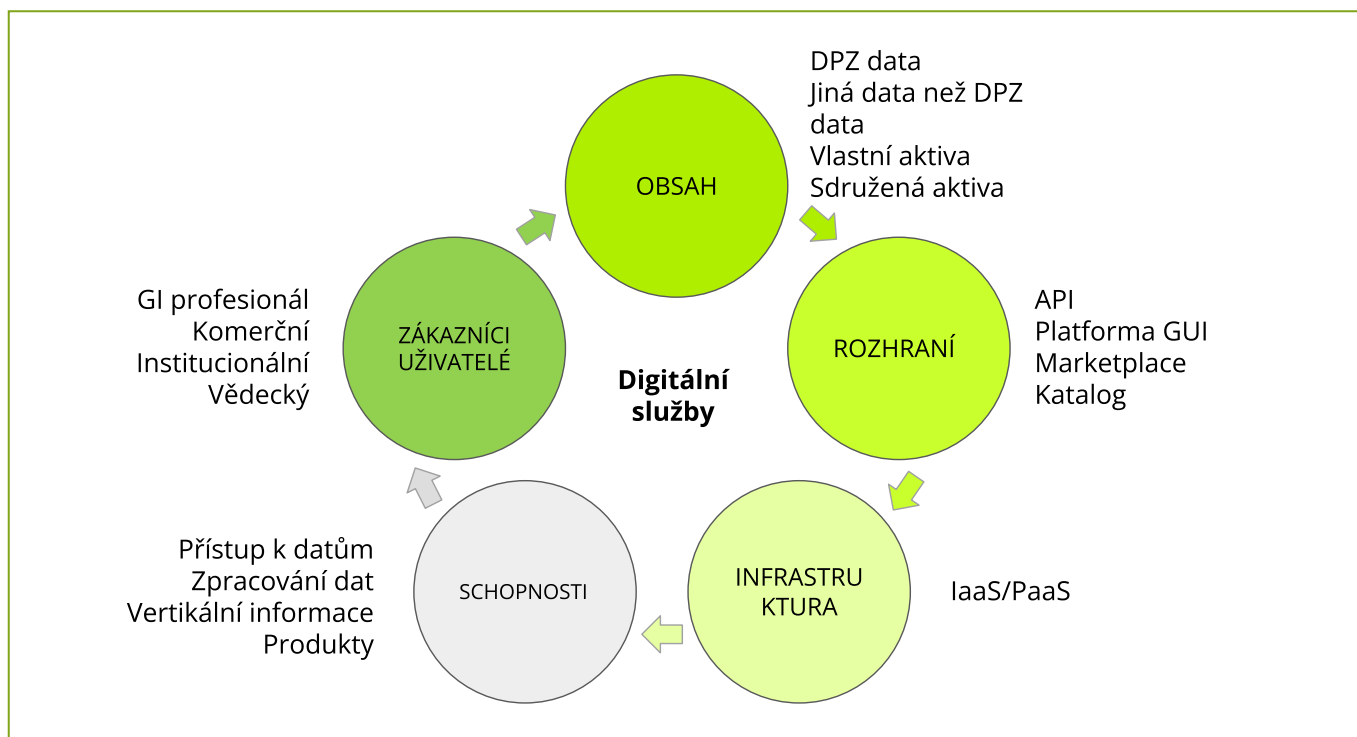
Tradičními družicovými systémy používanými pro zemědělské účely jsou družice Landsat. Tyto družice monitorující povrch Země již od roku 1972 (Landsat MSS), a potom od roku 1982 (Landsat Thematic Mapper). Ty poskytují snímky v optické a infračervené části spektra. Jejich výhodou spočívá především ve volné dostupnosti snímků s historickou posloupností, která zaručuje statisticky ověřitelné výsledky po dobu několika desetiletí. Snímky obsahují více než 7 spektrálních pásem. Nevýhodou je jejich nízké až střední prostorové rozlišení a opakování snímání stejné oblasti přibližně každých 16 dní.

V současné době lze bezplatně využívat snímky z družice Landsat 8 s 11 spektrálními pásmy. Na družicový systém Landsat navázal evropský program Copernicus s družicemi Sentinel. Copernicus je evropský nástroj pro DPZ. Jedná se o komplexní systém družic pro různé účely a zahrnuje i pozemní systémy. Data jsou zpracovávána a poskytují uživatelům spolehlivé a aktuální informace prostřednictvím souboru služeb souvisejících s otázkami životního prostředí, zemědělství a bezpečnosti. Klíčem k uvolnění obrovského potenciálu programu Copernicus je snadný přístup k jeho datům a informačním produktům. Družice Sentinel-2 shromažďuje v optické a infračervené části spektra data ze 13 spektrálních pásem s prostorovým rozlišením od 10 m do 60 m a s opakováním 5 dnů. V mikrovlnné části spektra se družice Sentinel-1 používají k měření výšky terénu a vegetace. Optické snímky z družic Sentinel-2 a radarové záznamy z družic Sentinel-1 jsou volně dostupné. Vedle dat Landsat a Copernicus, existuje i řada komerčních družic s vyšším rozlišením. Některé platformy budou zmíněny dále. [87]

S rostoucí dostupností bezplatných dat DPZ (satelitní snímky Copernicus atd.) roste počet uživatelů, kteří mohou tato data efektivně využívat. Aby bylo možné data interpretovat jako informace, jsou zapotřebí různé platformy pro poskytování dat DPZ, jejich integraci, přístup a analýzu. Pro přístup k datům DPZ potřebujeme mít k dispozici dodavatelské platformy, kde budou mít zúčastněné strany přístup k vybraným datům a budou moci provádět specifické služby (Obrázek 35).

Pro usnadnění a standardizaci přístupu k družicovým datům financuje Evropská komise vývoj pěti cloudových platform poskytujících centralizovaný přístup k datům, informacím a nástrojům pro jejich zpracování v rámci programu Copernicus. Tyto platformy jsou známé pod názvem Data and Information Access Services (DIAS) [88].

¹⁹ UAV = Unmanned Aerial Vehicle



Obrázek 35: Digitální služby (upraveno dle [88]).

DIAS nabízí zdroje pro zpracování, nástroje a doplňkové zdroje dat za komerčních podmínek, aby dále usnadnila přístup k datům Sentinel a službám a produktům programu Copernicus.

Podobně jako DIAS poskytuje Amazon Web Services (AWS) [89] cloudové výpočetní platformy a rozhraní API na komerčním základě. Kromě výpočetních, úložných a databázových technologií nabízí také služby pro strojové učení a umělou inteligenci. Poskytuje i přístup k satelitním datům Landsat-8, Sentinel-2 a CBERS-4.

Příklady platform, které poskytují správu, ukládání, přístup a zpracování velkých dat DPZ na serveru bez nutnosti stahovat velké množství velkých datových sad DPZ, jsou Earth Observing System Data and Information System (EOSDIS) NASA, Google Earth Engine (GEE), Open Data Cube, European Data Cube, Sentinel Hub, System for Earth Observation Data Access, Processing and Analysis for Land Monitoring (SEPAL), OpenEO a JEODPP. [90]

NASA podporuje úplné a otevřené sdílení všech svých dat pro výzkumné a aplikační komunity, soukromý průmysl, akademickou obec a širokou veřejnost. Aby se vyhovělo potřebám těchto různých komunit, poskytl NASA v rámci systému EOSDIS (Earth Observing System Data and Information System) různé způsoby, jak data objevovat, přistupovat k nim a využívat je. Poskytuje komplexní možnosti správy vědeckých dat NASA o Zemi z různých zdrojů - družic, letadel, terénních měření a různých dalších programů. Pro družicové mise EOS poskytuje EO-

SDIS funkce pro řízení a kontrolu, plánování, sběr dat a počáteční (úroveň 0) zpracování. [91]

Google Earth Engine (GEE) je cloudová platforma, která umožňuje zpracování dat dálkového průzkumu Země ve velkém měřítku. Mezi aplikace patří: detekce odlesňování, klasifikace půdního pokryvu, odhad lesní biomasy a uhlíku a mapování světových bezlesí. Výhoda spočívá v rychlosti výpočtů, protože zpracování je zadáváno na serverech společnosti Google. Platforma poskytuje řadu neustále aktualizovaných datových souborů. Jeden z nástrojů GEE, Earth Engine Code Editor, je webové integrované vývojové prostředí, které usnadňuje vývoj složitých geoprostorových pracovních postupů. [92]

Open Data Cube (ODC), dříve známá jako Australian Geoscience Data Cube (AGDC), je analytický rámec složený z řady datových struktur a nástrojů, které usnadňují organizaci a analýzu dat DPZ. ODC umožňuje katalogizaci rozsáhlých souborů dat DPZ a přístup k nim a manipulaci s nimi prostřednictvím sady nástrojů příkazového řádku a rozhraní API v jazyce Python. [93]

Euro Data Cube (EDC) je služba pro využívání dat DPZ, která poskytuje všechny důležité produkty DPZ a odvozených dat na jednom místě. Poskytuje možnost zpracovávat úlohy pro rozsáhlé analýzy a strojové učení pro analýzu jevů z různých hledisek, poskytování více zdrojů dat a porovnávání a korelaci několika proměnných současně. [94]

Sentinel Hub (SH) je soukromá platforma nabízející veřejný přístup vyvinutá společností Sinergise, která poskytuje

přístup k datům Sentinel a vizualizační služby. Na rozdíl od Google Earth Engine omezuje SH přístup k funkcím v různých platebních plánech. Bezplatný plán umožňuje omezené možnosti - prohlížení, výběr a stahování nezpracovaných dat. Placený přístup umožňuje přístup k datům prostřednictvím protokolů OGC a specifického rozhraní API, zpracování dat, přístup k datům mobilní aplikace, vyšší limity přístupu ke zdrojům a technickou podporu. Funkce platformy SH jsou zpřístupněny prostřednictvím služeb OGC a rozhraní RESTful API. K dispozici je také webové rozhraní, které umožňuje konfiguraci konkrétních služeb [90, 95].

Digital Globe GBDX [96] je jednou z nejpokročilejších platform pro poskytování informačních služeb založených na EO s využitím obrovského archivu senzorů Digital Globe VHR a poskytuje pokročilé prostředky PaaS a IaaS pro vývoj komplexních pracovních postupů zpracování pomocí specializovaného API a specifické metody pro integraci svých algoritmů do pracovního postupu (pomocí kontejnerů Docker). Kromě toho nabízí i hotové algoritmy pro zpracování dat (např. ortorektifikace, kompenzace atmosféry).

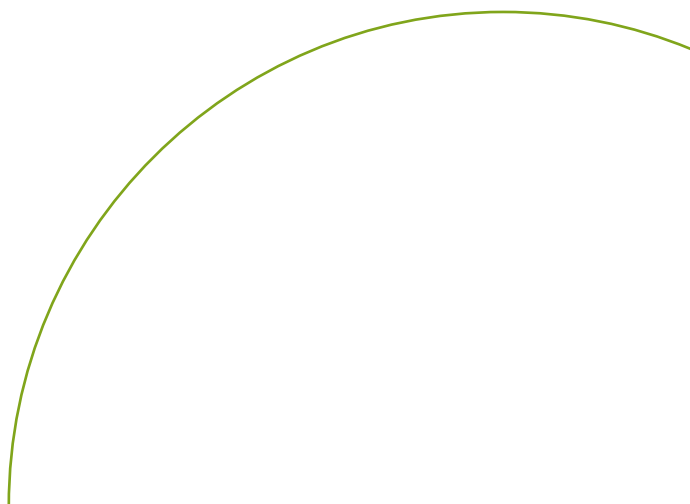
Platforma Planet Labs [97] nabízí obrovský archiv snímků. Kromě toho nabízí API pro vyhledávání a stahování, které umožňuje vytvářet pracovní postupy automatického zpracování schopné získávat informace ze snímků. Výhodou Planet oproti tradičním satelitním provozovatelům je nižší cena snímků a jejich inovativní obchodní modely.

European Open Science Cloud (EOSC) je digitální platforma pro vědeckou komunitu, která poskytuje bezproblémový přístup k datům a interoperabilním službám, jež se zabývají celým cyklem výzkumných dat, od jejich objevení a vytěžování až po ukládání, správu, analýzu a opakované použití napříč hranicemi a vědními obory [98]. Datové centrum a platforma NextGEOSS jsou součástí EuroGEO, evropského příspěvku ke GEO. Datové centrum obsahuje jak družicová data, tak data in-situ, pomocná data a katalogizují se také aplikace a služby využívající data programu Copernicus. NextGEOSS je napojen na několik cloudových služeb včetně DIAS. [9]

10.1. LETECKÉ A UAV PLATFORMY

Dalším důležitým zdrojem dat z pozorování Země pro zemědělství jsou letecké snímky pořízené pilotovanými letadly. Tato data jsou podstatně dražší než družicová, ale jsou poskytována s vyšším rozlišením - od 0,10 m do 0,50 m. Tato data se využívají především pro stanovení velikosti produkčních bloků půdy pro systém identifikace pozemků (LPIS) a pro kontrolu zemědělských postupů na pozemcích zemědělců. Letecké údaje většina zemí Evropské unie obvykle získává s jednoletým nebo dvouletým opakováním.

Bezpilotní letadla (UAV,) jsou v současnosti nejslibnějším zdrojem dat s vysokým potenciálem a efektem pro přímé zásahy do rostlinné výroby díky své operativnosti a ceně dat na hektar (s klesajícími pořizovacími náklady na UAV a klesajícími cenami senzorů). [9]



11. SENZOROVÉ TECHNOLOGIE

O senzorech v zemědělství pojednává samostatná publikace technologické platformy Využití senzorů v zemědělství [99], proto je zde uveden pouze stručný úvod do problematiky. Senzor je zařízení, které detekuje a reaguje na určitý typ vstupu z fyzického prostředí. Konkrétním vstupem může být světlo, teplo, pohyb, vlhkost, tlak nebo kterýkoli z mnoha dalších jevů prostředí. Výstupem je zpravidla signál, který je převeden na člověkem čitelný displej v místě senzoru nebo je elektronicky přenášen po síti ke čtení nebo dalšímu zpracování.

Zemědělství je velmi nepředvídatelné, a to z důvodu velké závislosti na počasí a podmínkách prostředí (např. déšť, teplota, vlhkost, krupobití), nepředvídatelných událostech (např. nemoci zvířat, škůdci), jakož i na kolísání cen na zemědělských trzích. Kombinace a analýza datových toků poskytovaných senzory v reálném čase může pomoci při informovanějším rozhodování a rychlé reakci na změny a nepředvídatelné události. Například kombinací údajů ze senzorů o úrodnosti půdy spolu s webovými službami pro předpověď počasí by bylo možné lépe rozhodovat o přesnějším zavlažování a hnojení plodin; údaje ze senzorů lze využít k monitorování agropotravinářských parametrů v reálném čase, jako je pH, teplota, vlhkost půdy nebo tok kyslíku.

Pomocí senzorů ve spojení s internetovým připojením je možné nepřetržitě sledovat různé plodiny a pozemky, i když jsou vzdálené, předpovídat a zároveň kontrolovat výnosy a kvalitu potravin. Dalším důležitým aspektem je georeferencování, které umožňuje zemědělským strojům přesně zásobovat denní potřeby různých druhů plodin, a to bez lidského zásahu nebo jen s minimálním zásahem [100]. IoT se používá i v živočišné výrobě. [101]

On-line měřicí zařízení je často i součástí strojů. Spektrální charakteristiky plodin lze získat pomocí přístrojů umístěných na strojích, např. N senzor²⁰, CropSpec²¹, Crop circle ACS²², OptRX, ISARIA²³ nebo GreenSeeker²⁴. Jednou z důležitých informací je nasycení půdy srážkami. Zemědělci nebo specializované služby (někdy celostátní) měří hodnotu nasycení půdy vlhkoměrem.

Hodnoty vlhkosti půdy a předpovědi počasí jsou nejdůležitějšími informacemi pro provádění agrotechnického opatření - "použij optimální množství ve správný čas". [9]

²⁰ Více viz <https://www.yara.co.uk/crop-nutrition/farmers-toolbox/n-sensor/>

²¹ Více viz <https://www.topconpositioning.com/crop-sensing/canopy-sensing/cropspec>

²² Více viz <https://hollandscientific.com/portfolio/crop-circle-ac-430/>

²³ Více viz <https://www.plantsystems.co.uk/isaria/177>

²⁴ <https://agriculture.trimble.com/product/greenseeker-system/>



12. OBČANSKÁ VĚDA A CROWDSOURCING

Citizens Science neboli Občanská věda je výzkum prováděný, zcela nebo částečně, amatérskými nebo profesionálními vědci [102]. V zemědělské praxi se nejčastěji setkáváme se sběrem dat prováděným amatéry tzv. crowdsourcing nebo také dobrovolná geografická informace (Volunteered Geographic Information, VGI). I tato problematika je řešena v publikaci [99], proto opět jenom stručně.

Crowdsourcing je model získávání zdrojů, při kterém jednotlivci nebo organizace získávají zboží a služby, včetně nápadů, hlasování, mikroúkolů a financí, od velké, relativně otevřené a často rychle se vyvíjející skupiny účastníků. Crowdsourcing lze chápat jako jednu z metod Citizens Science [103] umožňující zapojení veřejnosti do vědeckého výzkumu [104].

V oblasti sběru prostorových informací nebo pozorování Země se často používá termín občanské observatoře [105]. Tímto termínem se obvykle rozumí metody komunitního monitorování využívající nové aplikace a senzory pozorování Země zabudované do přenosných nebo mobilních osobních zařízení [106, 107, 108].

Dalším termínem, který se v této souvislosti často používá, je dobrovolná geografická informace (Volunteered Geographic Information, VGI) [109], což je využití nástrojů k vytváření, shromažďování a šíření geografických dat poskytovaných dobrovolně jednotlivci [110]. Příklady tohoto fenoménu jsou WikiMapia, OpenStreetMap a Google Map Maker. VGI lze také považovat za rozšíření kritických a participativních přístupů ke geografickým informačním systémům a za specifické téma v rámci on-line nebo webové spolehlivosti. Tyto stránky poskytují obecné základní mapové informace a umožňují uživatelům vytvářet vlastní obsah označováním míst, kde došlo k různým událostem nebo kde existují určité prvky. Při dobrovolném sběru dat je důležitou součástí způsob jejich zpracování. Příkladem je Neogeografie (New Age Geography) zaměřená na kombinaci geograficky označených dat (např. Keyhole Markup Language - KML) [111] s mapovým rozhraním pro kontextuální průzkum [112].

13. VIZUALIZACE VELKÝCH DAT

Vizualizace dat je grafické znázornění informací a dat. Jedná se o jednu z relativně jednoduchých a široké veřejnosti srozumitelných metod prezentace znalostí obsažených v datech a informacích. Pomocí vizuálních prvků, jako jsou grafy, diagramy a mapy, poskytují nástroje pro vizualizaci dat přístupný způsob, jak vidět a pochopit stav, trendy, odlehle hodnoty a vzorce v datech [113].

Množství dat vytvořených soukromým i veřejným sektorem na celém světě každoročně prudce roste. Jedná se jak o data sbíraná pomocí senzorů (IoT technologie stacionární i na strojích) a dále i z programu Copernicus a dalších družicových systémů. Veškerá tato data jsou užitečná pouze tehdy, pokud z nich lze získat cenné znalosti a na jejich základě přijímat rozhodnutí. Vizualizace velkých dat není jediným způsobem, jak lze data analyzovat, ale techniky vizualizace velkých dat nabízejí rychlý a efektivní způsob [114].

Důležitým aspektem vizualizace velkých objemů dat výběr neefektivnějšího způsobu vizualizace dat tak, aby byly dobře viditelné znalosti v nich obsažené. Za určitých okolností mohou stačit jednoduché grafické nástroje, jako jsou koláčové grafy nebo histogramy, ale u rozsáhlých a různorodých souborů dat mohou být vhodnější pokročilejší vizualizační techniky. Mezi různé příklady grafické vizualizace velkých dat patří např.:

- Lineární: Seznamy položek, položky seřazené podle jednoho prvku jako např. textové.
- 2D/Planární/geoprostorové: Kartogramy, mapy bodového rozložení, mapy proporcionálních symbolů, obrysově mapy.
- 3D/objemové: 3D počítačové modely, počítačové simulace.
- Časové: Časové osy, grafy časových řad, spojené grafy rozptylu, obloukové diagramy, kruhové grafy.
- Vícerozměrné: Koláčové grafy, histogramy, matice, mračna značek, sloupcové grafy, stromové mapy, tepelné mapy, pavoučí grafy, plošné grafy, krabicové a šikmé grafy, bublinové mračno, kulový graf, kruhové zobrazení, Ganttův diagram, síť, polární plocha, rozptylový graf (2D nebo 3D), proudový graf, klínový stohový graf.
- Stromové/hierarchické: Dendrogramy, radiální stromové grafy, hyperbolické stromové grafy.

Tyto všechny metody lze i kombinovat ve formě dashboardu [115].

13.1. PŘÍKLADY POKROČILÝCH VIZUALIZACÍ

V následujícím textu stručně popíšeme některé méně obvyklé vizualizační postupy testované v projektu DataBio a také jednu ukázkou z projektu SmartAgriHubs.

13.1.1. Komplexní integrovaná vizualizace dat

Komplexní integrovaná vizualizace dat byla důležitou součástí českých pilotních projektů v oblasti zemědělství a tato technologie byla testována také v pilotních projektech v oblasti rybolovu. Použitá technologie byla HSLayers NG²⁵. HSLayers NG je webová mapovací knihovna napsaná v jazyce JavaScript. Rozšiřuje funkčnost otevřené knihovny OpenLayers a přebírá základní myšlenky z předchozí knihovny HSLayers, ale na frontendu používá moderní JS frameworky místo ExtJS 3 a poskytuje lepší přizpůsobitelnost. Proto se k jejímu názvu přidává NG ("Next Generation"). Je stále ve vývoji a je poskytována jako software s otevřeným zdrojovým kódem. HSLayers NG je postaven modulárně, což umožňuje volné připojování a odebírání modulů, pokud jsou splněny závislosti pro každý z nich. Kontrola závislostí se provádí automaticky. Jádro frameworku je vyvíjeno s použitím AngularJS²⁶, requireJS²⁷ a Bootstrap²⁸. Tato kombinace frameworků byla zvolena především kvůli zajištění rychlého a škálovatelného vývoje a kvůli poskytnutí moderního responzivního vzhledu aplikace. Na Obrázku 36 je uveden příklad komplexní integrované vizualizace dat výnosového potenciálu, spojující běžné funkce mapového klienta s vizualizací trojrozměrných dat (3D) a podporu vizualizace senzorů a agrometeorologických dat pro zemědělce může pomoci s předpovědí počasí a lepším plánováním operací.

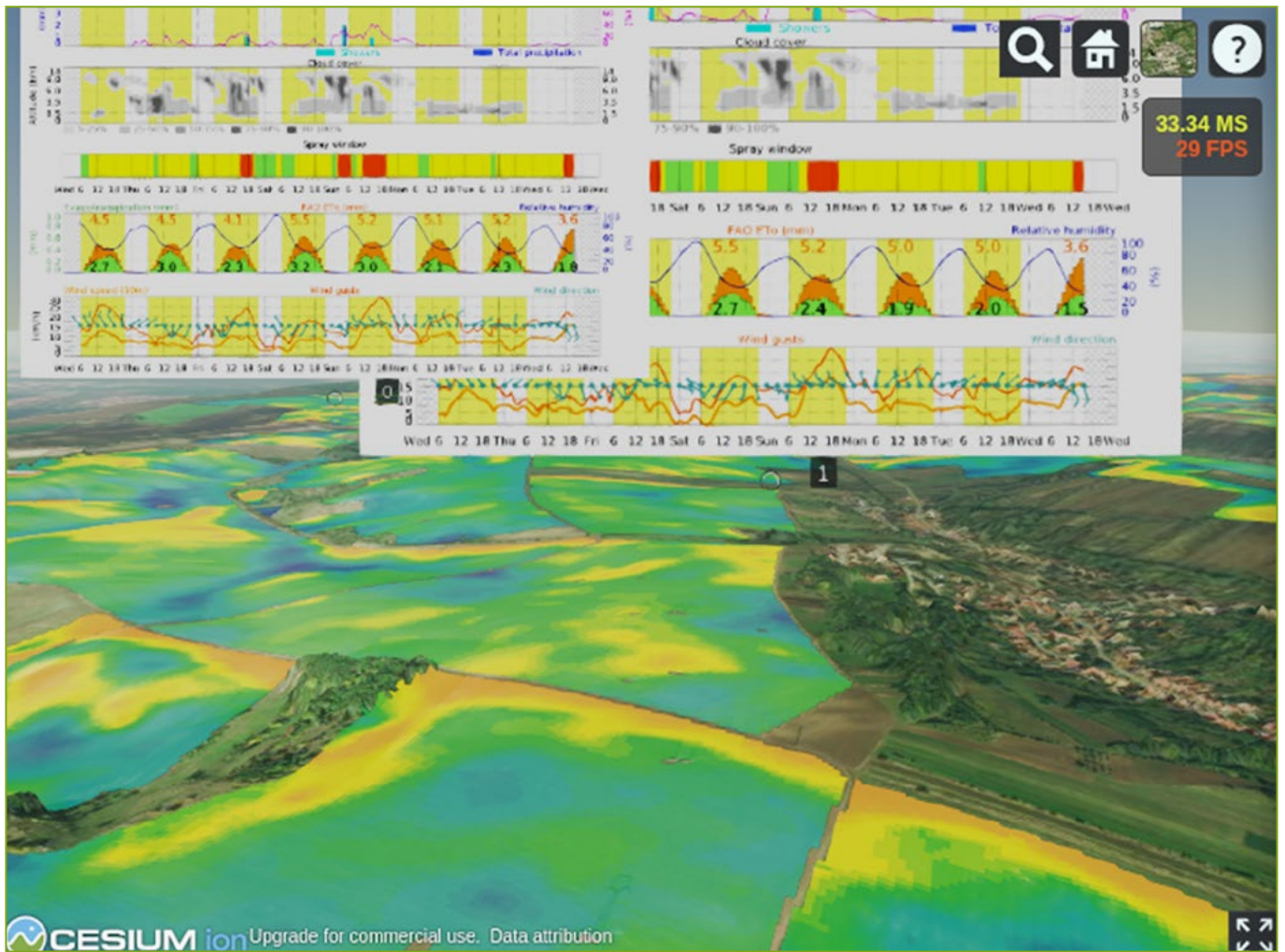
Další příklad pokročilé vizualizace pochází z experimentu s monitorováním podzemních vod a jejich vlivu na zemědělskou produkci v rámci projektu SmartAgriHubs [117].

²⁵ Více informací na <https://ng.hslayers.org/>

²⁶ Více informací na <https://angularjs.org/>

²⁷ Více informací na <https://requirejs.org/>

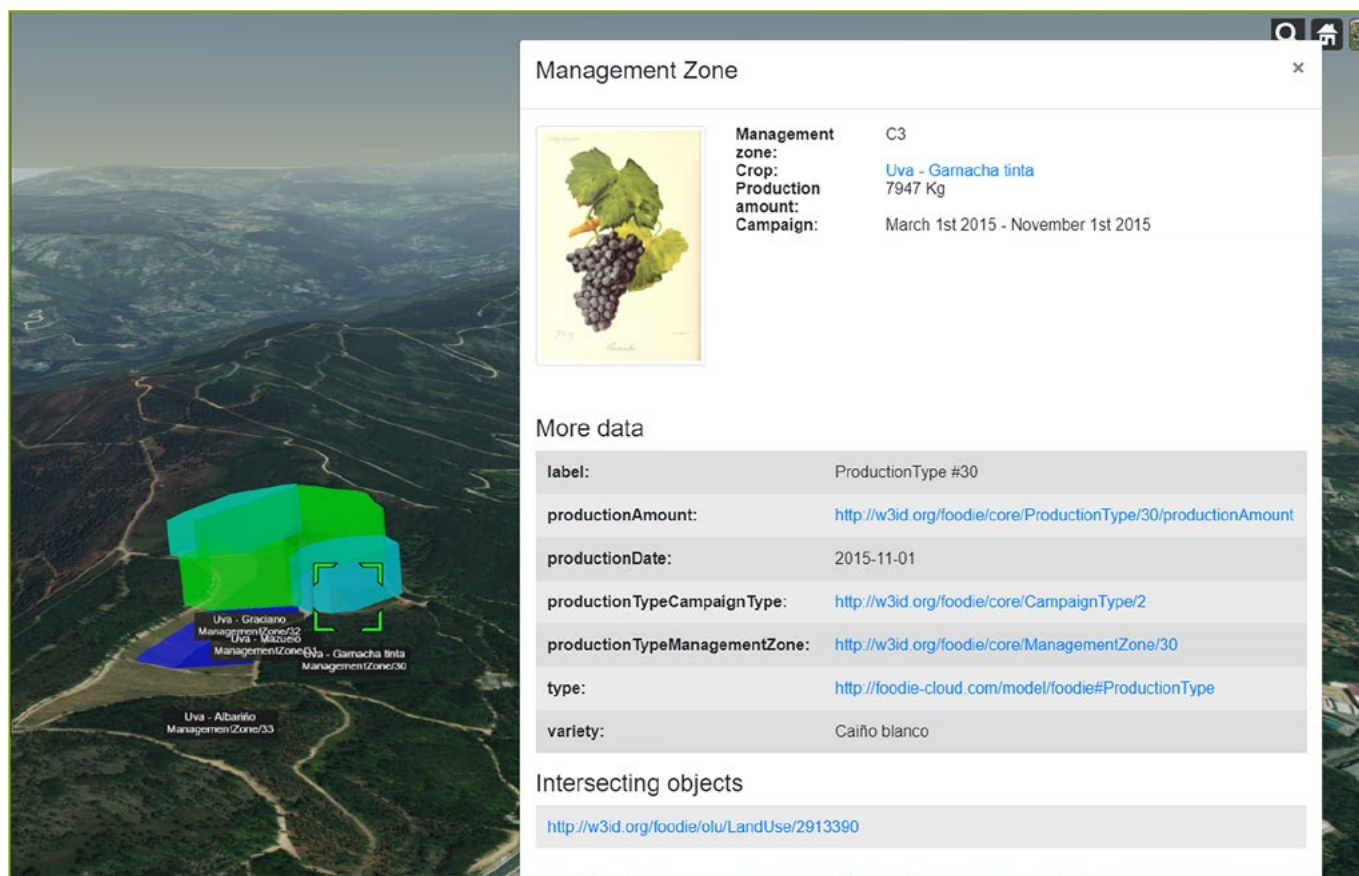
²⁸ Více informací na <https://getbootstrap.com/>



Obrázek 36: Integrace dat o výnosovém potenciálu (3D mapy) s meteorologickými daty (časové řady) (převzato z [115, 116]).



Obrázek 37: Vizualizace senzorových a satelitních dat.



Obrázek 38: Snímek obrazovky aplikace zobrazující výsledek případu použití typů plodin na základě Linked Data [převzato z [115]].

Cílem experimentu bylo vyvinout webový systém pro integraci, transformaci a využití velkého množství dat a modelů pro agrometeorologická a podzemní měření. Tato technologie umožní získat přesné informace o klimatických podmínkách a stavu vegetace, (Obrázek 37).

13.1.2. Vizualizace Linked Open Data

V kapitole 5.5. Otevřená propojená data – Linked Open Data jsme popisovali nový přístup k datům ve formě propojených dat. Vizualizace propojených dat spočívá v poskytování grafických reprezentací zajímavých aspektů v rámci tzv. sémantického webu. Rozmanitost propojených dat a jejich typů je obrovská. Příkladem otevřených propojených dat v oblasti zemědělství je datový model FOODIE (viz kapitola 9.1. FOODIE data model), který byl původně vyvinut v rámci projektu FOODIE a později rozšířen v projektu DataBio.

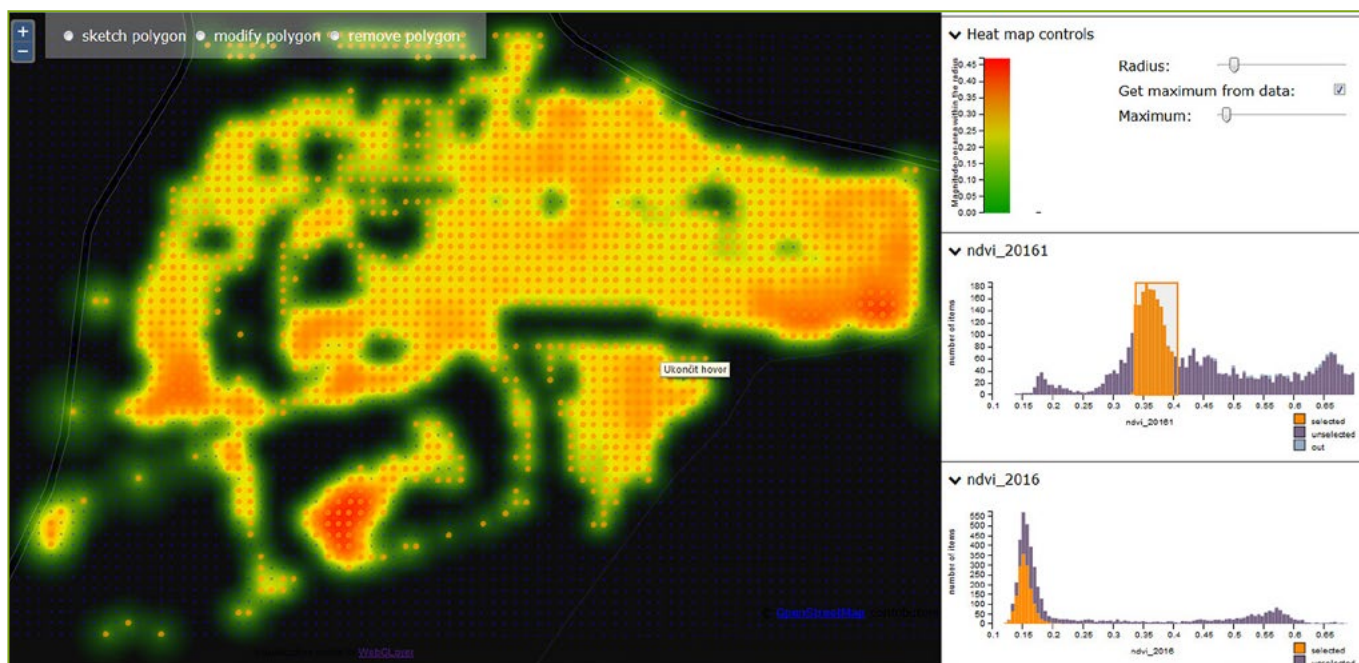
Vizualizace takovýchto dat je specifická, protože vyžaduje dotazy pomocí specifického jazyka GeoSPARQL. Výstupy jsou pak ve formátech RDF nebo JSON-LD, které prozatím běžné mapové aplikace zobrazují. Data tudíž musí být transformována do jiného formátu jako např. GeoJSON. (Obrázek 38).

13.1.3. Vysvětlující vizualizace (Explanatory visualisation)

Vysvětlující vizualizace (explanatory visualization) je proces, při kterém odborník vytváří mapy a další grafické prvky, přičemž pracuje s relativně neznámými geografickými daty. Tyto mapy zpravidla slouží k jedinému účelu a fungují jako pomůcka při snaze experta vyřešit určitý (geografický) problém. Při práci s daty by se měl expert opírat o kartografické znalosti, aby byl schopen nahlížet na data z různých perspektiv. Výsledné mapy a grafika jako takové jsou k dispozici v interaktivním prohlížečím prostředí, které stimuluje vizuální myšlení a podporuje informované rozhodování. WebGLayer je open sources knihovna v jazyce JavaScript zaměřená na rychlou interaktivní vizualizaci velkých vícerozměrných prostorových dat prostřednictvím propojených zobrazení. Knihovna je založena na WebGL a využívá GPU pro rychlé vykreslování a filtrování [118]. Příklad takové vizualizace je na Obrázku 39.

13.1.4. Kolaborativní vizualizace

Kapitola 12. Občanská věda a crowdsourcing popisuje metody občanské vědy a crowdsourcingu. Více kolaborativní vizualizace pomocí MapWhiteboard se zrodila z řady



Obrázek 39: WebGLLayer zobrazující výnosový potenciál (převzato z [115]).

rozsáhlých aktivit při psaní pomocí Dokumentů Google [119]. Na rozdíl od tradičních offline nástrojů pro zpracování textu umožňují Dokumenty Google upravovat stejný dokument více lidem [120] - ve stejnou dobu - a umožňují všem připojeným klientům vidět změny provedené v dokumentu v reálném čase díky synchronizaci všech změn mezi všemi připojenými klienty prostřednictvím serveru.

Možnost pracovat na sdíleném textu a vyhnout se nutnosti integrovat fragmenty z více zdrojových dokumentů s více styly odstranila mnoho překážek [121] spojených s tradiční úpravou dokumentů. Technologie MapWhiteboard [122] se snaží o totéž v případě tradičního používání nástrojů GIS. Celková vize této technologie je taková, že Map Whiteboard bude pro GIS tým, čím je Google Docs pro zpracování textu. MapWhiteboard se snaží vylepšit manuální a lineární pracovní postupy GIS, které byly zděděny z dob, kdy v této oblasti převládal software pro stolní počítače. Dalším rozměrem technologie MapWhiteboard je význam sdílení, nikoliv pouze sdílení dat. Prostorová data podléhají (ne)správné interpretaci; mapy jsou prostředkem k vyjádření těchto interpretací způsobem vhodným pro sdělení účelu, záměru a společného porozumění. MapWhiteboard se neomezuje pouze na úpravu sdílené datové sady. Tato technologie sdílí celou mapu a prezentuje data v kontextu všech podkladových informací, které mohou být užitečné. Umožňuje více klientům vidět identické mapové rozhraní a současně do něj kreslit, což napodobuje rozšířený případ použití, kdy sedíme u jedného stolu a společně si prohlížíme mapu, komentujeme ji, ukazujeme na "věci" a navrhujeme změny.

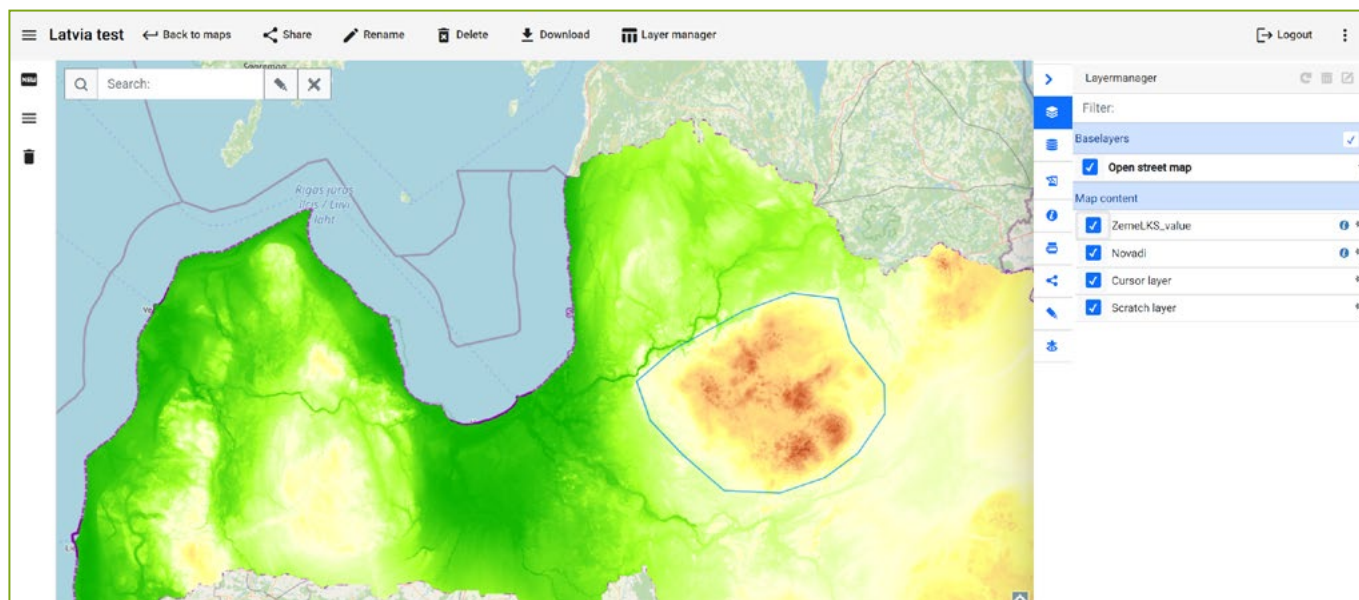
Aplikace GIS obvykle věnují většinu uživatelského rozhraní ovládacímu prvku, který zobrazuje obsah mapy. Mapy se skládají z jedné nebo více vrstev "výkresů", které jsou vykresleny nebo "naskládány" na sebe v určitém

pořadí. Vrstva může být buď předem vykreslený rastrový obrázek (nebo sada obrázků), nebo může jít o matematickou vektorovou geometrii, kterou nakreslí klientská aplikace. Mapy mají sadu přidružených ovládacích prvků, které mohou buď spustit akci po kliknutí na ně, nebo vyžadují určitou interakci s mapou v podobě kreslení, výběru, přetahování nebo zvětšování pomocí zařízení lidského rozhraní, jako je myš, stylus nebo povrch citlivý na dotyk. Všechny objekty vysílají události, které umožňují sledovat měnící se stav mapy a šířit změny do připojených klientů, aby tito mohli mezi sebou synchronizovat uživatelské rozhraní.

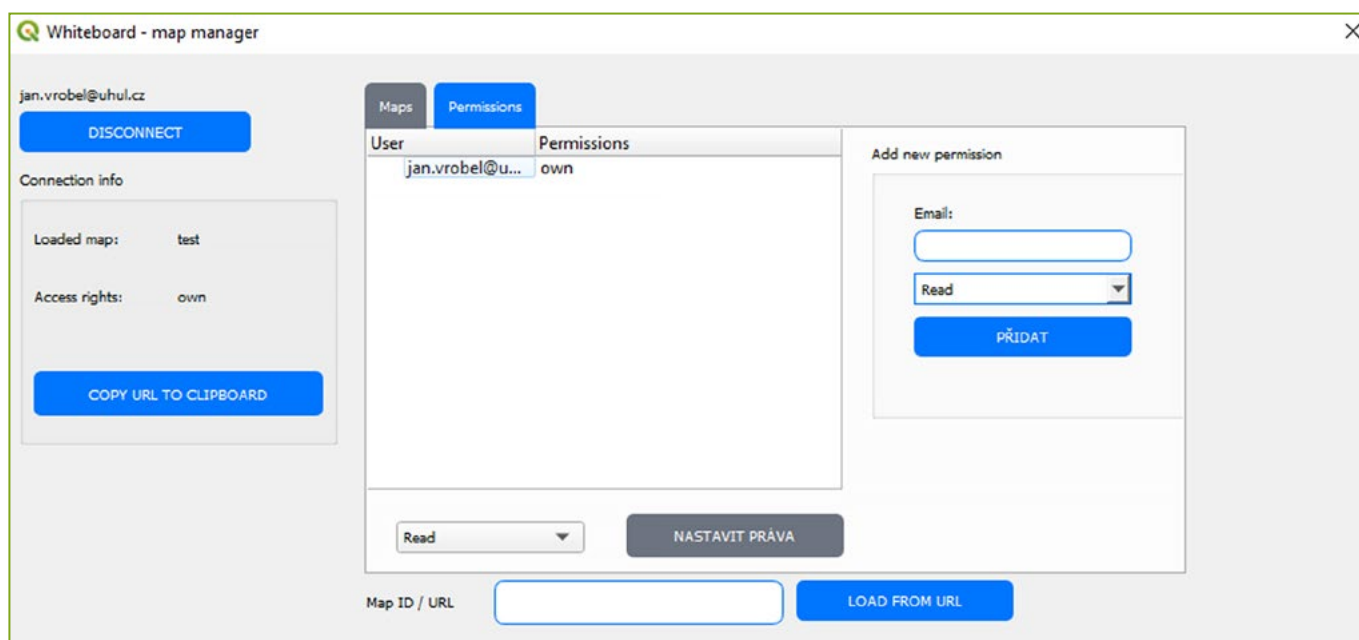
Každý klient musí mít možnost samostatného posouvání a pohybu na mapě. Všichni ostatní uživatelé však musí vidět kurzory ostatních připojených uživatelů, aby se napodobilo "ukazování", jak by to dělali lidé stojící nad stolem v zasedací místnosti.

Klienti musí být schopni kreslit na mapu a rozdělit to, co kreslí, do samostatných vrstev s odlišnými datovými modely. Uživatelé musí mít možnost vzájemně upravovat kreslené prvky a upravovat atributy spojené s jednotlivými prvky. Do mapy musí být možné přidávat informace, například přidat datovou vrstvu, kterou lze použít jako podkladové mapy, nebo doplňující informace k projednávanému tématu. A konečně musí být možné vyjmout data "z" MapWhiteboard a převést je do běžných pracovních postupů GIS, jako jsou profesionální úpravy, vizualizace a analýzy.

Klienti, kteří využívají Map Whiteboard, musí implementovat sadu instrukcí, které se předávají mezi jejich mapovými komponentami a serverem MapWhiteboard. Existuje velmi velké množství možných integrací, které lze provést, ale abychom mohli začít s technologií, která má



Obrázek 40: Webový klient (převzato z [122]).



Obrázek 41: Správa projektu v QGIS (převzato z [122]).

na trhu značné rozšíření, zaměřuje se MapWhiteboard na začátku na integraci s jakýmkoli mapovým klientem založeným na OpenLayers.

Technologie Map Whiteboard má za cíl zvýšit hodnotu profesionálních pracovních postupů. Proto je klíčovou součástí počáteční integrace prokázání, že koncept funguje napříč platformami. To znamená nejen mezi webovými klienty, ale také mezi profesionálním desktopovým klientem GIS a odlehčeným webovým klientem. (Obrázek 40)

Stejně jako v případě webových aplikací existuje velké množství stávajících aplikací GIS, které jsou široce využívány ve státní správě i v soukromém sektoru. Pro ověření koncepce, aby se projekt prosadil mezi potenciálními uživateli, se rozhodl zaměřit na QGIS (Obrázek 41).

14. ANALÝZY, MODELOVÁNÍ A UMĚLÁ INTELIGENCE

Cílem analýzy dat je zkoumat velké množství dat a vyvozovat z nich závěry. Lze použít několik technik, z nichž každá využívá podobné metody, ale má trochu jiné zaměření. Mezi tyto metody patří např. statistika, dolování dat a strojové učení. (Obrázek 42).

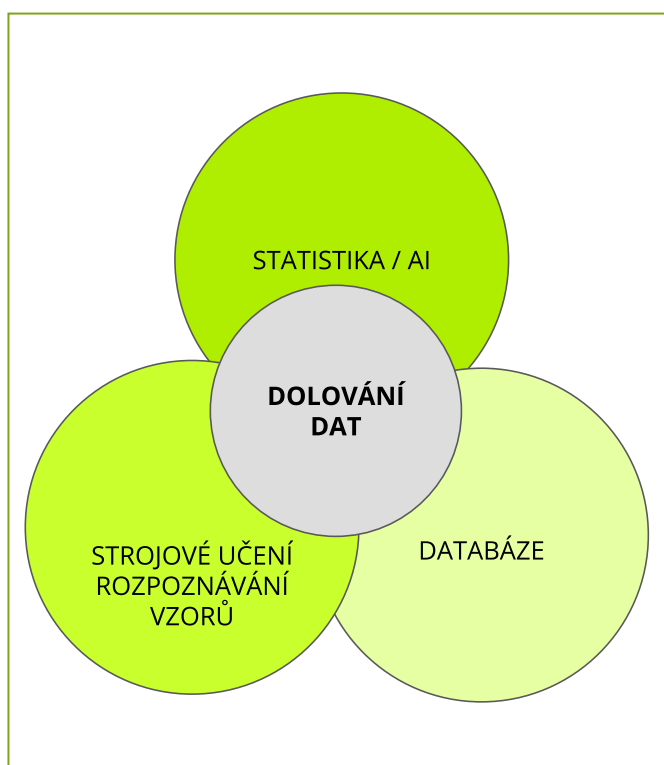
14.1 DATA MINING

Data mining je definován jako "věda o získávání užitečných informací z velkých souborů dat nebo databází" [123]. Strojové učení je "programování počítačů k optimalizaci výkonnostního kritéria na základě příkladových dat nebo minulých zkušeností" [124]. Někdy se dělení mezi strojovým učením a dolováním dat provádí na základě souborů

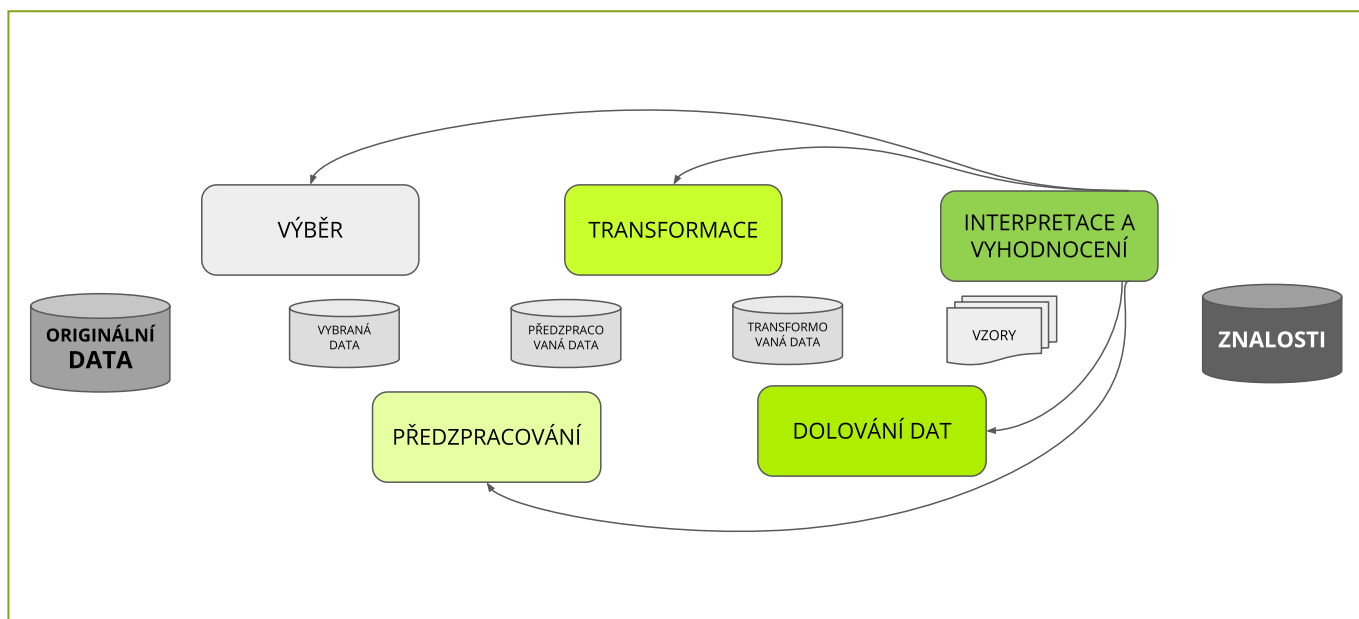
dat. Data mining je zaměřen na analýzu velkých databází, zatímco ve strojovém učení se klade důraz na učení vzorů z dat. Kořeny analýzy dat jsou ve statistice. Rozvoj počítačů a schopnost uchovávat a spravovat velké objemy dat umožnil rozsáhlé statistické výpočty a odstartoval vývoj nových metod, jejichž ruční provádění by bylo zdlouhavé.

Nejnovější oblastí analýzy dat je vizuální dolování dat. Vizualizace informací, dolování dat a interakce s uživatelem se v minulosti vyvíjely jako samostatné oblasti, ale od přelomu tisíciletí se stále více integrují jako vizuální dolování dat. Myšlenka vizuálního dolování dat se poprvé objevila v roce 1999, kdy Wong [125] tvrdil, že místo používání vizuálního zkoumání dat a analytických algoritmů dolování jako samostatných nástrojů by silnější strategií dolování dat bylo spojení vizualizací a analytických procesů do jednoho nástroje dolování dat. Podobná technologie je zmíněna v kapitole 13.1.3. Vysvětlující vizualizace (Explanatory visualisation). Mnoho technik dolování dat zahrnuje matematické kroky, které vyžadují zásah uživatele, a vizualizace by mohla tyto procesy podpořit. Vizuální dolování dat není jen o používání vizualizace k využívání dat, ale je to analytický proces dolování, v němž vizualizace hrají hlavní roli [126].

Analýza dat je iterativní proces, který začíná výběrem cílových dat ze surového materiálu a jejich předběžným zpracováním a transformací do vhodné podoby znalostí (Obrázek 43). Analýza dat využívá několik typů dat: databázové záznamy, maticová data, dokumenty, grafy, odkazy, transakční data, sekvence transakcí, data sekvencí DNA, celogenomové informace, časoprostorová data. Kvalita dat může často způsobovat problémy. Data mohou obsahovat šum, mohou se v nich vyskytovat chybějící hodnoty a duplicitní údaje, a proto je před jejich použitím nutná fáze čištění dat. Mohou být nutné i další druhy předběžného zpracování, jako je agregace dat, vzorkování, redukce dimenzionality, výběr podmnožiny, vytvoření příznaků a transformace atributů [123]. Dále se data proženou algoritmem dolování dat, který z nich vytvoří vzory. Uživatel interpretuje a vyhodnotí výsledky a spustí novou iteraci s případnými úpravami výchozích dat, algoritmu a jeho parametrů.



Obrázek 42: Techniky analýzy dat (upraveno dle [123]).



Obrázek 43: Proces dolování dat – přeměna dat ve znalost (upraveno dle [123]).

14.2. STATISTICKÉ METODY

Statistické metody se používají pro zkoumání dat, aby bylo možné lépe porozumět jejich vlastnostem [123]. Mezi ústřední metody patří např. souhrnné statistiky, korelace a vizualizace. Souhrnné statistiky jsou čísla, která shrnují vlastnosti dat. Amar a další [127] klasifikovali statistické metody jako:

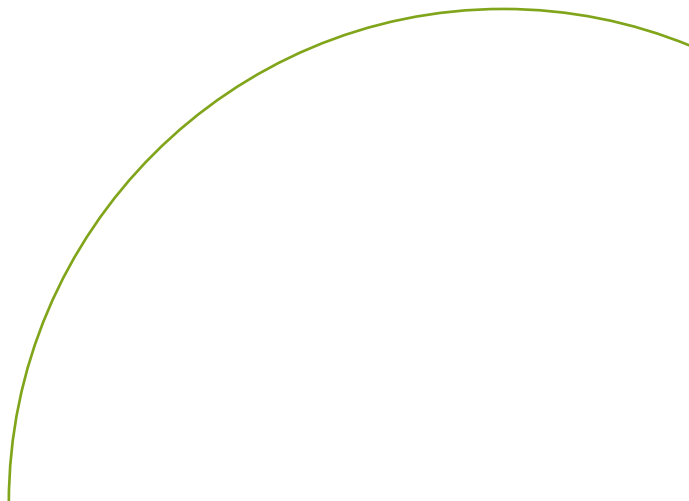
1. Hodnoty získané počítačem; průměr, medián, počet, složitější hodnoty,
2. Hledání extrémů; hledání případů dat, které mají nejvyšší a nejnižší hodnotu definovaného atributu,
3. Určení rozsahu: nalezení rozpětí hodnot atributu datových případů,
4. Charakterizace rozdělení: vytvoření rozdělení souboru datových případů s kvantitativním atributem, např. pro pochopení "normality". Vizualní metody využívají schopnost člověka rozpoznávat vzory. Jednotlivé proměnné se vyjadřují vizuální formou, např. jako histogramy a spojnicové grafy.

14.3. UMĚLÁ INTELIGENCE

Umělou inteligenci (AI) lze definovat jako "schopnost systému správně interpretovat vnější data, učit se z těchto dat a využívat tyto poznatky k dosažení konkrétních cílů a úkolů prostřednictvím flexibilního přizpůsobení" [128].

Strojové učení používají výzkumníci již od 50. let 20. století k analýze a získávání informací z dat. Teprve v posledním desetiletí s nástupem všeobecného využití grafických procesorů (GPU) umožnilo skutečný rozvoj neuronových sítí a zejména toho, co se dnes označuje jako hluboké učení (Deep Learning) [128]. Tento nově objevený výpočetní výkon dal vzniknout metodám, které jsou schopny řešit složité problémy reálného světa. Kapacita moderních počítačů nejenže umožňuje používat výpočetně náročné metody, ale také usnadňuje analýzu obrovského množství dat, tzv. velkých dat, v rozsahu, který byl dříve neřešitelný. Na rozdíl od předchozích metod využívá hluboké učení (Deep Learning) více vrstev neuronových sítí k vytvoření struktur schopných provádět konkrétní úkol, jako je klasifikace, segmentace, detekce, predikce a generování dat.

Hluboké učení je schopno objevovat korelace v datech bez potřeby ručně vytvořených rysů. Nedostatek heuristiky spolu s množstvím výpočetních zdrojů činí metody hlubokého učení ideálně vhodnými pro zpracování problémů velkých dat. Strojové učení dále nabízí možnost celoživotního učení, kdy se systém dokáže přizpůsobit měnícím se podmínkám. Zde je nutné konstatovat, že v souladu s tím, co jsme již napsali v kapitole 3.2. Kdo skutečně potřebuje data, ačkoli je strojové učení často prezentováno jako náhrada lidské inteligence, jedná se pouze o nástroj pro digitalizaci lidských odborných znalostí do počítačového modelu. Tento model je tak dobrý, jak dobré jsou informace, které mu lidé poskytnou.

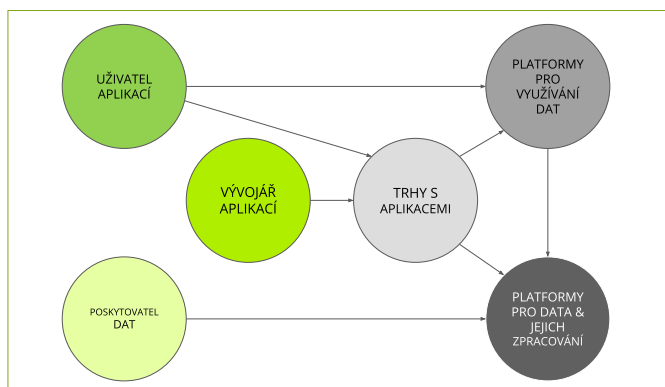


15. ŽIJEME V DOBĚ CLOUDU

V posledních letech roste počet snadno přístupných úložišť velkých dat umístěných v cloudových infrastrukturách. Mezi ty známější patří např. úložiště družicových dat z dálkového průzkumu Země, která jsou přístupná veřejnosti. Tato úložiště se v současné době transformují z čistě datových přístupových platforem na platformy, které nabízejí další sady cloudových produktů/služeb, jako jsou výpočetní, úložné nebo analytické služby. Zkušenosti ukazují, že kombinace dat a odpovídajících služeb je klíčovým faktorem umožňujícím efektivní zpracování velkých objemů dat.

Pokud již není přenos velkého množství dat proveditelný nebo nákladově efektivní, je třeba procesy (nebo aplikace) nasadit a provádět co nejlépe skutečným datům. [29]

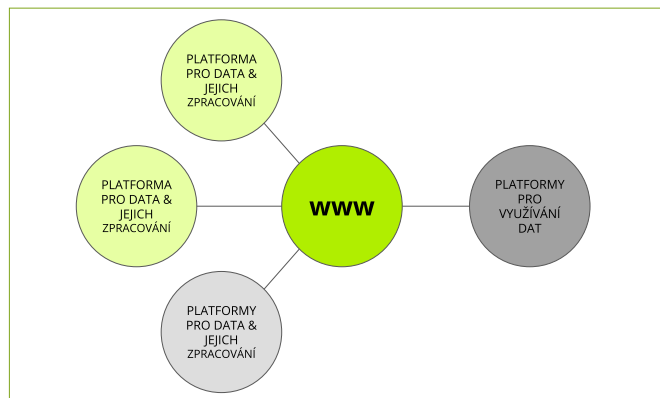
Poskytovatelé dat v levé části Obrázku 44 dole zpřístupňují svá data na veřejně přístupných datových a zpracovatelských platformách v cloudu. V ideálním případě tyto platformy poskytují přístup k větším souborům nezpracovaných dat a datových produktů od více poskytovatelů dat. Spotřebitelé aplikací (vlevo nahoře), tj. zákazníci se specifickými potřebami, které lze uspokojit zpracováním dat, identifikují vhodnou aplikaci (aplikace), která zpracováním (velkých) dat vytváří požadované výsledky. Aplikace vytvářejí vývojáři aplikací a nabízejí je na trzích s aplikacemi,



Obrázek 44: Vysokoúrovňová architektura cloudu (upraveno dle [129]).

které fungují dost podobně jako trhy s chytrými telefony, s tím rozdílem, že aplikace jsou nasazovány na vyžádání na cloudových platformách, nikoli stahovány a instalovány do chytrých telefonů. Platformy pro využívání výsledků (vpravo nahoře) podporují spotřebitele aplikací pomocí jednotného přihlášení, usnadňují řetězení aplikací i na více platformách pro zpracování dat a zajišťují co nejlépe uživatelskou zkušenost. [29]

V současné době probíhají standardizační snahy konsorcia OGC o Cloudové architektuře pro dálkový průzkum Země (Earth Observation Cloud Architecture), které splňují výše uvedené požadavky na vytvoření tržišť pro zpracování velkých objemů dat v cloudu pro jednotlivé domény a mezi nimi. Tato architektura podporuje paradigma velkých dat přibližující aplikace k datům, která jsou uložena a distribuována na nezávislých datových a zpracovatelských platformách. Základní myšlenkou je, že každá platforma poskytuje standardizované rozhraní, které umožňuje nasazení a parametrizované provádění aplikací, které jsou zabaleny jako softwarové kontejnery. Na datové a zpracovatelské platformy navazují platformy pro využívání dat, které umožňují řetězení kontejnerů/aplikací do pracovních postupů, tj. kdy a co se po zadání uživatele provede [29, 129]. Příklad takovýchto platform a jejich možná propojení je zobrazen na Obrázku 45.

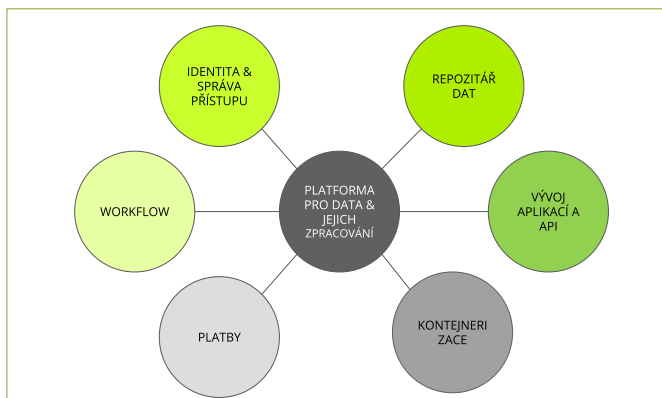


Obrázek 45: Platformy cloudové architektury pro dálkový průzkum Země (upraveno dle [29, 129]).

15.1. DATOVÉ A ZPRACOVATELSKÉ PLATFORMY

Platforma pro zpracování dat znázorněná na Obrázku 46 má šest hlavních součástí [129]:

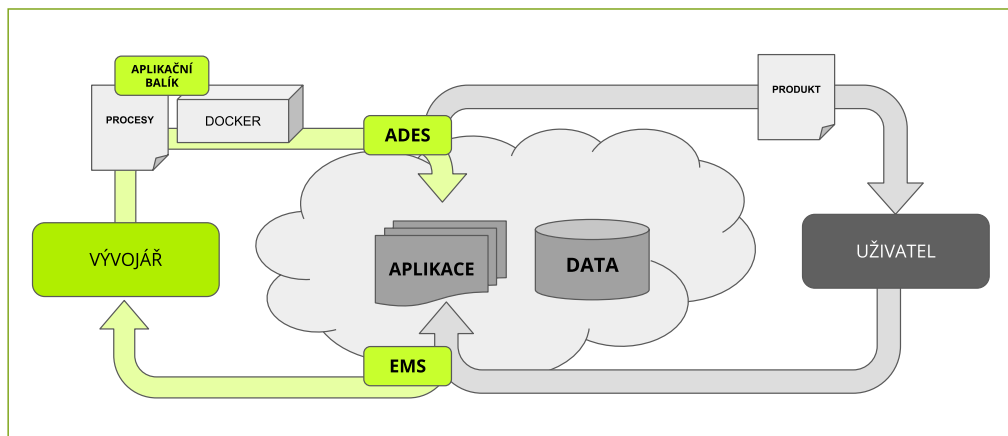
- Úložiště dat (Data repository).
- Rozhraní API (Application Programming Interfaces) pro nasazení a spuštění aplikací nebo provádění dotazů. Všechny aplikace jsou zapouzdřeny jako tzv. Docker kontejnery, které umožňují snadné a bezpečné nasazení a spuštění aplikací v cizích prostředích.
- Služba nazvaná “Docker Daemon” poskytuje prostředí Docker pro instanci a spuštění jednotlivých kontejnerů typu Docker.
- Komponenta pro výpočet nákladů za spuštění aplikace (Billing and Quoting) umožňuje získávat cenové nabídky a konečné vyúčtování za použití cloudových aplikací a služeb. Tato komponenta je zásadní, protože cenu za běh aplikace není vždy snadné vypočítat. Některé aplikace mají jednoduchý cenový model, který závisí pouze na parametrech, jako je oblast zájmu nebo časové období. Jiné aplikace, nebo ještě více složitě celé pracovní postupy s mnoha aplikacemi, mohou vyžadovat heuristiku pro výpočet úplné ceny za jejich provedení.
- Spouštěč pracovních postupů (Workflow Runner) může spouštět kontejnerové aplikace Docker. Spravuje dynamické načítání dat a perzistenci výsledků v nestálém kontejnerovém prostředí.
- Správa identit a přístupů (Identity & Access Management) poskytuje funkce pro správu uživatelů.



Obrázek 46: Struktura datové a zpracovatelské platformy (upraveno dle [29, 129]).

15.2. PLATFORMY PRO VYUŽÍVÁNÍ DAT

Platforma pro využívání výsledků je odpovědná za registraci a správu aplikací a za nasazení a provádění aplikací na platformách pro zpracování dat. Dále podporuje vytváření pracovních postupů na základě registrovaných aplikací a agreguje fakturační prvky, které jsou součástí těchto pracovních postupů. V ideálním případě vybírá platforma pro využívání výsledků nejvhodnější datovou a zpracovatelskou platformu na základě potřeb spotřebitele.



Obrázek 47: Kontext architektonických komponent cloudové architektury pro dálkový průzkum Země (upraveno dle [29]).

15.3. STANDARDY PRO EARTH OBSERVATION CLOUD ARCHITECTURE

Výše popsaná architektura je založená na třech klíčových součástech: Služba pro nasazení a provádění aplikací (Application Deployment and Execution Service – ADES), služba pro správu provádění aplikací (Execution Management Service – EMS) a balíček aplikací (Application Package – AP). V rámci aktivit OGC byla tato architektura upravena pro zpracování satelitních snímků. Následující schéma na Obrázku 47 znázorňuje vysokoúrovňový pohled na dvě oddělené smyčky vývoje aplikace (vlevo) a využití.

- cloudy výrobců zemědělských strojů (každý velký výrobce má vlastní cloud),
- cloudy lokalizačních služeb včetně RTK,
- cloudy výrobců chemikálií,
- cloudy klimatických služeb,
- cloudy poskytovatelů senzorů,
- a další.

Takže situace se zatím spíše komplikuje. Je otázka, jaká z tohoto vede cesta. V podstatě se dají vydefinovat dvě potenciální varianty. Většina velkých firem, včetně provozovatelů výše uvedených cloudů si žije svůj velký sen, že jednoho dne se oni stanou tím hlavním integrátorem, který propojí všechny služby a budou je nabízet celosvětově (Obrázek 50).

15.4. ŘEŠÍ DNEŠNÍ CLOUD SITUACI “CHUDÁKA ZEMĚDĚLCE”?

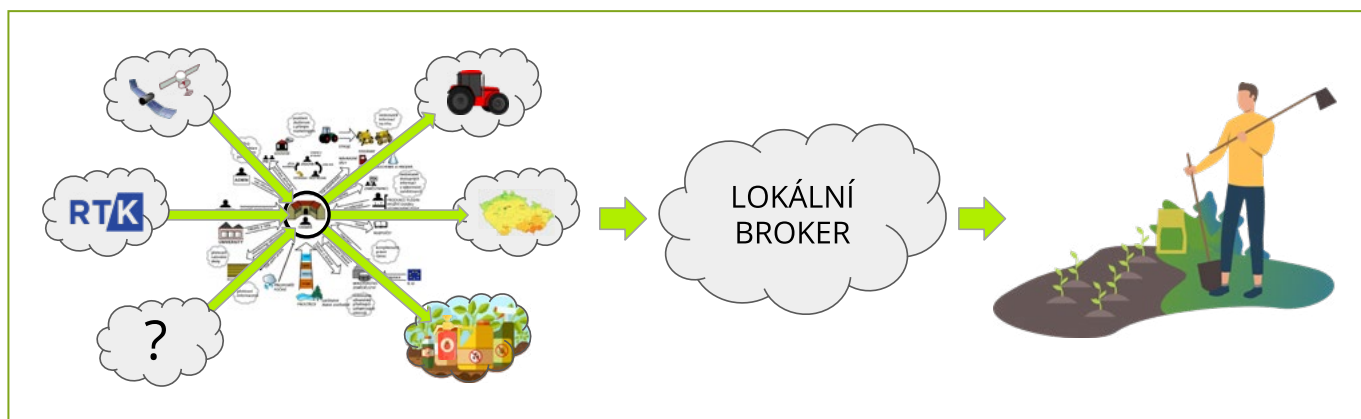
Bohužel na tuto otázku v dnešní době existuje jednoznačná odpověď: NERĚŠÍ. Naopak situace se mohla ještě zkomplikovat. Pokud se podíváme na schéma z kapitoly “2.4 Chudák farmář”, můžeme jej rozšířit následovně (Obrázek 49).

Do hry dnes vstupují:

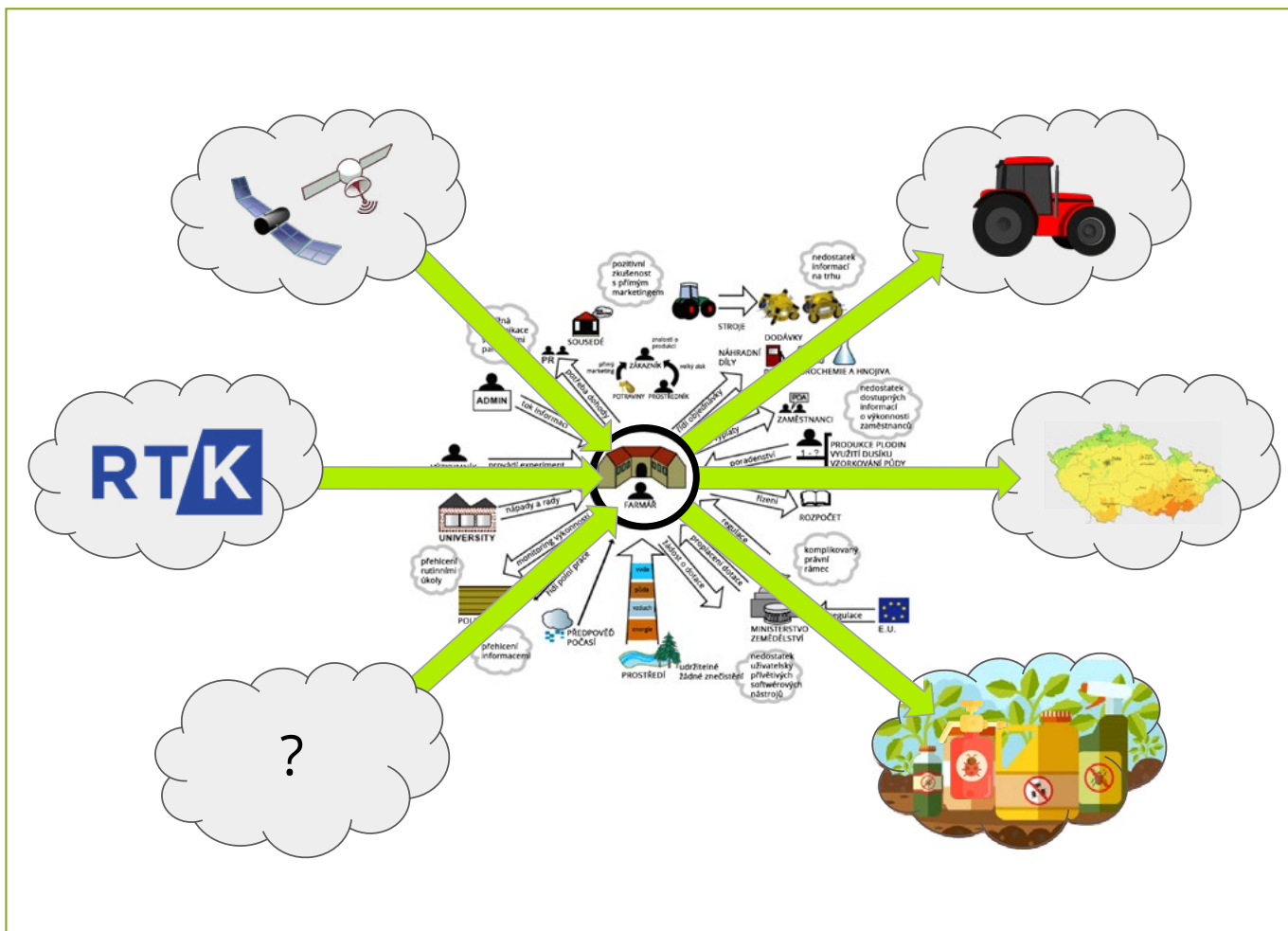
- provozovatelé platform pro zpracování satelitních snímků (viz kapitola 10. Dálkový průzkum Země),

Druhou variantou je (a v tuto variantu věří i autoři této publikace) vznik množství lokálních, regionálních nebo národních integrátorů - data brokerů, kteří budou integrovat a zprostředkovávat služby pro komunity farmářů ve svém okolí, viz. Obrázek 48. To je ostatně v souladu s myšlenkami o řetězcích přidaných v kapitole 3.3. Zpět ke znalostní pyramidě.

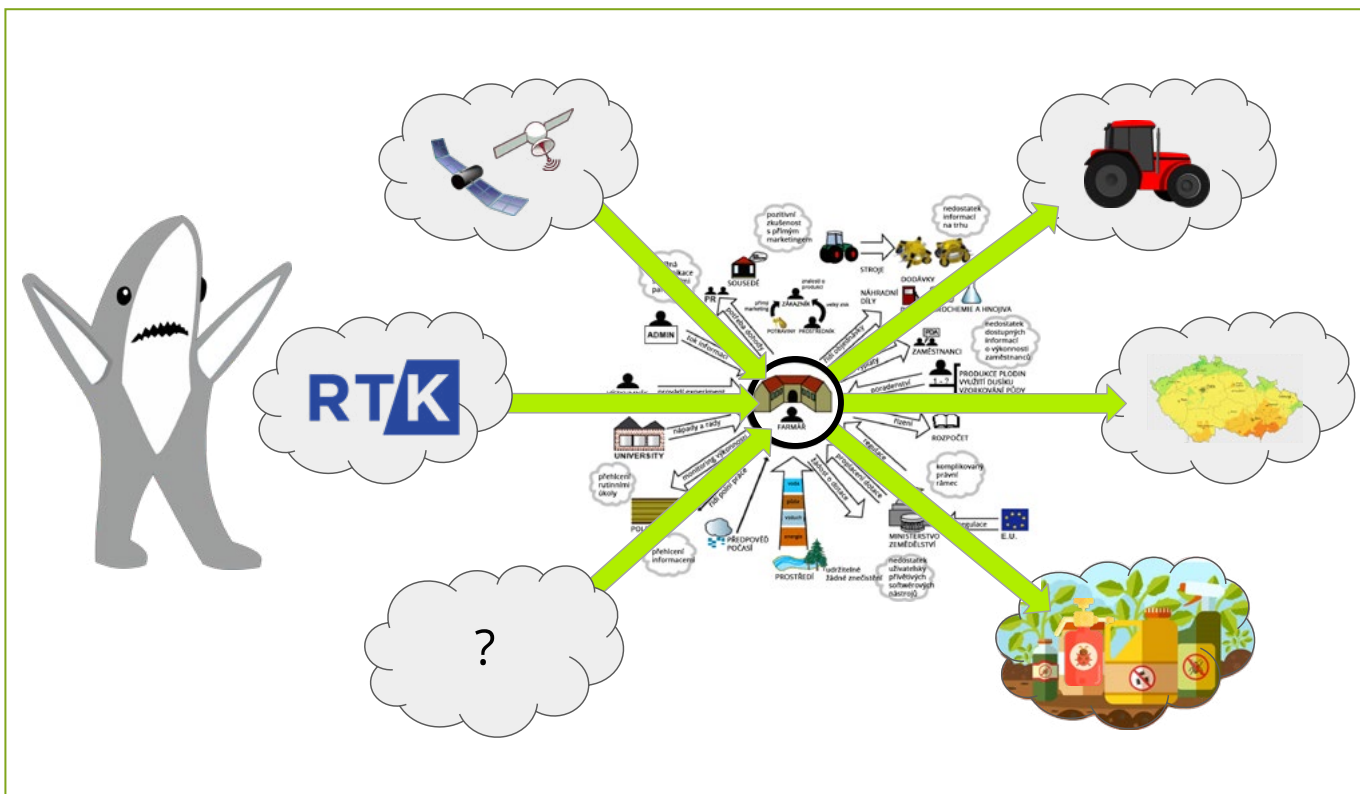
Kapitola věnovaná projektu Demeter, ukazuje koncept takového řešení.



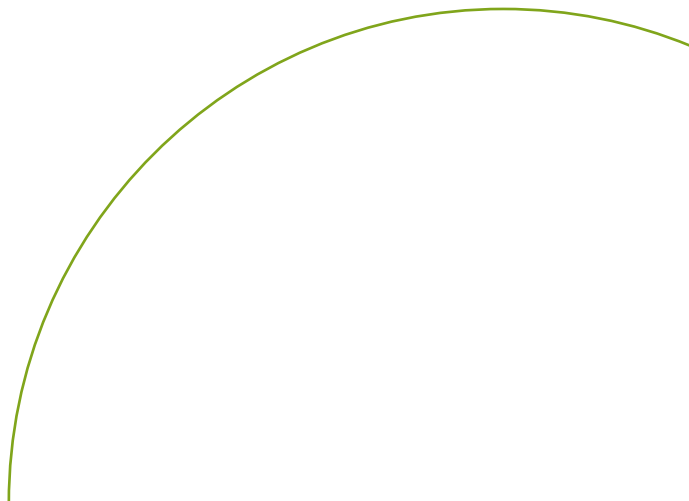
Obrázek 48: Koncept data brokerů.



Obrázek 49. Chudák farmář v době cloudu.



Obrázek 50. Sen velkých hráčů o ovládnutí informací v zemědělském sektoru.



16. PŘÍKLADY ŘEŠENÍ Z PROJEKTŮ HORIZONT 2020

Evropská komise podporuje svými grantovými programy mj. i výzkum v oblasti zemědělství. Největší objem finančních prostředků je uvolňován prostřednictvím programu Horizont 2020, v jehož rámci bylo podpořeno několik významných zemědělských projektů využívajících velká data, na kterých se náš tým podílel.

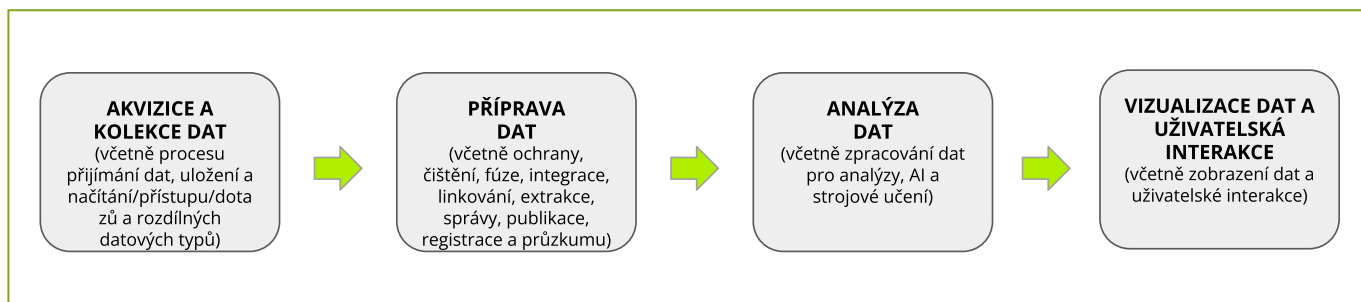
16.1. DATABIO A NÁSTROJE VELKÝCH DAT

Projekt DataBio integroval nejmodernější technologie pro zpracování velkých objemů dat s využitím infrastruktury a řešení stávajících partnerů projektu z oblasti zemědělství, lesnictví a rybníkářství. Tato řešení umožňují velká data inteligentně zpracovávat, analyzovat a vizualizovat. Prostředí DataBio umožňuje všem těmto třem sektorům selektivně využívat různé softwarové komponenty a datové sady podle jejich požadavků. Realizace proběhla díky kontinuální spolupráci koncových uživatelů a společností poskytujících technologie, výzkumných organizací, univerzit a zúčastněných stran z programu EU Big Data Value Public Private Partnership.

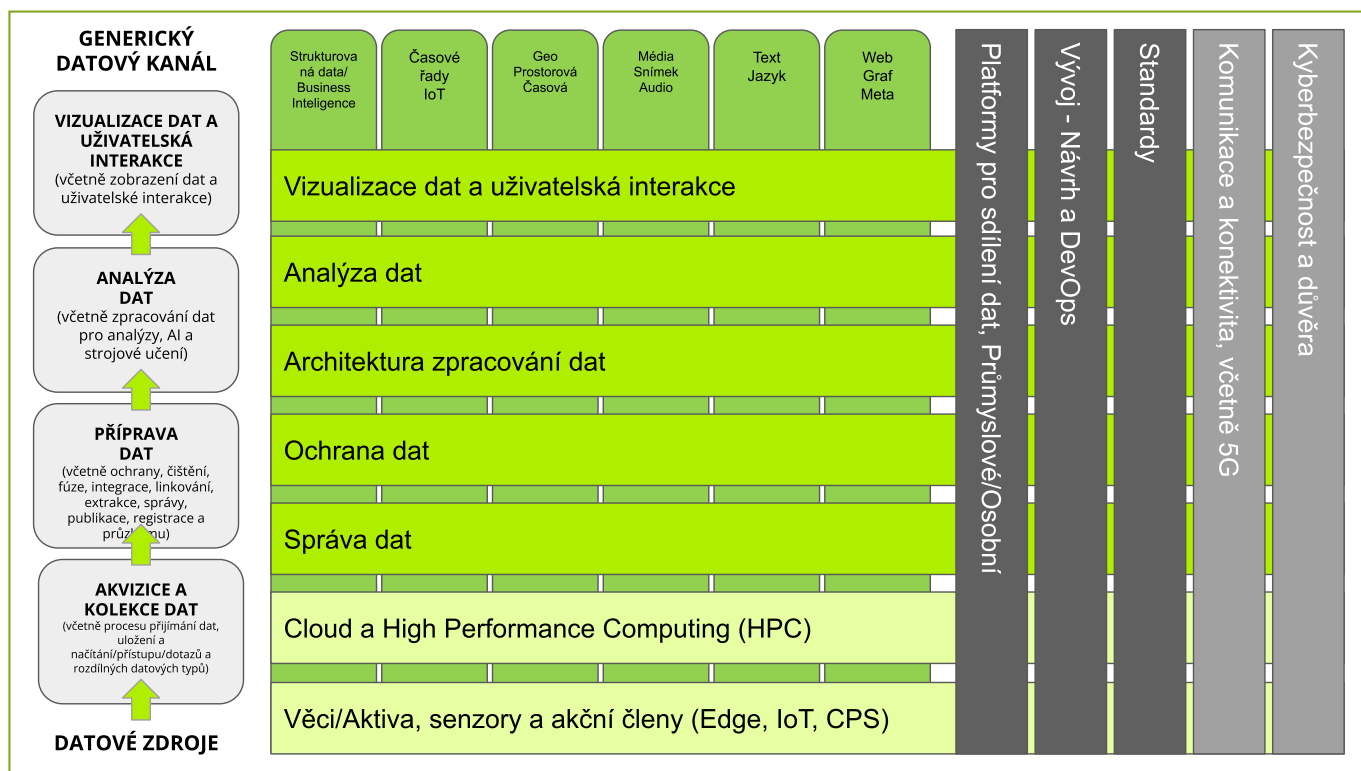
16.1.1. Zemědělské piloty

V oblasti zemědělství se DataBio projekt zaměřil na následující oblasti [130]:

- Precizní zahrady včetně vinné révy a oliv
 - Přesné zemědělství v oblasti oliv, ovoce, vinné révy a zeleniny se zaměřovalo na dálkovou diagnostiku a hodnocení chorob rostlin na základě zpracování satelitních snímků; systém upozornění na stav počasí, který povede k rozhodování o konkrétních opatřeních (např. ochrana plodin); systém o přesném zavlažování, precizní zemědělství podpora účinného hnojení půdy a postřiků v souladu se specifickými potřebami zemědělského podniku a ochranou životního prostředí; Pilot se zaměřoval na kombinované využití údajů o půdě, údajů o počasí, mapových podkladů, družic (LR, HR, VHR, SAR), zemědělských deníků, UAV, údajů o profilu farmy a údajů shromážděných mobilními audiovizuálními zařízeními,
- zároveň byly zemědělcům poskytovány poradenské služby týkající se diverzifikace plodin, které je nasměrují k produktivnějšímu a odolnějšímu pěstování.
- Správa velkých dat ve skleníkových ekosystémech - cílem navrhovaného pilotního projektu bylo poskytnout znalosti, know-how a nástroje týkající se toku informací, správy a analýzy dat ve skleníkovém zahradnictví. Za tímto účelem byla kombinovaná genomická a metabolická data. Během tohoto projektu bylo využito již vytvořených genomických dat, která byla integrována s novými, aby bylo možné posoudit genetický potenciál nových odrůd rajčat a jejich výkonnost ve sklenících.
- Přesné zemědělství na orné půdě
 - Obiloviny a plodiny - cílem tohoto pilotního projektu bylo poskytovat informace pro precizní zemědělství, především na základě časových řad družicových snímků s vysokým rozlišením (typu Sentinel-2), doplněných snímků z bezpilotních letounů, a polními (senzorovými) daty. Tyto informace byly použity jako vstupní údaje pro řízení zemědělských podniků (operativní rozhodnutí, taktická rozhodnutí).
 - Správa strojů - cílem byl monitorovací systém, který zahrnuje sledování polohy vozidel pomocí GPS v kombinaci se sběrem informací z palubního terminálu (CAN-BUS) a jejich online nebo off-line přenos do prostředí GIS. Monitorovací systém byl testován ve velkých, středních a malých zemědělských podnicích na základě úrovně zpracování informací a jejich interakce s dalšími údaji zemědělského podniku.
- Podpora Společné zemědělské politiky a finančního sektoru
 - Pojištění - Cílem tohoto pilotního projektu bylo poskytování a vyhodnocení služeb pro trh zemědělského pojištění na základě využití řady družicových dat programu Copernicus. Tato data byla integrována s meteorologickými údaji a dalšími dostupnými pozemními daty.



Obrázek 51: Obecná linka řetězení operací v projektu DataBio (upraveno dle [20]).



Obrázek 52: BDV referenční architektura (upraveno dle [20]).

- Podpora zemědělské politiky [CAP] - Cílem pilotního projektu bylo poskytování produktů a služeb založených na specializovaných vysoce automatizovaných procesorech zpracovávajících velké objemy dat se zaměřením na data Copernicus Sentinel 2.

16.1.2. DataBio architektura

Základní metodou při zpracování toků dat v projektu DataBio, bylo řetězení operací, tzv. pipelines, neboli technologická linka. Termín linka byl v projektu DataBio použit pro popis kroků zpracování dat. Každý krok má svá vstupní a výstupní data. Tato linka se vytváří řetězením jednotlivých článků za sebou, kdy výstup z předchozího kroku zpracování je přiváděn do kroku následujícího. Typicky v aplikacích pro zpracování velkých objemů dat zahrnují kroky této linky sběr dat, jejich zpracování, analýzu a vizualizaci výsledků. V projektu DataBio tyto kroky nazýváme obecnou linkou (Obrázek 51). Tato generická linka byla přizpůsobena pro oblast zemědělství, lesnictví a rybolovu.

Pro podrobnější popis řetězení operací nad velkými daty byl v projektu DataBio přijat referenční model Big Data Value (BDV) (Obrázek 52). Referenční model BDV byl vyvinut průmyslovým sdružením Big Data Value Association (BDVA) s přihlédnutím k příspěvkům technických odborníků a zúčastněných stran z různých oborů a také v interakci s dalšími průmyslovými sdruženími a s EU.

Dle [20] lze popsat následující úrovně BDV referenční architektury (viz Obrázek 52):

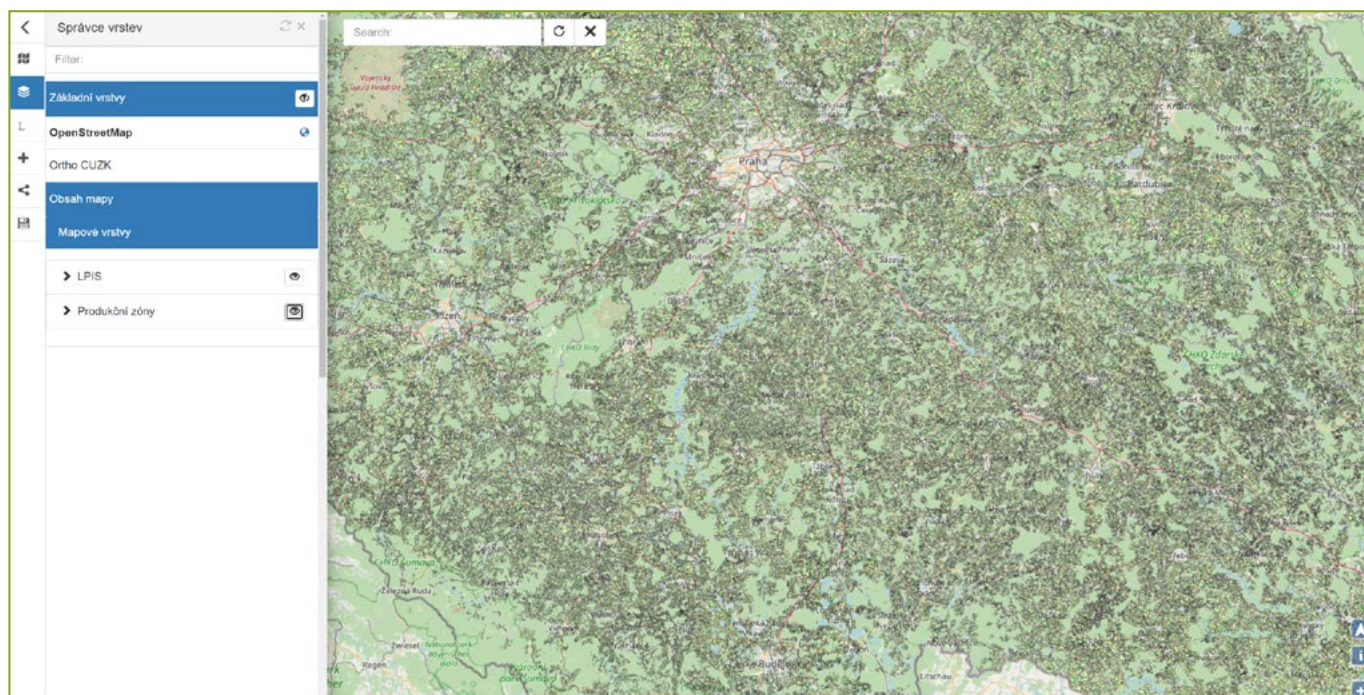
- **Sběr dat:** tato vrstva zajišťuje rozhraní pro interakci s poskytovateli dat a zahrnuje přenos dat z různých zdrojů tak, aby bylo možné data zpřístupnit, využít a analyzovat. Hlavním zdrojem velkých dat jsou data ze senzorů v kontextu internetu věcí v kyberneticko-fyzikálních systémech. Úlohy v této vrstvě v závislosti na typu shromážděných dat a na implementaci aplikace zahrnují přijímání nebo provádění specifických sběrů dat, stahování dat nebo přijímání dat od poskytovatelů dat a ukládání nebo ukládání dat do vyrovnávací paměti. Během procesu mohou být také vytvořena metadata,

kteřá usnadňují následné agregace nebo metody vyhledávání standardizovanými metodami. Do tohoto kroku lze zahrnout také zohlednění bezpečnosti a ochrany soukromí, protože během sběru dat se obvykle provádějí činnosti ověřování a autorizace a také činnosti zaznamenávání a udržování informací o původu dat.

- **Cloud**, vysoce výkonná výpočetní technika (HPC) a správa dat: Efektivní zpracování velkých objemů dat a správa dat může znamenat efektivní využití cloudových a HPC platforem. Tradiční relační databáze (RDB) obvykle nejsou dobře škálovatelné. Nejsou také nijak zvláště vhodné pro zpracování nestrukturovaných dat, jako jsou obrázky a video. Proto se doplňují nerelačními databázemi, jako jsou databáze s ukládáním klíčů, sloupcové, dokumentové a grafové databáze [131]. Z nich se sloupcově orientované architektury používají např. v softwaru Apache Cassandra a Hbase pro ukládání velkého množství dat. Dokumentové databáze zaznamenaly v posledních letech obrovský rozvoj. Nejpoužívanější dokumentovou databází poslední doby je MongoDB, která byla použita i v projektu DataBio, např. v DataBio Hub pro správu aktiv projektu a v databázové komponentě GeoRocket.
- **Příprava dat:** Úkoly prováděné v tomto kroku zahrnují validaci dat, například kontrolu formátů, čištění dat, například odstranění odlehlých hodnot, extrakci užitečných informací, organizaci a integraci dat shromážděných z různých zdrojů. Kromě toho se úkoly skládají z využití metadatových klíčů k vytvoření rozšířené a vylepšené datové sady, anotace, publikace a prezentace dat, aby byla k dispozici pro nalezení, opětovné použití a uložení, standardizace a přeformátování, jakož i zapouzdření.
- **Zpracování a ochrana dat:** Klíčem ke zpracování velkých objemů dat pomocí algoritmů, které jsou v některých případech složité tak, aby bylo dosaženo vysoké propustnosti, je uspořádání výpočtů tak, aby probíhaly paralelně. Hardware pro paralelní výpočty se skládá z 10, 100 nebo několika tisíc procesorů, často sdružených do karet GPU (Graphical Processing Unit) nebo TPU (Tensor Processing Unit). GPU a TPU se používají zejména v oblasti strojového učení a vizualizace. Paralelizace je přímočará při zpracování obrazu a videa, kde se stejné operace obvykle aplikují na různé části

obrazu. Paralelní výpočty na GPU's se používají v DataBio, např. pro vizualizaci dat. Ochrana dat zahrnuje mechanismy ochrany pro zajištění soukromí a anonymizace. Tato oblast se nachází mezi správou a zpracováním dat, ale může být spojena i s oblastí kybernetické bezpečnosti.

- **Analýza dat:** Tato vrstva slouží k objevování nových vzorců a vztahů, které poskytují nové poznatky. Získávání znalostí z dat vychází z požadavků aplikace, která specifikuje algoritmy pro zpracování dat. Analýza dat je zásadním krokem, protože jejím cílem je poskytování doporučení, či přímo provádění rozhodování. K metodám potřebným pro analýzu velkých objemů dat patří např. hašování, indexování a paralelní výpočty. V mnoha případech se používají také techniky strojového učení a další metody umělé inteligence. Analýza využívá data z minulosti i ze současnosti. Cílem je často pochopit mechanismus, který utvářel minulost a/nebo předpovědět budoucnost. Data z minulosti se používají pro popisnou a diagnostickou analytiku a klasické dotazování a reporting. Patří sem údaje o výkonu, transakční údaje, údaje o postojích, údaje o chování, údaje související s polohou a interakční údaje. Data ze současnosti se využívají při monitorování a analýze v reálném čase. To vyžaduje rychlé zpracování a v mnoha případech zpracovávání dat v reálném čase, k detekci krizových událostí. Využití dat pro budoucnost zahrnuje predikci a doporučení. To obvykle vyžaduje zpracování velkých objemů dat, rozsáhlé modelování a také kombinování znalostí z minulosti a přítomnosti, aby bylo možné získat náhled do budoucnosti.
- **Vizualizace dat a interakce s uživatelem:** Vizualizace pomáhá při interpretaci dat vytvářením grafických reprezentací sdělovaných informací. Přidává tak datům větší hodnotu, protože lidský mozek lépe vstřebává informace, pokud jsou prezentovány v podobě grafů, a nikoli v tabulkách nebo přehledech. Uživatelé díky vizualizaci mohou porozumět rozsáhlým souborům složitých dat, pracovat s nimi a činit rozhodnutí. Efektivní vizualizace dat musí zachovávat rovnováhu mezi vizuálními prvky, které poskytuje, a způsobem, jakým je poskytuje, aby přitáhla pozornost uživatelů a předala jim správná sdělení.



Obrázek 53: Portál s mapou výnosového potenciálu pro celou ČR (upraveno dle [132]).

16.1.3. Česká pilotní aplikace

Cílem jedné z pilotních aplikací bylo vyvinout platformu pro mapování výživového stavu plodin za pomoci využití dat DPZ (Landsat, Sentinel-2) jako podpůrného nástroje pro variabilní dávkování hnojiv a ochranu plodin. [132]

To zahrnuje identifikaci stavu plodin, mapování prostorové variability a vymezení management zón. Pro experimenty potřebné pro vývoj a testování platformy byl zvolen jeden zemědělský podnik v ČR, nicméně základní datové sady jsou již připraveny pro využití v rámci celé ČR. Řešení, které je výsledkem projektu, je tedy využitelné řešení pro jakýkoliv zemědělský podnik v České republice. V projektu byla využita metodika pro odhadování variability výnosů v rámci půdních bloků. Tato metodika byla vyvinuta na Mendelově univerzitě [133]. V rámci pilotu byla vyvinuta aplikace a byl realizován experiment, kde byla odhadovaná variabilita výnosů porovnávána s daty z výnosoměrů umístěných na zemědělské technice. Byla využita osmiletá časová řada multispektrálních snímků Landsat a Sentinel, která slouží k výpočtu vegetačních indexů, na jejichž základě je odhadována variabilita výnosů v rámci pozemku. Dle této variability jsou následně vymezeny takzvané produkční zóny. [132]

Získaná mapa odhadované výnosové variability znázorňuje procentuální podíl výnosu v daném místě ku průměrné hodnotě výnosu daného pozemku.

Mapa je potom překlasifikována na mapu produkčních zón zahrnující tři kategorie – oblasti s vysokým, středním a nízkým očekávaným výnosem, případně pět kategorií, pokud pozemek vykazuje větší variabilitu s výraznými extrémy a je vhodné použít přesnější klasifikaci.

Mapa produkčních zón potom může sloužit například jako podklad pro variabilní aplikaci dusíku a v případě použití posilovací strategie variabilní aplikace, při které jsou v oblastech s vysokým očekávaným výnosem dávky dusíku navýšeny, a naopak v místech s nízkým očekávaným výnosem sníženy. Může sloužit i jiným účelům dle uvážení agronoma a managementu zemědělského podniku.

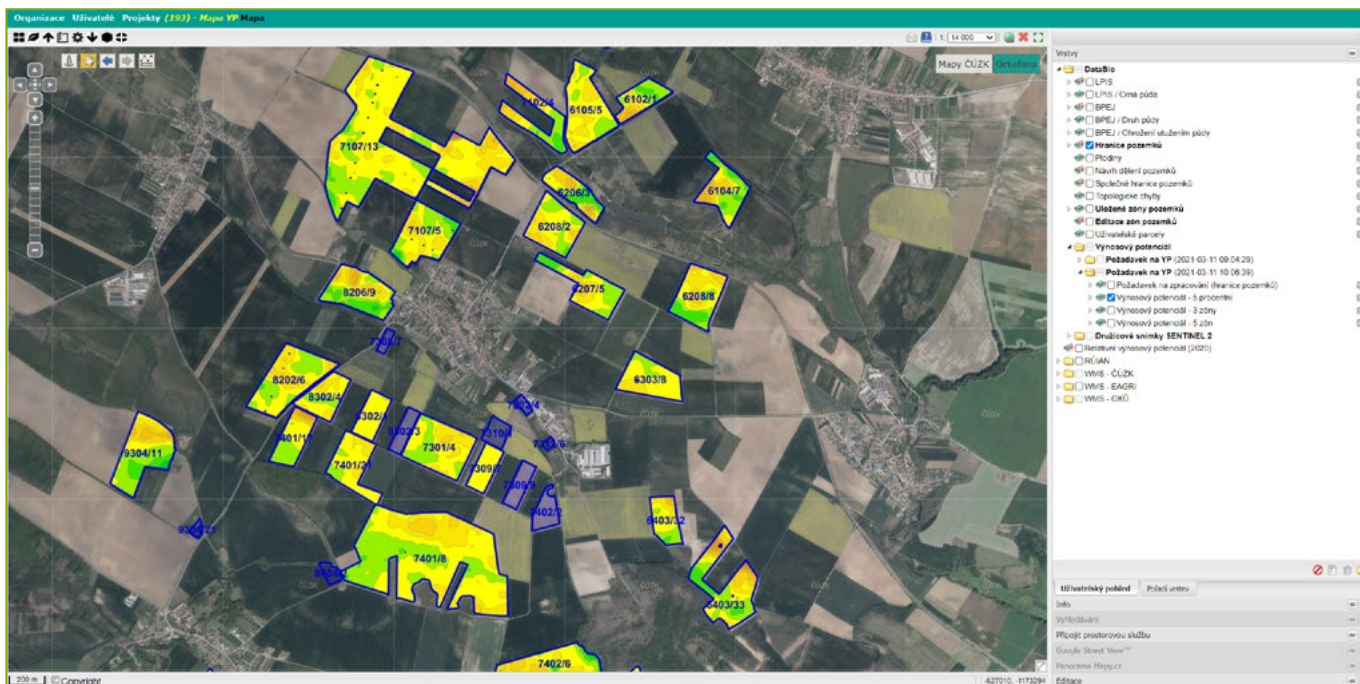
Provoz tohoto nástroje je dále podporován platformou pro automatické stahování a management multispektrálních dat z družice Sentinel 2 a dat atmosférické korekce.

Další část vývoje platformy byla zaměřena na mapování datového modelu LPIS do ontologie FOODIE, následný převod LPIS dat a na vývoj nástrojů pro získávání relevantních propojených dat (linked data) pomocí dotazů. Lesprojekt – služby, s.r.o. realizoval tyto aktivity společně s PSNC a v současné době systém podporuje otevřený přístup k datům LPIS prostřednictvím ontologie FOODIE a také zabezpečený přístup k datům farem. [132]

Zemědělci mohou tato základní data volně testovat (Obrázek 53) nebo si je i stahovat.

Řešení je nabízeno v podobě webového GIS portálu pro zemědělce, kde mohou uživatelé kromě zadání požadavků na vytvoření map produkčních zón tyto zóny a aplikační mapy vytvářet i svépomocí a k tomu využít hrubou nerevidovanou mapu produkčních zón (výnosového potenciálu) pokrývající ornou půdu v celé ČR a zpřístupněnou veřejnosti prostřednictvím WMS.

Portál nabízí také možnost zobrazení a připojení dalších vrstev, které mohou společně s hrubou revidovanou vrstvou produkčních zón a hranic pozemků sloužit jako další podklady při přípravě aplikačních map.



Obrázek 54: Webová aplikace pro přípravu doporučení (upraveno dle [132]).

16.2. EUXDAT

Mezi tyto další pomocné vrstvy, které jsou v portálu zahrnuty, patří například veřejný LPIS, který může uživatel využít pro import hranic pozemků pro své projekty, dále mapy družicových snímků v přirozených barvách, i vegetační indexy.

Jako podpůrné vrstvy jsou připojeny například WMS služby publikované ministerstvem zemědělství zahrnující erozní ohroženost pozemků a další, ale prostřednictvím formuláře a rozhraní pro standardizované webové služby může uživatel připojit i jakékoliv další dostupné vrstvy dle svého uvážení. [132] (Obrázek 54)

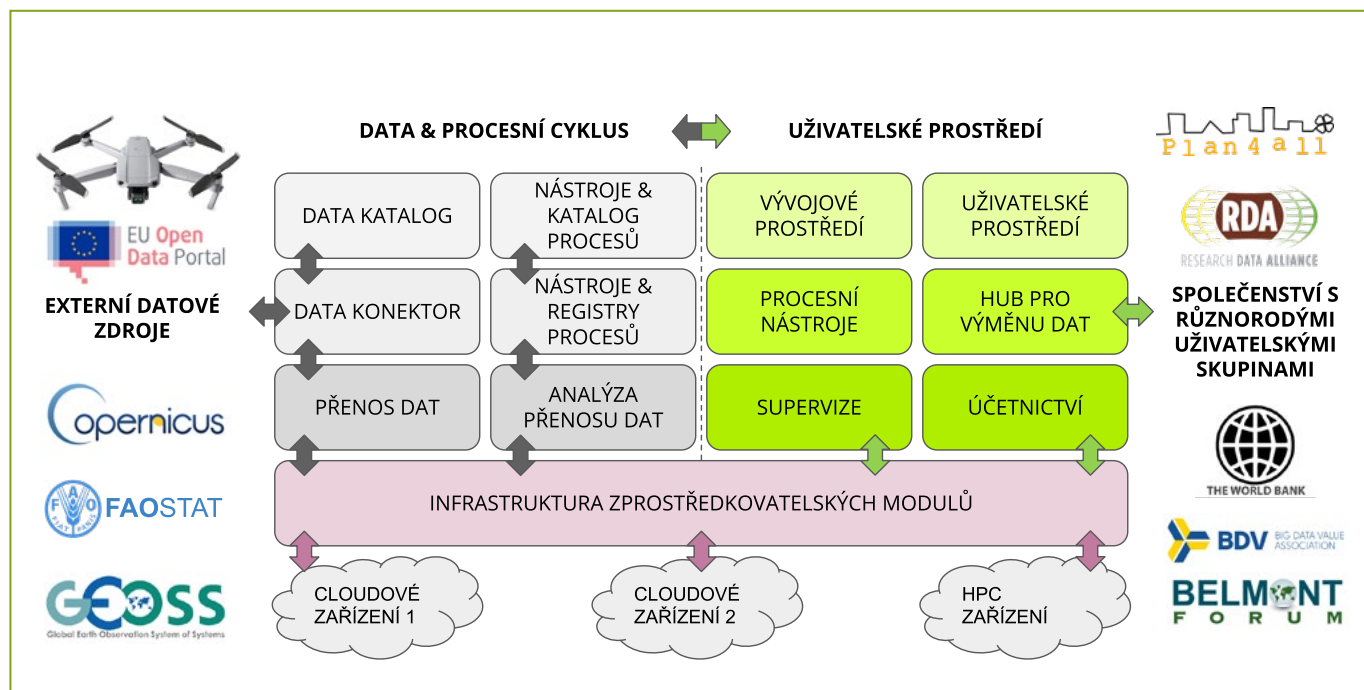
Mapy odhadované výnosové variability (výnosového potenciálu) pro sezónu 2017 a nyní i pro sezónu 2020 pro celou ČR a zdrojová data jsou nyní k dispozici na AgriHubu (viz kapitola 17. Digitální inovační Huby - budoucnost, jak transformovat velká data do znalostí).

Projekt European e-Infrastructure for Extreme Data Analytics in Sustainable Development (EUXDAT) navrhl cloudovou e-infrastrukturu, zaměřenou na zemědělství, monitorování půdy a energetickou účinnost pro udržitelný rozvoj a slouží také jako způsob podpory plánovacích politik pro různé činnosti v zemědělství.

Za tímto účelem bylo třeba řešit problémy související se současným a budoucím obrovským množstvím heterogenních dat, která je třeba spravovat a zpracovávat. EUXDAT stavěl na stávajících vyspělých řešeních a poskytuje pokročilý frontend, kde mohou uživatelé vyvíjet aplikace nad infrastrukturou založenou na High Performance Computing (HPC) a cloudu. Projekt EUXDAT využil spolupráce s vědeckými komunitami, aby identifikoval nové trendy a datové sady, což napomohlo vývoji e-infrastruktury. Konečným výsledkem projektu je integrovaná e-infrastruktura, která podporuje koncové uživatele k vytváření nových aplikací pro udržitelný rozvoj. [134]

Projekt EUXDAT nasadil nejmodernější platformu pro využívání velkých dat, dat ve formě hybridního HPC-Cloudu s využitím stávající infrastruktury projektových partnerů. Tato e-infrastruktura EUXDAT umožňuje různým uživatelům (vědci, zemědělci, plánovací a rozhodovací orgány) plně využívat cloudové výpočetní kapacity ke zkoumání nových přístupů, nových inovativních služeb a provádění předpovědí a simulací s extrémně velkými a heterogenními soubory dat. Mezi cíle projektu patřilo [135]:

- Cíl 1: Vyvinout sadu nástrojů pro správu extrémně velkých datových souborů s přihlédnutím k požadavkům na jejich ukládání, různým formátům a politikám řízení pro snížení latence pohybu dat a ochranu informací.



Obrázek 55: Celková architektura projektu EUXDAT (upraveno dle [135]).

- Cíl 2: Přizpůsobit a podle potřeby vyvíjet již dostupné nástroje pro zpracování dat a přidat do nich nové funkce tak, aby je bylo možné poskytovat formou analýzy velkých objemů dat ve formě služby. Tyto hlavní změny budou zaměřeny na možnosti využití možností HPC s novými nástroji pro správu dat, zlepšení portálu pro uživatele a přizpůsobení správy zdrojů.
- Cíl 3: Provádět servisní činnosti založené na integrované e-infrastruktúře, kde budou navržena řešení ověřena na třech datově náročnými pilotními projekty z oblasti udržitelného rozvoje.
- Cíl 4: Provádět důležitou síťovací činnost, zejména v oblasti udržitelného rozvoje, s cílem motivovat k přijetí navrhovaných nástrojů širší evropskou komunitu.

16.2.1. Výstupy projektu EUXDAT

Jako hlavní technologický výstup byla vytvořena e-infrastruktura pro analýzu velkých dat, která je poskytována jako služba a to s několika softwarovými vrstvami podporujícími udržitelné a produktivní zemědělství, ochranu půdy, ochranu vod, regionální biologickou rozmanitost a udržitelný rozvoj zelené infrastruktury vybudováním robustního propojení e-infrastruktury s heterogenními zdroji dat, kombinací odborných znalostí z různých oborů a výsledků předchozích úspěšných projektů v oblasti zemědělství, lesnictví, zelené infrastruktury, územního a environmentálního plánování, geomatiky a ekosystémů. Přehled této komplexní e-infrastruktury je zobrazen na Obrázku 55.

16.2.2. Piloty a scénáře

Specifikované piloty, jak je popsáno v návrhu projektu EUXDAT [136], jsou technické piloty ~ vize každého pilotu nejprve popisuje jeho technologickou povahu a poté ma-

puje pilot na níže identifikované scénáře mající obchodní potenciál. Kompletní popis pilotů a scénářů je uveden v [137]. Výstupy a přístupy k jednotlivým scénářům a jejich webovým rozhraním jsou dostupné skrze uživatelskou platformu projektu EUXDAT podrobně popsanou dále.

Projekt měl tři pilotní části:

- Pilot 1: Monitorování půdy a udržitelné hospodaření - Služba poskytne mapy skutečného napadení cílových olivových plantáží na úrovni stromů. Pomocí těchto map a nových ošetření budou moci zemědělci a agronomové efektivněji a účinněji zasahovat při odhalování a ošetřování škůdců a nemocí. Jejich identifikace bude dosaženo pomocí hyperspektrálních snímků pořízených pomocí UAV a jejich korelace s pozemní kontrolou.
- Pilot 2: Analýza energetické účinnosti - Pilotní projekt bude na jedné straně zaměřen na sběr a integraci dat z různých zdrojů dat (in-situ senzory, monitorování strojového parku, systémů řízení zemědělských podniků atd.). Dále bude zaměřen na poskytování analýz a vizualizací jejich výsledků jako podpůrného nástroje pro rozhodování zemědělců a agronomů na straně druhé.
- Pilot 3: 3D farmaření - Hlavní cíl tohoto pilotu spočívá v rozdělení polí do zón s různou produktivitou na základě vhodné vlhkosti, sklonu a orientace pozemku, obsahu hnojiv, dobré přístupnosti atd.

Pro výše uvedené piloty jsou definovány scénáře, které přispívají do náplně jednotlivých pilotů:

- **Vylepšení otevřené datové sady pro data o využití půdy (Open Land Use Map Improvement)**

Scénář byl sémanticky rozdělen na dvě části. Obohacení otevřené datové sady o využití půdy (OLU) o charakteristicky důležité pro zemědělce v Evropě. A také vývoj modelů,

kteří umožnily identifikaci polí v regionech, kde tato data nejsou k dispozici pomocí dat z dálkového průzkumu Země a pomocí nich vypočítat homogenitu půdního pokryvu polí. Více o otevřených datech o využití půdy viz kapitola 9.2. Otevřená data využití půdy – Open Land Use.

- **Monitorování stavu plodin (Monitoring of crop status)**

Monitorovací systém zahrnovalo komponentu pro detekci anomálií plodin založenou na specifických algoritmech analýzy obrazu Sentinel 2, která monitoruje stav plodin a přítomnost anomálií jako odchylky v prostorovém rozložení indexů, které souvisejí se zdravotním stavem, jako jsou CRI2 a NDVI. Pomocí hyperspektrálních dat a dalších zdrojů dat pak byly anomálie přiřazovány stresu, chorobám nebo hmyzu. Indexy Sentinel-2 vymezují možná místa výskytu chorob na rozsáhlých územích. Ohniska byla snímána hyperspektrální kamerou na bezpilotním letounu, aby bylo možné přesně identifikovat chorobu nebo stres, a ne pouze anomálie multispektrálních indexů S-2, které pracují s hrubším rozlišením než hyperspektrální indexy, které jsou namontovány na bezpilotním letounu. Hyperspektrální indexy mají schopnost přesně identifikovat typ nemoci nebo stresu.

- **Vymezení agroklimatických zón (Delimiting Agro-climatic zones)**

Systém agroklimatické klasifikace umožnil zařadit pozemky do agroklimatických tříd pro rostlinnou výrobu a hospodaření na základě dlouhodobých klimatických údajů (teplota, vlhkost, osvit atp.), půdního pokryvu a topografie.

- **Hledání změn klimatických vzorců (Looking for climatic patterns changes)**

Tento systém umožnil posoudit trendy, rozložení četností a extrémy hlavních proměnných souvisejících s počasím (teplota, srážky, evapotranspirace, půdní vlhkost) pro zemědělské oblasti. Rozdíly oproti stávajícím klimatickým databázím budou následující: 1. Analýza byla zaměřena na zemědělskou půdu (filtr). 2. Analýza byla zaměřena na období pěstování plodin (filtr použitelný pro různé plodiny). 3. Analýza umožnila uživatelům nahrát vlastní záznamy (např. měření teploty) pro porovnání

regionálních modelových dat. 4. Služba byla propojena s existujícími modely změny klimatu, které kvantifikují budoucí změny.

Tato služba vytvářela jedinečnou možnost posoudit potenciální dopad změny klimatu na místní produkci.

- **Systém pro podporu rozhodování při plánování zásahů v terénu (Information support for field use recommendations)**

Systém propojil soubory prostorových dat (aktuální počasí a předpověď, půdní data, typ půdního pokryvu, zóny hospodaření atd.) a poskytl zemědělcům doporučení pro skutečné využití a kontrolu dodržování předpisů v minulosti pro jednotlivá pole a plodiny. Tyto informace jsou důležité pro plánování několika činností: a) zásahy do hospodaření na poli, b) dostupnost pole těžkou technikou, jako je kombajn, rozmetadlo hnoje nebo rozmetadlo základních hnojiv, c) stav půdy (např. obsah NPK a vody). Aktuální počasí na poli je nutné pro rozhodnutí o tom, zda je možné a účinné v daném čase aplikovat pesticidní postřiky nebo hnojiva. Tento scénář byl zaměřen především na ochranu polí a plodin.

- **Efektivní využívání přírodních zdrojů (Effective utilization of natural resources)**

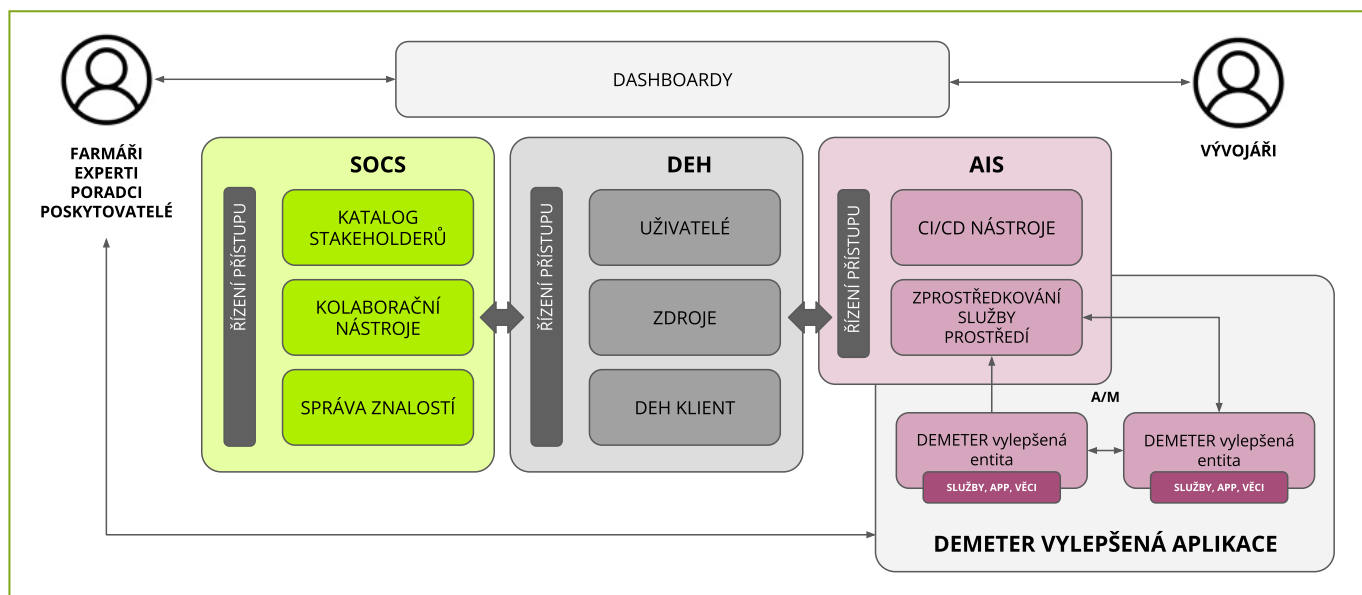
Scénář byl zaměřen na přípravu analytického a vizualizačního nástroje pro výdaje na přírodní zdroje a vstupy (např. hnojiva, pesticidy, pohonné hmoty, elektrickou energii, krmné suroviny) na konkrétní farmě pro agronomy a zemědělce.

- **Monitor růstových fází rostlin (Plant Growth Stage Monitor)**

Pro zemědělské poradce, kteří radí, kdy sázet, hnojit, zasahovat a sklízet, je monitor růstových fází rostlin službou, která kombinuje data ze satelitů, půdy, klimatu, počasí, modelu plodin a senzorů a poskytuje informace o růstových fázích specifických pro danou lokalitu a střednědobé předpovědi.

- **Systém pro podporu rozhodování v oblasti ochrany plodin (Decision Support System for Crop Protection)**

Modelování vývoje hromadných změn u hmyzu v závislosti na klimatu a zeměpisné poloze.



Obrázek 56: Technický ekosystém DEMETER (upraveno dle [138]).

16.3. DEMETER

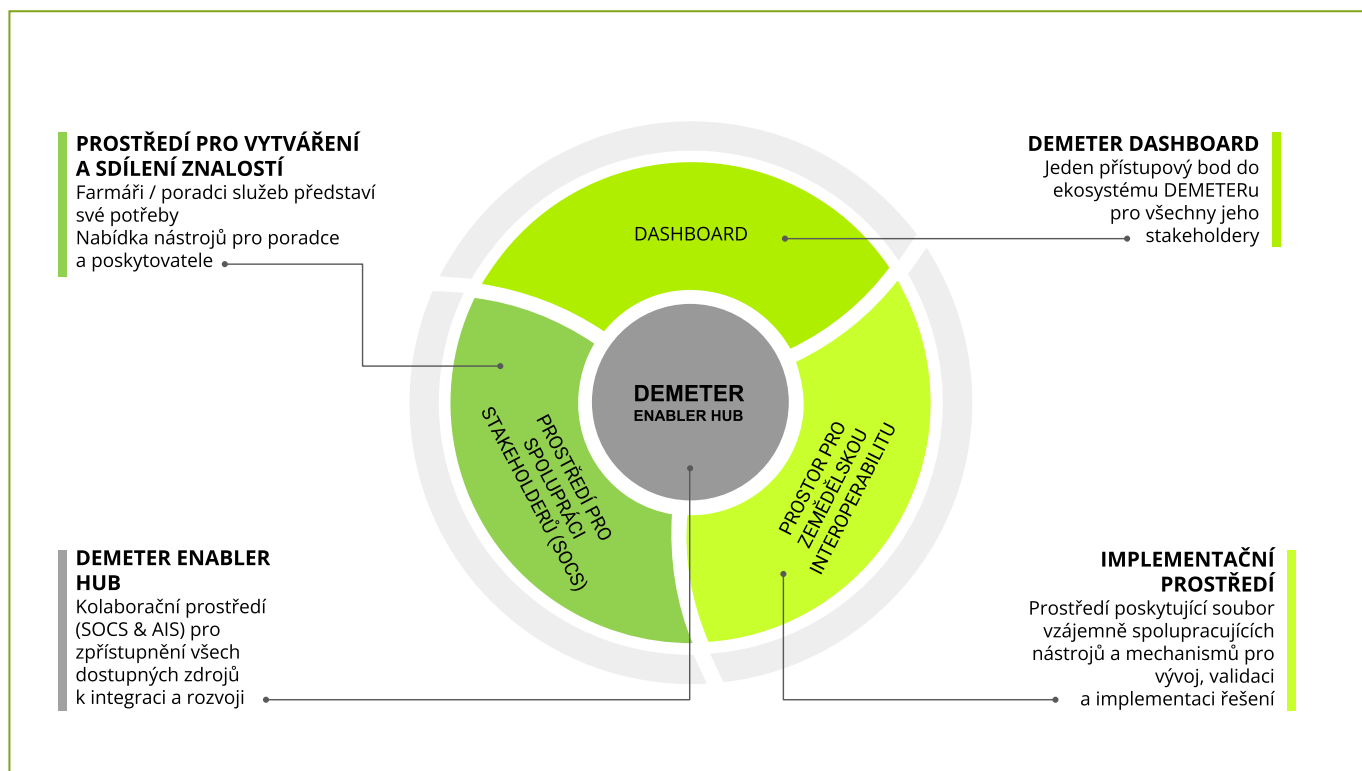
Demeter je projekt z programu Horizont 2020 (857202) podporovaný Evropskou unií [138, 139]. Cílem projektu DEMETER je vést digitální transformaci evropského zemědělsko-potravinářského odvětví prostřednictvím rychlého zavádění pokročilých technologií internetu věcí, datové vědy a inteligentního zemědělství a zajistit tak jeho dlouhodobou životaschopnost a udržitelnost.

Projekt DEMETER poskytuje zemědělcům digitální prostředky:

- Pomocí modelu "člověk v cyklu", který se neustále zaměřuje na propojení lidských znalostí a zkušeností s digitálními informacemi.
- Zaměření na interoperabilitu jako hlavní digitální prostředek, rozšíření záběru interoperability na data, služby, platformy, komunikaci M2M (machine-to-machine) a on-line inteligenci, ale také na lidské znalosti a realizace interoperability propojením zemědělců, poradců a poskytovatelů ICT řešení a strojů.
- Transformace odvětví postavením řešení na řadě digitálních technologií: internetu věcí, pozorování Země, velkých dat, umělé inteligence a digitálních postupů: spolupráce, mobility a otevřených inovací.

Tato rozhodnutí byla učiněna ve spolupráci s velkou uživatelskou základnou DEMETER (přibližně 6 000 zemědělců) a rozsáhlým pilotním pokrytím (20 pilotních projektů v 18 zemích - Belgie, Česká republika, Finsko, Gruzie, Německo, Řecko, Irsko, Itálie, Lotyšsko, Černá Hora, Norsko, Polsko, Portugalsko, Rumunsko, Srbsko, Slovinsko, Španělsko, Turecko).

V současné době již existuje celá řada systémů a platformem precizního zemědělství, které využívají mnoho různých technologií komunikace, snímání a zpracování dat. Přístup spočívající v budování nového hlavního systému, který by zahrnoval ostatní, není pro projekt s rozsáhlými pilotními činnostmi proveditelný kvůli možným problémům se škálovatelností (např. udržování stavu v rámci přístupu pub/sub) a správou (např. přístup k zemědělským datům). Proto byl v rámci projektu DEMETER zvolen zastřešující přístup integrující různorodé technologie, platformy a systémy a zároveň podporující plynulou výměnu dat v celém zemědělsko-potravinářském řetězci jako strategický směr. Toho je dosaženo prostřednictvím prostoru interoperability v zemědělství **Agricultural Interoperability Space (AIS)**, který se zaměřuje na poskytování úplného souboru mechanismů interoperability pro vývoj, ověřování a následné nasazení řešení. DEMETER nedefinuje zcela nové mechanismy interoperability, ale využívá (a rozšiřuje) širokou škálu již existujících mechanismů na úrovni senzorů, dat a služeb. AIS je navíc podporován prostředím zprostředkovatelských služeb DEMETER **Brokerage Service Environment (BSE)**, které usnadňuje nasazení aplikace s podporou DEMETER tým, že poskytuje informace o koncových bodech nabízených různými entitami rozšířeními o DEMETER (např. koncové body pro získávání dat, pro zpracování informací v nabízených prostředcích), které byly objeveny prostřednictvím **DEMETER Enabler Hub - DEH** a mají být spotřebovány. Obsažené nástroje CI/CD (vyhrazené pouze pro partnery projektu) pomáhají při integraci a nasazení různých modulů a enablers. (Obrázek 56)



Obrázek 57: Přehled hlavních konceptů systému DEMETER (upraveno dle [138]).

Protože interoperabilita dat má zásadní význam, navrhované řešení poskytuje nezbytné mechanismy převodu dat kombinující použití sémantického datového modelu **Agriculture Information Model - AIM** spolu s příslušnými mechanismy převodu/správy/odvozování dat, které přebírají široce rozšířená standardizovaná řešení. (Obrázek 57)

Všechny registrované zdroje jsou vývojářům zpřístupněny prostřednictvím DEMETER Enabler Hub (DEH), aby bylo možné řídit nasazení na základě možností (nebo omezení) přijatých technologií i faktorů vlastnictví zdrojů. Hub rovněž propojuje zdroje s otevřeným prostorem pro spolupráci zainteresovaných stran **Stakeholders Open Collaboration Space - SOCS**, aby bylo možné jejich objevování z hlediska funkcí / vlastností, které budou využity při řešení výzev prostřednictvím mechanismu spoluvytváření. A konečně systém řízení přístupu **Access Control System - ACS** je složka, která má na starosti řízení přístupu ke zdrojům DEMETER Cloud bezpečným a efektivním způsobem.

Cílem DEMETERu je usnadnit vytváření aplikací z existujících dat, komponent a systémů a podle potřeby vyvíjet nové. Aby spolu všechny tyto aplikace mohly komunikovat a spolupracovat, je nutné používat společný datový model pro všechna vyměňovaná data. V opačném případě je nelze používat společně. Model AIM poskytuje tento společný datový formát (a jazyk) pro všechna vyměňovaná data. Kromě toho, protože si přejeme interoperabilitu s nejnámějšími a nejpoužívanějšími ontologiemi, modely a slovníky, které se již používají, jsme ve skutečnosti

vytvořili AIM tak, že jsme převzali koncepty a slovníky ze všech těchto existujících modelů a sloučili je dohromady, čímž jsme zajistili zpětnou interoperabilitu se všemi těmito modely.

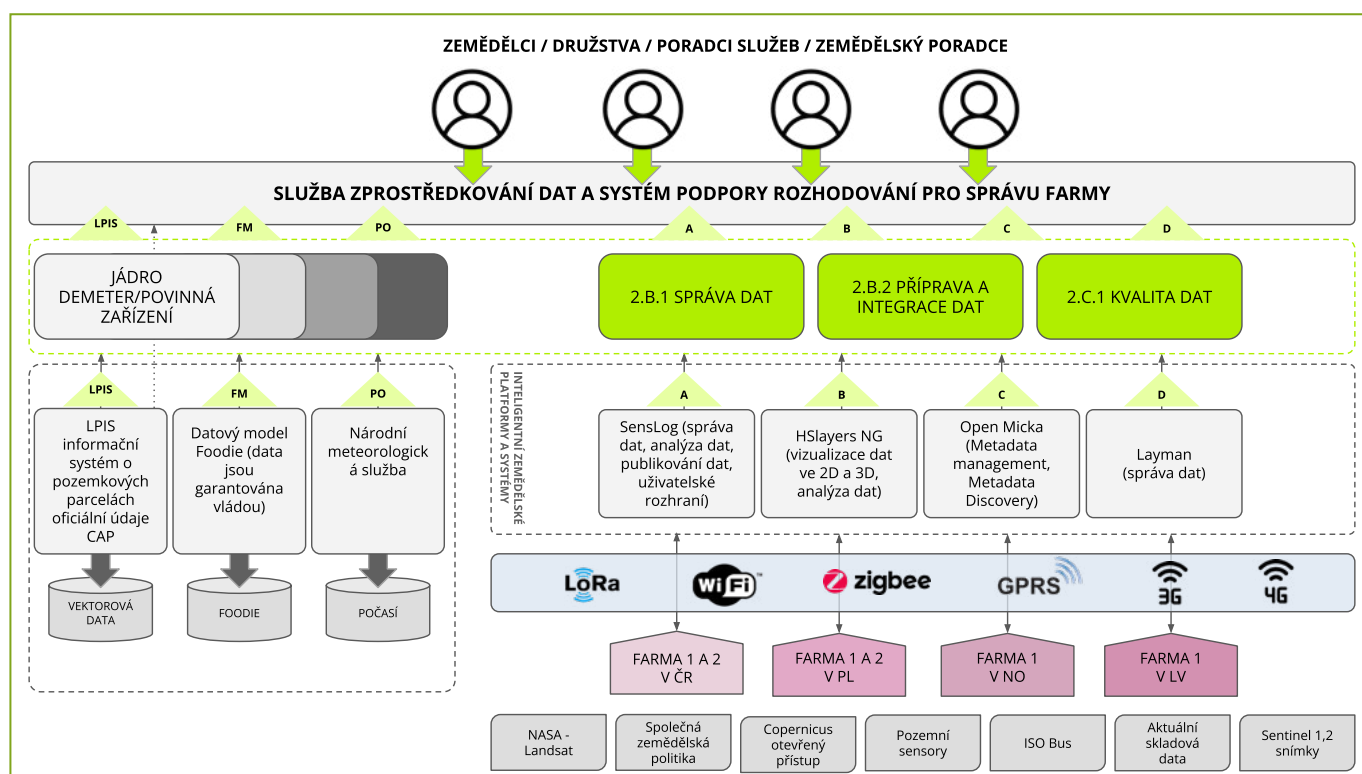
16.3.1. Český pilot

Cílem českého pilotního projektu [140] je vytvořit důvěryhodný a kompatibilní trh s daty zemědělských podniků, který bude stát mezi vlastníky a provozovateli zemědělských datových cloudů a zemědělci a který bude zahrnovat technickou platformu a poradenské služby, které zajistí snadné přijetí dat a technologií zemědělci.

Plánem je využití aplikací pro integraci, analýzu a vizualizaci dat pro systém podpory rozhodování. Pilotní projekt bude zpočátku nasazen minimálně na 6 místech v následujících zemích: Česká republika - 2 farmy, Polsko - 2 farmy, Lotyšsko - 1 farma nebo regionální organizace a Norsko - 1 farma. Důvodem pro stanovení tohoto cíle je, že zemědělci využívají mnoho informačních a technických systémů a technologií pro:

- Sběr, analýzu a přípravu dat pro využití v zemědělských podnicích.
- Organizaci práce na farmě.
- Řízení zemědělských procesů a řízení strojů.
- Organizace života farmy, plánování, prognózování.
- Ukládání a archivaci dat.

Mnohdy jsou tyto systémy a služby vyvíjeny nebo provozovány různými společnostmi. Používají nezávislé komunikační protokoly a na jejich základě - systému všech zařízení nejsou schopny organizovat zprostředkování dat na



Obrázek 58: Referenční architektura instancovaná pro pilot Lesprojekt – služby, s.r.o. (upraveno dle [138, 140]).

farmě. Na základě popisu této skutečnosti je třeba hledat řešení, která tento stav zlepšují. Je třeba hledat společné řešení, kdy zemědělec bude mít přístup ke kompletním datům.

Projekt DEMETER usnadní dosažení těchto cílů, neboť Agriculture Information Model - AIM je založen na sémantickém mapování mezi různými datovými modely a standardy a prostředí DEMETER Enabler Hub – DEH a Brokerage Service Environment – BSE umožňuje uživatelům a poskytovatelům dat nebo služeb objevit a využívat jiné nástroje Demeter, které poskytují data a umožňují využívat data mapovaná na AIM v různých analýzách prováděných jinými nástroji DEMETER. (Obrázek 58)

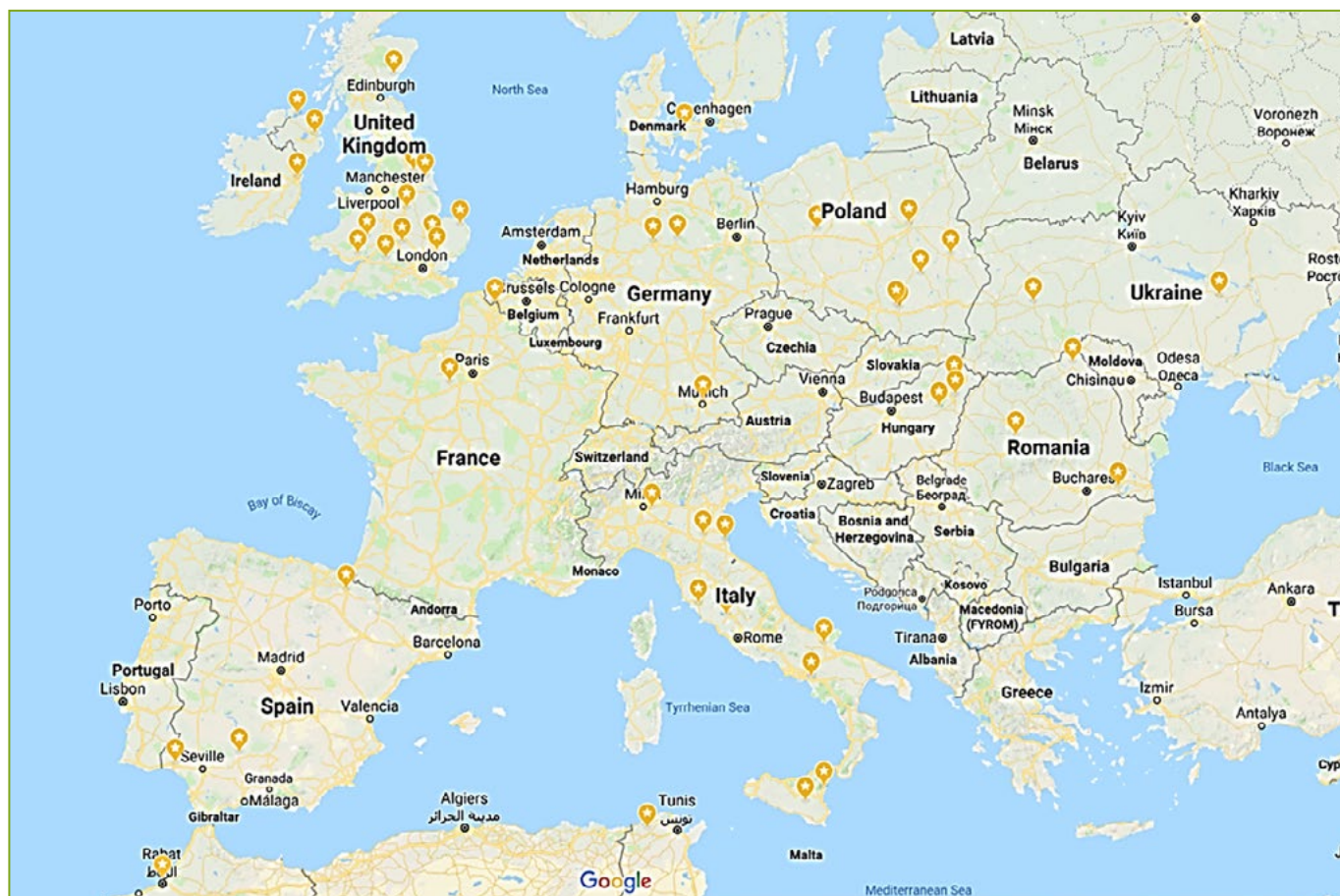
16.4. INNOVAR

InnoVar je projekt [141] reagující na výzvu Evropské komise k hledání inovací v oblasti testování plodin. Na projektu se podílí přes dvacet partnerů napříč EU se zastoupením ve všech dominantních evropských agroklimatických zónách.

V zemědělství (včetně zahradnictví a dalších forem prvovýroby) je stále větší tlak na snížení závislosti na vnějších vstupech, minimalizaci ekologické stopy a vyrovnávání se s proměnlivými klimatickými podmínkami. V této souvislosti je nutné šlechtění rostlin dále vyvíjet a systematictěji zohledňovat ty vlastnosti, které přispívají k odolnosti plodin vůči biotickým a abiotickým stresům. To přináší také zavedení kritérií a metod pro testování výkonnosti nových odrůd rostlin za podmínek spojených s udržitelnými a variabilnějšími zemědělskými postupy. To zahrnuje:

- Identifikaci charakteristik plodin a „kritérií udržitelnosti“ spojených se schopností nových odrůd udržet výnos za proměnlivějších podmínek a za udržitelnějších postupů hospodaření s plodinami.
- Vyvinutí metody a nástrojů pro integraci „kritérií udržitelnosti“ do testování výkonnosti v různých agroekologických prostředích, typech půdy a podmínkách pěstování.
- Zlepšení přesnosti a rychlosti metod testování odrůd.
- Hledání synergie mezi dvěma klíčovými metodami hodnocení odrůd (VCu, DUS).

Projekt InnoVar si dává za cíl vyvinout novou generaci testů a metodik testování zemědělských plodin umožňující vyhledání optimálních odrůd pro farmy v evropském kontextu.



Obrázek 59: Testovací lokality projektu InnoVar (upraveno dle [141]).

Součástí projektu je i vývoj databáze integrující historická data (výzkumné ústavy, univerzity, pěstitelé) a data získaná v průběhu projektu (desítky testovacích lokalit - Obrázek 59, stovky odrůd v několika pěstebních kontextech a replikacích, čtyři úrody během trvání projektu). Pilotní fáze projektu je zaměřená na pšenici, následně je plánovaná aplikace vyvinutých nástrojů na další plodiny.

Projektová databáze obsahuje:

- Seznam plodin a odrůd (genotypů).
- Seznam geolokací, kontrolních bodů a osevních míst.
- Fenotypická data včetně historických.
- Fenomická data z vybraných lokací - vizuální data sbíraná z dronů.
- Půdní data - od chemického složení po průběžný stav vlhkosti a teploty v různých půdních horizontech.
- Meteorologická data.
- Genetická data - SNP analýzy kultivovaných genotypů.
- Analytická data
 - Vztahy mezi fenotypickými a fenomickými daty a genetickými determinanty.
 - Výstupy genové asociativní studie mezi genetickými markery a zájmovými VCU a DUS vlastnostmi.
 - Výstupy navržených analytických modelů.

Celkově je v databázi ukládáno přes čtyři sta atributů (mimo fenomická, genetická a analytická data), které jsou řízeny v předem definovaném kontextu (cca. deset dimenzí).

Uvedené datové úložiště slouží také k dotazování, agregaci a exportování dat pro vyvinuté modely strojového učení a standardní analytické účely (Rdata, JSON).

17. DIGITÁLNÍ INOVAČNÍ HUBY - BUDOUCNOST, JAK TRANSFORMOVAT VELKÁ DATA DO ZNALOSTÍ

Digitální inovační huby neboli centra pro digitální inovace jsou definovány jako ekosystémy, které se skládají z malých a středních podniků, velkého průmyslu, start-upů, univerzit, výzkumníků, vývojářů, akceleratorů a investorů. Jejich cílem je vytvořit co nejlepší podmínky pro dlouhodobý technologický růst a obchodní úspěch všech zúčastněných. Tato centra by měla sloužit jako kontaktní místa s řadou podpůrných služeb pro malé a střední firmy v regionu (i mimo něj) a umožňovat přístup k digitálním znalostem, technologiím, prototypovým řešením, testovacím zařízením, kolaboračním nástrojům a zapojení do inovačního ekosystému v regionu. Tvorba digitálních inovačních hubů je podporována programem Digital Europe a rámcovým programem Horizon Europe, který bude stěžejní pro podporu výzkumu, vývoje a inovací v Evropské Unii.

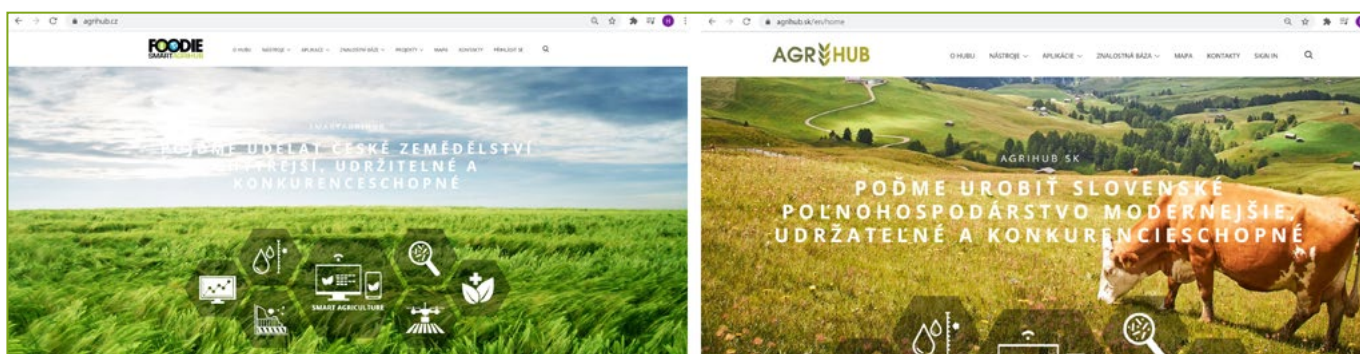
Projekt AgriHub CZ&SK, digitální inovační centra pro Českou republiku a Slovensko [142], podpoří digitální inovace v zemědělsko-potravinářské oblasti a realizuje celkem 7 inovačních experimentů (IE) s využitím dodatečných finančních prostředků (Obrázek 60). Hlavní myšlenkou dvanáctiměsíčního projektu je integrovat aktivity stávajících inovačních hubů v Česku a na Slovensku a podpořit interakci všech aktérů v zemědělském výrobním řetězci (zemědělci a farmáři, poradci, výrobci strojů, výzkumníci, vývojáři) s cílem posunout české a slovenské zemědělství směrem k SmartFarming 4.0. Cílem této

platformy je vybudovat prostředí, které pomůže podpořit propojení organizací a jednotlivců, mobilizaci talentů, v zemědělsko-potravinářském řetězci a usnadní spolupráci za účelem vybudování nového chytrého, ekologicky, sociálně a ekonomicky udržitelného zemědělství.

Cílem tohoto projektu je vyvinout inovační ekosystém pro tvorbu nových datových služeb pro SmartFarming 4.0 prostřednictvím následujících nástrojů:

Technologické

- Budování otevřené cloudové a datové infrastruktury pro více zemí.
- Agregace otevřených dat jako jsou geoprostorová data shromážděná podle směrnice INSPIRE, družicová data (COPERNICUS Sentinel-2, Sentinel-1, data Landsat, data od komerčních poskytovatelů družic), data o zemědělských podnicích, data o strojích, data Corine Land Cover (OLU - OpenLandUse), data LPIS atd. a zajištění jednotného přístupového místa ke všem potřebným informacím.
- Podpora interoperability se stávajícími digitálními inovačními centry v regionu a se stávajícími veřejnými a soukromými platformami. Podpora standardizace informací souvisejících se zemědělstvím, podpora standardizačního úsilí mezinárodních organizací, jako jsou OGC, IGAD RDA, BDVA, a také úsilí INSPIRE a DestinationE.



Obrázek 60: Vstupní webové stránky <https://www.agrihub.cz/> (vlevo) a https://www.agrihub.sk (vpravo).

- Rozvoj produktů a služeb pro SmartFarming 4.0. a propagace a propojování existujících řešení.
- Stimulace spoluvytváření nových datových služeb pro malé, střední i velké zemědělce a poskytovatele SW.
- Spolupráce s výrobou potravin a podpora lokálního odbytí.

Organizační

- Vytvoření vazeb osob, firem a jiných subjektů se znalostmi a technologiemi.
- Vytvoření prostředí, ve kterém budou moci všechny zúčastněné strany sdílet své znalosti.
- Vytvoření přímých odkazů na výzkumníky, vývojáře a hlavně na konečné uživatele - zemědělce a poradce.
- Přizvání dalších výzkumných organizací, vývojářů i farmářů ke spolupráci a provádění společných experimentů.

Vzdělávání a osvěta

- Zajistit školení pro zemědělce v oblasti precizního zemědělství.
- Zvyšování povědomí a využívání datových služeb precizního zemědělství pro vládní nebo veřejný sektor.
- Zvyšování povědomí o zemědělství ve veřejnosti jako o oboru, který může být velice moderní a přívětivý pro životní prostředí.

17.1. DEMONSTRAČNÍ APLIKACE S DATY DRUŽIC SENTINEL-1 A SENTINEL-2

Konkrétní součástí technologických nástrojů Digitálních inovačních hubů jsou praktické aplikace dat družic Sentinel do zemědělství. V současné chvíli Digitální inovační hub pro Českou republiku²⁹ nabízí několik aplikací a další jsou dále vyvíjeny a přidávány.

Jednou ze zmíněných ukázkových aplikací je Klasifikace plodin na základě časových řad snímků Sentinel-1 a Sentinel-2. Cílem satelitní klasifikace plodin je rozlišení a identifikace druhů plodin na poli v různých fázích s co nejvyšší přesností. Podle zvolené metody můžeme technologii rozdělit na dva druhy:

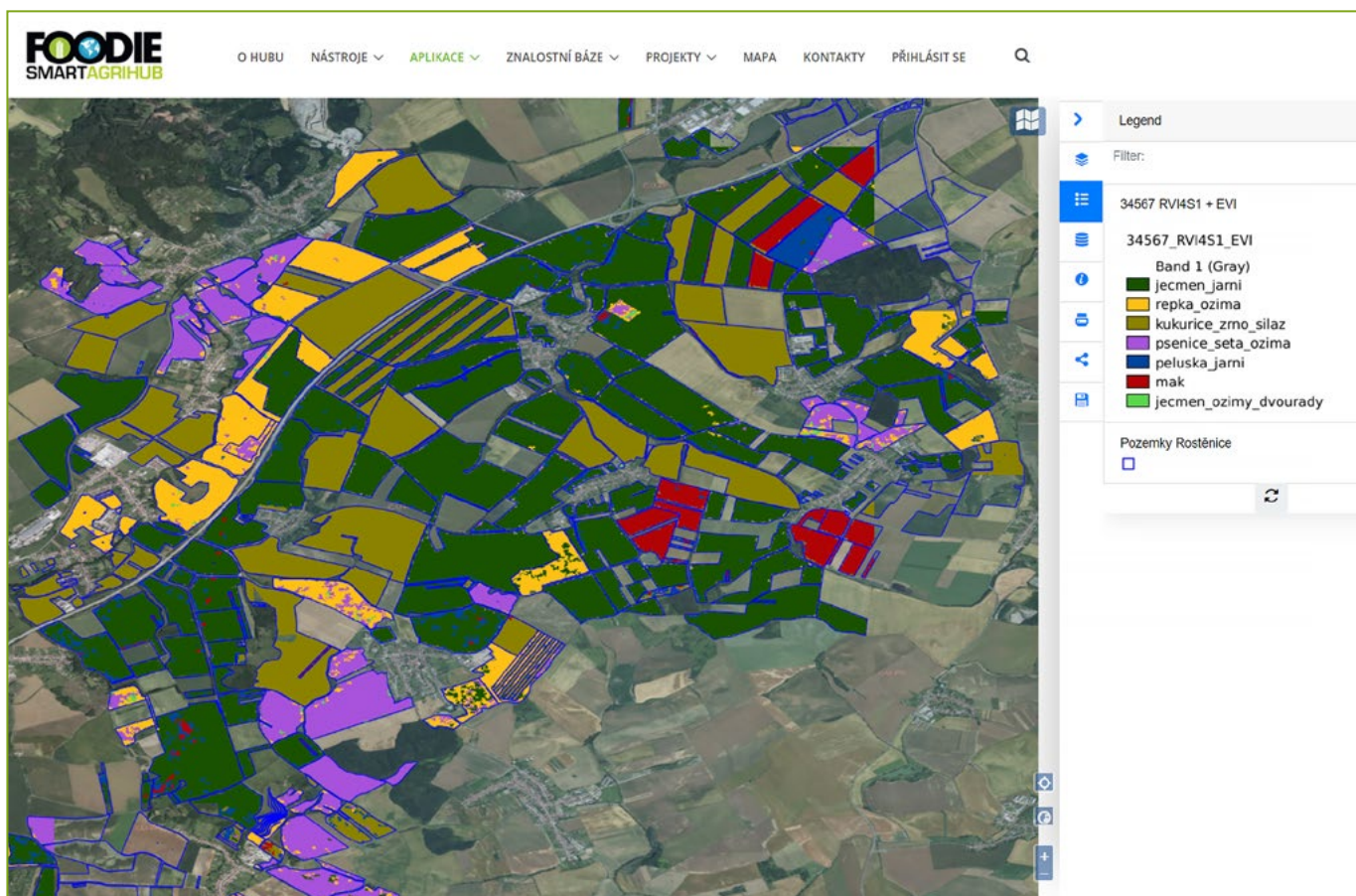
1. Neřízená klasifikace shlukuje pixely s podobnými spektrálními projevy do klastrů. V praxi to znamená, že podobný půdní pokryv (plodiny) je zahrnut do jedné skupiny a odlišeny tak od nepodobných plodin. Podle zájmového území se odlišuje i využití této technologie. Při aplikaci metody na celý snímek dochází k přesnému odlišení různých pokryvů půdy. Velmi dobře tak lze odlišit zastavěné území, vodní plochu, lesy, trvalý travní porost a ornou půdu (Obrázek 61). Zajímavých výsledků je dosaženo při aplikaci na zemědělské plochy. Ty jsou neřízenou klasifikací rozděleny do skupin na základě podobnosti jednotlivých plodin. Takto například nerozlišíme pšenici od ječmene, zato spolehlivě a jednoduše odlišíme kukuřici od obilnin. Důležitým aspektem neřízené klasifikace je její časová nenáročnost a relativně krátká potřeba přípravy dat. Nevýhodou při podrobnějších analýzách s větším množstvím výstupních tříd je neznalost významu třídy. Algoritmus od sebe odliší odlišné, ale přiřadí jednotlivé třídy k jejich významům nedokáže.

2. Řízená klasifikace přiřazuje jednotlivé objekty (pixely) do předem definovaných tříd na základě podobnosti v odrazivosti vstupních snímků. Na rozdíl od neřízené klasifikace jsou výstupem třídy s předem definovaným významem. Toho je dosaženo za cenu větší náročnosti při přípravě dat. Aby mohl algoritmus přiřazovat jednotlivé

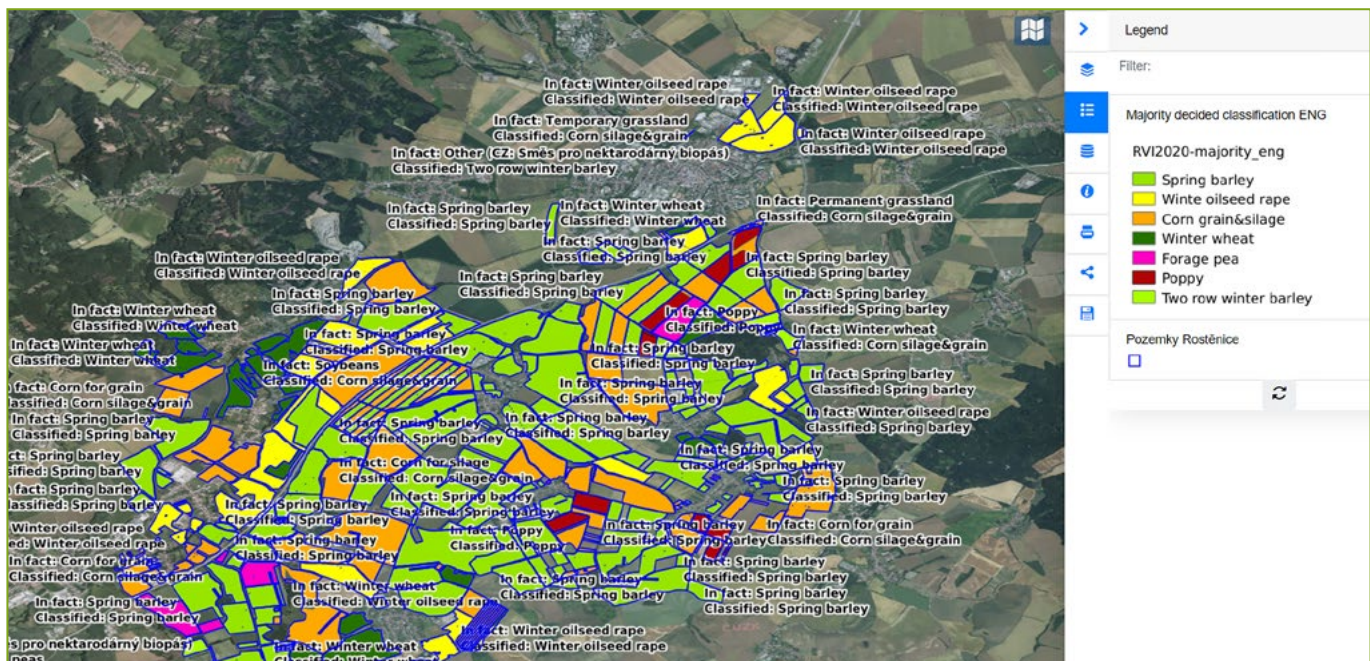
²⁹ <https://www.agrihub.cz/>



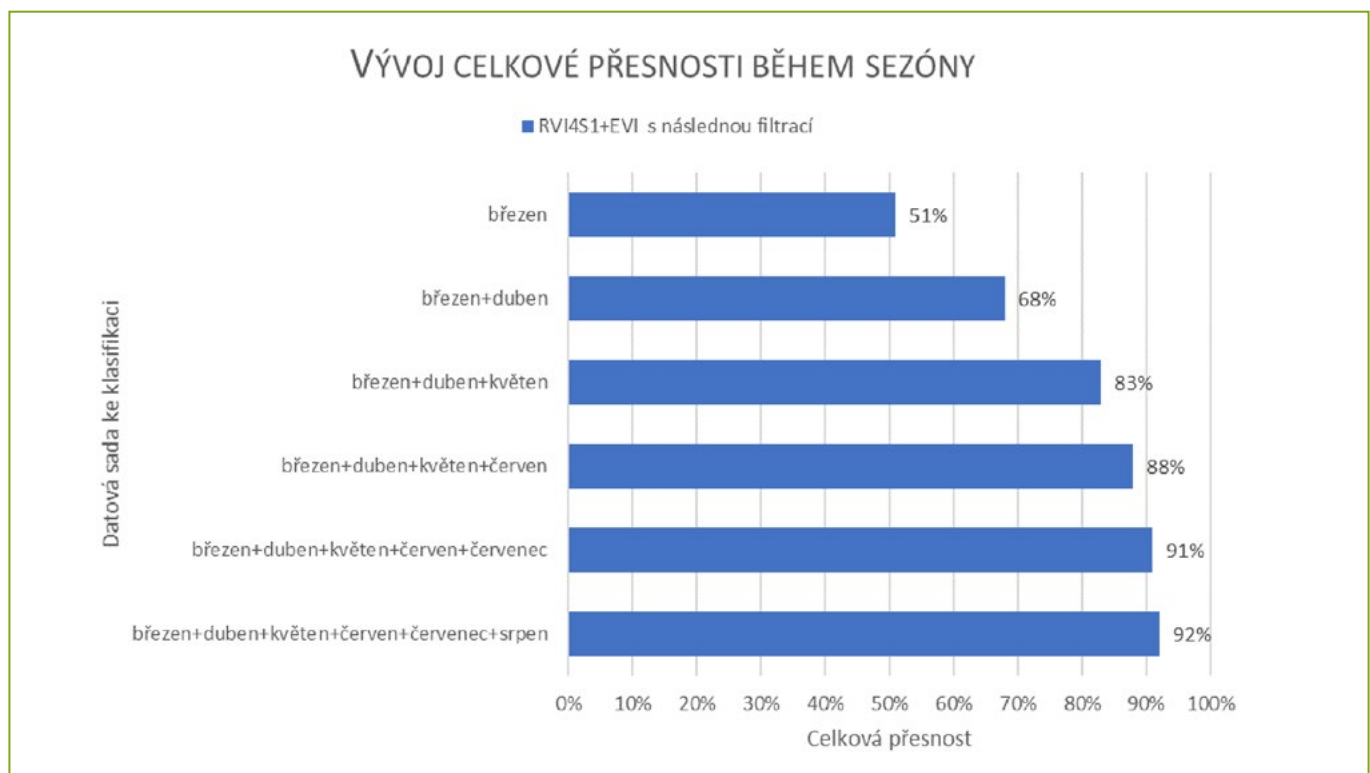
Obrázek 61: Příklad neřízené klasifikace ze slovenského Digitálního inovačního hubu.



Obrázek 62: Příklad řízené klasifikace na farmě Rostěnice v roce 2020.



Obrázek 63: Řízená klasifikace s daty družice Sentinel-1 s majoritním určením plodiny.

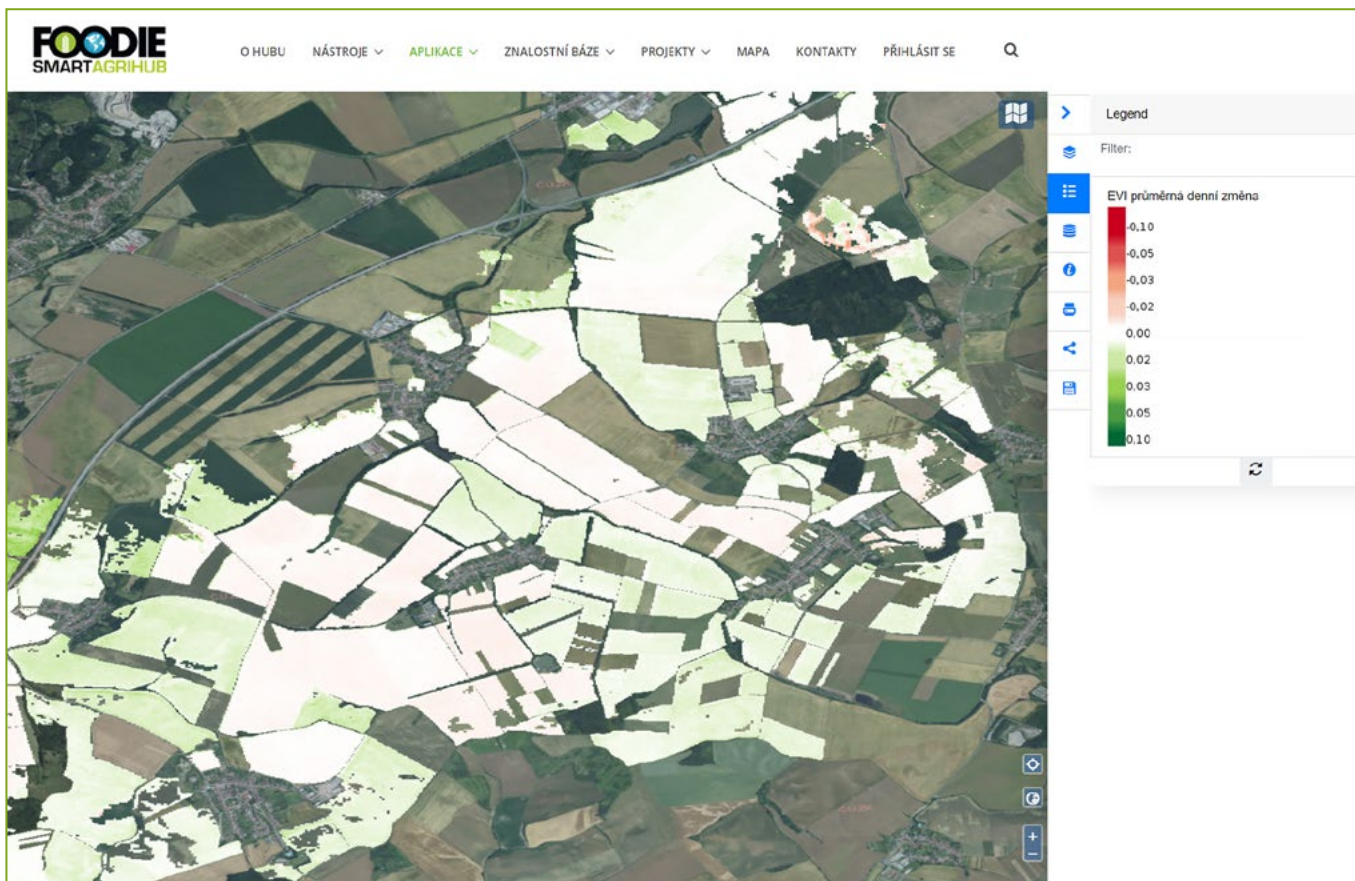


Obrázek 64: Vývoj přesnosti řízené klasifikace plodin podle kombinace dat družic Sentinel-1 a Sentinel-2.

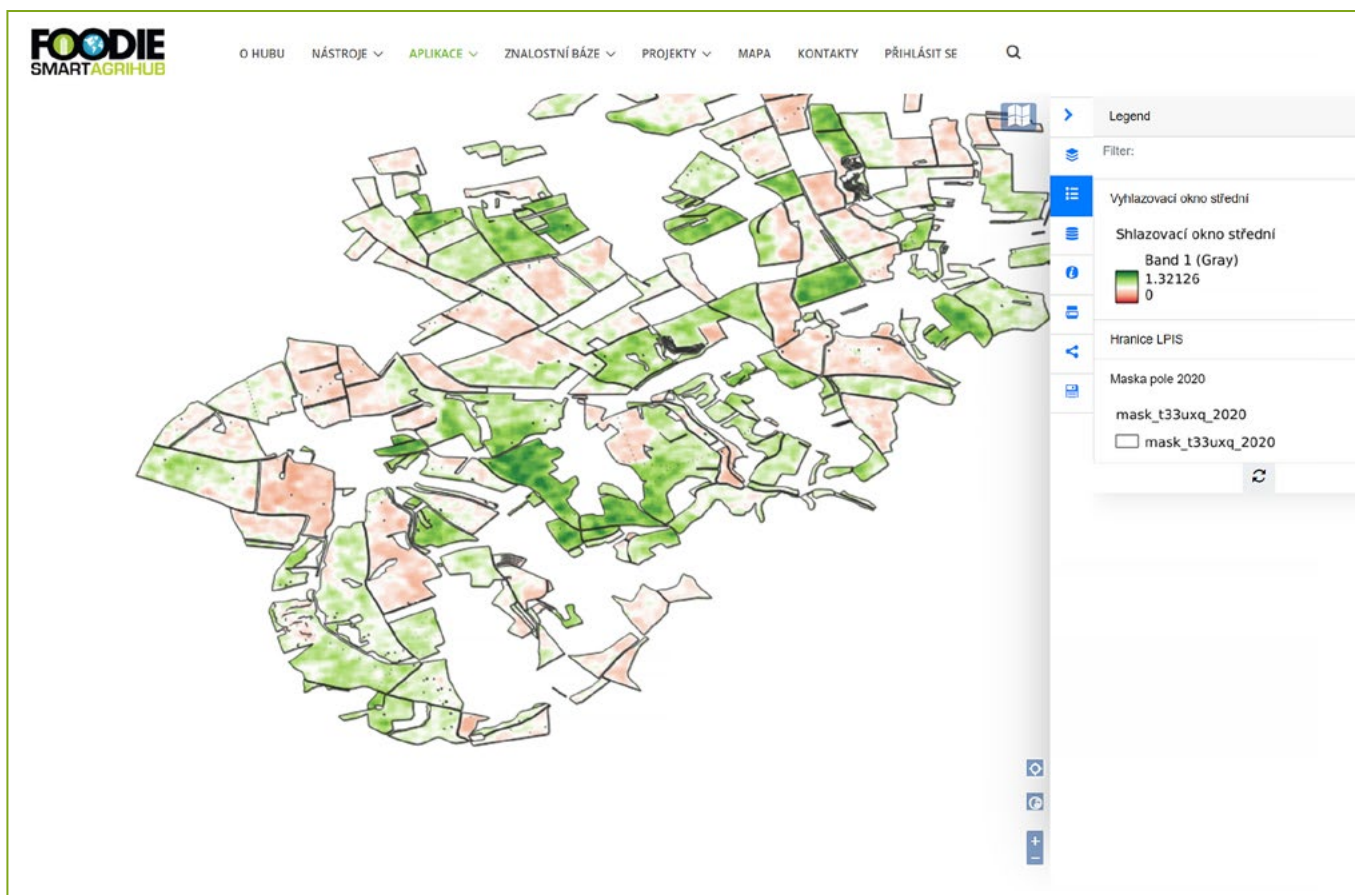
pixely do smysluplných tříd, musí být tyto třídy uživatelem předem definovány. Opět je možné tuto metodu použít pro rozsáhlé souvislé území, kdy výstupem budou již kategorie půdního pokryvu jako u neřízené klasifikace. Již zmíněná výhoda tohoto přístupu je znalost významu tříd. V evropském kontextu, kde je půdní pokryv detailně zmapovaný, je větší potenciál v řízené klasifikaci zemědělských plodin, neboť ty jsou každý rok jiné. Algoritmu dodáme satelitní snímky zemědělských pozemků, na známých pozemcích definujeme žádoucí výstupní třídy

(jednotlivé plodiny). Algoritmus poté k těmto plodinám přiřadí postupně všechny pixely (Obrázek 62). Půjdeme-li o krok dál, můžeme každému pozemku přiřadit plodinu s nejčtetnějším zastoupením (Obrázek 63).

Přesnost této klasifikace se logicky zlepšuje s množstvím vstupních snímků. Proti tomu jde zájem o výstupy a jejich využití, které jsou vyšší v raných fázích vegetační sezóny než po sklizni. Vývoj této přesnosti znázorňuje Obrázek 64.



Obrázek 65: Průměrná denní změna pro farmu Rostěnice, 2018-04-11 a 2018-04-18.



Obrázek 66: Zóny z družice Sentinel-1 k 15. 4. 2020.

Sledování stavu plodin je užitečná a žádoucí aplikace dálkového průzkumu země pro podporu rozhodování v zemědělské praxi. Zobrazení vegetačních indexů na zájmovém pozemku je v dnešní době poměrně běžnou praxí. Vegetační indexy jsou ukazatele kondice porostu vzniklé kombinací vybraných spektrálních pásem z multispektrálních snímků družic. V současné době se výzkum posouvá do pokročilejších analýz časových řad satelitních snímků v průběhu celé vegetační sezóny a tvorby zón správy pozemků. Český digitální inovační hub obsahuje příklady obou těchto inovativních přístupů:

1. V rámci výzkumu společnosti Lesprojekt-slужby s.r.o. byla vyvinuta metoda Průměrné denní změny indexu. Ta umožňuje vizualizovat vývoj kondice porostu během sezóny a zemědělcům tak pomáhá identifikovat aktuální problémy v porostu. Na Obrázku 65 je vidět průměrná denní změna Enhanced Vegetation Index mezi daty 11. a 18. dubna 2018. Pro agronoma je patrné, že vývoj porostu na zasetých pozemcích mezi zkoumanými daty probíhal standardně bez negativních jevů.
2. Zemědělství tlačeno nároky na vyšší produkci na stejné ploše s použitím méně umělých přípravků se za poslední roky významně zefektivnilo. Velkou zásluhu na tom má stále probíhající proces přesunu k preciznímu zemědělství. Jeho základem je odmítnutí uniformního pohledu na pozemek jako homogenní plochu a přijetí jeho variabilních charakteristik. Tyto vlastnosti pozemku lišící se v jeho částech podmiňují také různý přístup zemědělců k jednotlivým zónám, co se týče praktických úkonů jako setí, aplikace hnojiv a přípravků na ochranu rostlin a také sklizně. Krédem precizního zemědělství je ve správný čas na správném místě ve správném množství. Nezbytným předpokladem k tomuto rozdílnému přístupu k jednotlivým zónám na jednom poli je správné vytvoření těchto zón správy pozemku. Navíc se tyto zóny mohou vyvíjet během sezóny a měnit podle plánovaného zásahu. Na Obrázku 66 je příklad takových zón pro aplikaci dusíku vytvořený z družice Sentinel-1 a publikovaný na českém Digitálním inovačním hubu.



18. NĚCO NA ZÁVĚR

Pokusme se na závěr o malé shrnutí. Tato publikace v krátkosti představila problematiku Velkých dat ve vztahu k zemědělství. Vzhledem ke svému rozsahu nemůže být vyčerpávající, ale může alespoň pomoci orientovat se v základních pojmech. Snažili jsme se publikaci psát tak, aby poskytla orientaci i neoborníkovi na informační technologie (i když jsme se některým technickým detailům neubránili - netechnikům doporučuji přeskočit). Pro ty, co se chtějí v uvedené problematice orientovat podrobněji, jsme se snažili uvést dostatečné množství zdrojů. Tomu, kdo má zájem o hlubší orientaci v problematice, doporučujeme především knihu Big Data in Bioeconomy [1], ze které jsme do velké míry vycházeli. Tato publikace rozebírá celou problematiku do větších podrobností a technických detailů. Kromě této knihy jsme se opírali i o řadu předchozích mezinárodních výzkumů, na kterých se náš kolektiv podílel. Zde je velice zajímavé, že i přes dynamický rozvoj v informačních technologiích a i ve vědě, zůstávají mnohé věci v platnosti deset, patnáct i více let. To ukazuje na to, že někdy zavádění nových věcí do praxe probíhá pomaleji, než si původně představujeme. Přesto lze optimisticky říci, že v posledních letech u nás vnímáme posun kupředu.

Chtěli bychom se pokusit shrnout obsah publikace do několika stručných bodů:

1. Velká data mají již dnes pro zemědělství velký význam a jejich důležitost poroste, ale je potřeba si uvědomit, že velká data nejsou cílem, ale jenom prostředkem ke generování znalostí a k moudrým rozhodnutím.
2. Farmáři, vlastníci lesa, ale i úředníci na všech úrovních nejsou primární uživatelé dat, ale potřebují právě znalosti vytvořené na základě dat. Je třeba postoupit od shromažďování dat k jejich využívání.
3. Moudrá rozhodnutí lze vytvářet pouze na základě hluboké analýzy dat a bez této analýzy je značné riziko, že naše rozhodnutí nebudou moudrá. To neplatí jenom na úrovni farmy, ale i na úrovni veřejného sektoru, a to jak při formulování politik, tak při definování strategií.
4. Aby bylo možné z dat skutečně získat relevantní znalosti, je třeba spolupráce odborníků z různých oborů.
5. Velká data se vyznačují nejen svým rozsahem, ale i různorodostí a i faktem, že tato data jsou rozptýlená na mnoha místech a uložena v mnoha datových formátech a rozdílně přístupná. To značně snižuje využitelnost těchto dat a přístup k nim.
6. K tomu, aby byla velká data skutečně použitelná, je třeba důsledně uplatňovat FAIR principy, data musí být dohledatelná, přístupná, interoperabilní a znovu použitelná. To neznamená, že všechna data musí být otevřená, ale musí existovat jasná pravidla pro jejich sdílení.
7. Veřejný sektor by měl v maximální možné míře (tam, kde tomu nebrání ochrana osobních informací, bezpečnost státu apod.) zpřístupnit veškerá svá data jako volná.
8. V souladu s předchozím je nutné věnovat úsilí standardizaci na všech úrovních.
9. Dnes existuje řada zkušeností z produktů a je třeba předchozí výsledky optimálně využít.
10. Pro zemědělce je třeba zajistit lokální/regionální přístupový bod, který pro něho bude integrovat data a pomáhat mu vytvářet znalosti.



SEZNAM OBRÁZKŮ

- Obrázek 1: Cíle udržitelného rozvoje (převzato z [5]).
- Obrázek 2: Zelená dohoda pro Evropu [6].
- Obrázek 3: Hierarchie data-informace-znalosti-vědění podle Ackoffa (upraveno dle [10]).
- Obrázek 4: Úrovně řízení znalostí na farmě a vliv externích faktorů (upraveno dle [12]).
- Obrázek 5: Chudák farmář (upraveno dle [15]).
- Obrázek 6: Kdo pracuje a nepracuje s velkými daty (upraveno dle [15]).
- Obrázek 7: Řetězec organizaci přidávajících datům hodnotu (upraveno dle [22]).
- Obrázek 8: Hierarchie data-informace-znalosti-vědění v životě regionů (upraveno dle [15]).
- Obrázek 9: Linked Open Data cloud diagram [56].
- Obrázek 10: Životní cyklus dat (upraveno dle [31]).
- Obrázek 11: Ukázka nabídky připravených mapových kompozic na portále AgriHub.cz.
- Obrázek 12: Ukázka přidávání vrstev do mapové aplikace přímo z datového katalogu.
- Obrázek 13: Ukázka několika datových sad v prostředí metadatového katalogu Micka.
- Obrázek 14: Ukázka základního metadatového záznamu vrstvy integrovaného do mapového okna.
- Obrázek 15: Ukázka detailního metadatového popisu vrstvy.
- Obrázek 16: FOODIE Datový model - Farma (převzato z [65]).
- Obrázek 17: FOODIE Datový model - Díl půdního bloku (převzato z [65]).
- Obrázek 18: FOODIE Datový model - Zemědělská parcela (převzato z [65]).
- Obrázek 19: FOODIE Datový model - Management zóna (převzato z [65]).
- Obrázek 20: FOODIE Datový model - Mapa vzorků (převzato z [65]).
- Obrázek 21: Datový model prvku využití půdy.
- Obrázek 22: Přejech od INSPIRE modelu prvku využití půdy k modelu OLU.
- Obrázek 23: Algoritmus tvorby OLU České republiky.
- Obrázek 24: Odvození způsobu využití parcely z datové sady nižšího rozlišení.
- Obrázek 25: Ukázka vizualizace OLU mapy.
- Obrázek 26: Příklad výběru konkrétního pole a jeho atributů v Plzeňském kraji zobrazený v softwaru QGIS.
- Obrázek 27: Příklad polí v oblasti Plzeňského kraje vizualizovaných ve webové mapě SIEUSOIL Hubu.
- Obrázek 28: Mapa orientace svahů lokality v Karlovarském kraji.
- Obrázek 29: Zobrazení dle klasifikace využití území HILUCS.
- Obrázek 30: Zobrazení tříd krajinného pokryvu dle CORINE.
- Obrázek 31: Datový model SPOI – verze z února 2017 (převzato z [81]).
- Obrázek 32: Grafické uživatelské rozhraní Virtuoso SPARQL Query Editor pro dotazování pomocí jazyka SPARQL (převzato z [81]).
- Obrázek 33: Mapový klient SPOI (data z října 2021).
- Obrázek 34: Propojení databáze OLU, SPOI a 3D digitálního modelu terénu do jedné webové aplikace [83].
- Obrázek 35: Digitální služby (upraveno dle [88]).
- Obrázek 36: Integrace dat o výnosovém potenciálu (3D mapy) s meteorologickými daty (časové řady) (převzato z [115, 116]).
- Obrázek 37: Vizualizace senzorových a satelitních dat.
- Obrázek 38: Snímek obrazovky aplikace zobrazující výsledek případu použití typů plodin na základě Linked Data (převzato z [115]).
- Obrázek 39: WebGLayer zobrazující výnosový potenciál (převzato z [115]).
- Obrázek 40: Webový klient (převzato z [122]).

Obrázek 41: Správa projektu v QGIS (převzato z [122]).
Obrázek 42: Techniky analýzy dat (upraveno dle [123]).
Obrázek 43: Proces dolování dat – přeměna dat ve znalost (upraveno dle [123]).
Obrázek 44: Vysokourovňová architektura cloudu (upraveno dle [129]).
Obrázek 45: Platformy cloudové architektury pro dálkový průzkum Země (upraveno dle [29, 129]).
Obrázek 46: Struktura datové a zpracovatelské platformy (upraveno dle [29, 129]).
Obrázek 47: Kontext architektonických komponent cloudové architektury pro dálkový průzkum Země (upraveno dle [29]).
Obrázek 48: Chudák farmář v době cloudu.
Obrázek 49: Sen velkých hráčů o ovládnutí informací v zemědělském sektoru.
Obrázek 50: Koncept data brokerů.
Obrázek 51: Obecná linka řetězení operací v projektu DataBio (upraveno dle [20]).
Obrázek 52: BDV referenční architektura (upraveno dle [20]).
Obrázek 53: Portál s mapou výnosového potenciálu pro celou ČR (upraveno dle [132]).
Obrázek 54: Webová aplikace pro přípravu doporučení (upraveno dle [132]).
Obrázek 55: Celková architektura projektu EUXDAT (upraveno dle [135]).
Obrázek 56: Technický ekosystém DEMETER (upraveno dle [138]).
Obrázek 57: Přehled hlavních konceptů systému DEMETER (upraveno dle [138]).
Obrázek 58: Referenční architektura instancovaná pro pilot Lesprojekt – služby, s.r.o. (upraveno dle [138, 140]).
Obrázek 59: Testovací lokality projektu InnoVar (upraveno dle [141]).
Obrázek 60: Vstupní webové stránky <https://www.agrihub.cz/> (vlevo) a <https://www.agrihub.sk> (vpravo).
Obrázek 61: Příklad neřízené klasifikace ze slovenského Digitálního inovačního hubu.
Obrázek 62: Příklad řízené klasifikace na farmě Rostěnice v roce 2020.
Obrázek 63: Řízená klasifikace s daty družice Sentinel-1 s majoritním určením plodiny.
Obrázek 64: Vývoj přesnosti řízené klasifikace plodin podle kombinace dat družic Sentinel-1 a Sentinel-2.
Obrázek 65: Průměrná denní změna pro farmu Rostěnice, 2018-04-11 a 2018-04-18.
Obrázek 66: Zóny z družice Sentinel-1 k 15. 4. 2020.



SEZNAM ZKRATEK

ACS – Access Control System	GIS – Geografický informační systém
ADES – Application Deployment and Execution Service	GML – Geography Markup Language
AGDC – Australian Geoscience Data Cube	GMO – Genetically modified organism
AI – umělá inteligence	GPS – Global Positioning System
AIM – Agriculture Information Model	GPU – Graphical Processing Unit
AIS – Agricultural Interoperability Space	GSMML – GlobalSoilMap.net Geography Markup Language
AMI – Ambient Mobile Intelligence	HILUCS – Hierarchical INSPIRE Land Use Classification System
API – Application Programming Interface	HPC – High Performance Computing
AWS – Amazon Web Services	HR – High-spatial resolution
BDT – Big data technology	HTML – Hypertext Markup Language
BDVA – Big Data Value Association	IE – Inovační Experiment
BPEJ – bonitovaná půdně ekologická jednotka	IEEE – Institute of Electrical and Electronics Engineers
BSE – Brokerage Service Environment	IGAD – Agricultural Data Interest Group
CAP – Common Agricultural Policy	INSPIRE – Infrastructure for Spatial Information in Europe
CLC – Corine Land Cover	IoT – Internet věcí
CSW – Open Geospatial Consortium Catalogue Service	ISO – Mezinárodní organizace pro normalizaci
DCAT – Data Catalog Vocabulary	ISOBUS – communication protocol for the agriculture industry
DestinationE – Destination Earth	ISRIC – International Soil Reference and Information Centre
DIAS – Data and Information Access Services	IT – informační technologie
DIH – Digitální Inovační Hub	JSON – JavaScript Object Notation
DIKW – Data, Information, Knowledge, Wisdom	LiDAR – Light Detection and Ranging
DKM – digitální katastrální mapa	LPIS – Land Parcel Identification System
DNA – deoxyribonucleic acid	NASA – National Aeronautics and Space Administration
DOAP – Description of a Project	NG – Next Generation
DUS – Distinctness – Uniformity – Stability	NUTS3 – Nomenclature of Territorial Units for Statistics level 3
EGNOS – European Geostationary Navigation Overlay Service	O&M – Observations and Measurements
EK – Evropská komise	OASIS – Organization for the Advancement of Structured Information Standards
EMS – Execution Management Service	ODbL – Open Database License
EO – Earth observation	ODC – Open Data Cube
EOSC – European Open Science Cloud	OECD – Organisation for Economic Co-operation and Development
EOSDIS – Earth Observing System Data and Information System	OGC – Open Geospatial Consortium
EROS – Earth Resources Observation Satellite	OLU – Open Land Use
ESA – Evropská kosmická agentura	OSN – Organizace spojených národů
ESD – Evropská databáze půdy	OWL – Web Ontology Language
ESDAC – Evropské centrum pro data o půdě	PDF – Portable Document Format
EU – Evropská unie	PROBA – ESA Project for On-Board Autonomy
EU-DEM – digitální výškový model EU	RDA – Research Data Alliance
ExtJS – JavaScript framework	RDB – relační databáze
FAIR – Findable, Accessible, Interoperable and Re-usable principy	RDF – Resource Description Framework
FAO – Food and Agriculture Organization of the United Nations	RFID – Radio-frequency identification
GALILEO – EU global navigation satellite system	ROI – Return on investment
GEE – Google Earth Engine	RTK – Real-time kinematic positioning
GEOS – Global Earth Observation System of Systems	

RÚIAN – Registr územní identifikace, adres a nemovitostí
SAR – Specific absorption rate
SDGs – Sustainable Development Goals
SDI – spatial data infrastructure
SensorML – Sensor Model Language
SH – Sentinel Hub
SNP – single-nucleotide polymorphism
SOA – Service-oriented architecture
SOCS – Stakeholders Open Collaboration Space
SoilML – Soil Markup Language
SOS – Sensor Observation Service
SOTER (SoTerML) – Soil and Terrain Markup Language
SPOI – Smart Points of Interest
SPS – Sensor Planning Service
SQL – Structured Query Language
SRA – Strategický výzkumný program
SSN – XG - W3C Semantic Sensor Network Incubator Group
STA – Open Geospatial Consortium SensorThings API
SW – software
SZP – Společná zemědělská politika
TPU – Tensor Processing Unit
TWI – Topographic wetness index
UAV – Unmanned aerial vehicle
UML – Unified Modeling Language
URI – Uniform Resource Identifier
VAO – Value Added Organization
VDMA - Verband Deutscher Maschinen und Anlagenbau
VGI – Volunteered Geographic Information
VHR – Very-high-spatial-resolution
VoID – Vocabulary of Interlinked Datasets
VZU – Value for Cultivation and Use
W3C – World Wide Web Consortium
WaterML – Water Markup Language
WCS – Web Coverage Service
WFS – Web Feature Service, webová služba vzhledů
WMS – Web Map Service, webová mapová služba
WPS – Web Processing Service
WRB – World Reference Base
XML – eXtensible Markup Language



ZDROJE

- [1] Södergård, C., Mildorf, T., Habyarimana, E., Berre, A., Fernandes, J. and Zinke-Wehlmann, C., eds., 2021. Big Data in Bioeconomy [online]. Cham: Springer International Publishing. ISBN 9783030710682. Dostupné na: <<https://link.springer.com/book/10.1007/978-3-030-71069-9>> [cit. 1. 10. 2021].
- [2] Charvat, K., Gnip, P., 2009. Analysis of external drivers for farm management and their influences on farm management information systems. In: Precision agriculture '09, 9, s.915. doi: <https://doi.org/10.3920/978-90-8686-664-9>
- [3] Gnip, P., Charvat, K. Krocán, M., 2008. Analysis of external drivers for agriculture. In: World conference on agricultural information and IT, IAALD AFITA WCCA 2008, Tokyo University of Agriculture, Tokyo, Japan, 24-27 August, 2008 (s. 797-801). Tokyo University of Agriculture.
- [4] The Sustainable Development Agenda, 2021. United Nations Sustainable Development [online]. Dostupné na: <<https://www.un.org/sustainabledevelopment/development-agenda/>> [cit. 10. 10. 2021].
- [5] Osn.cz, 2021. Cíle udržitelného rozvoje (SDGs) – OSN Česká republika [online]. Dostupné na: <<https://www.osn.cz/osn/hlavni-temata/sdgs/>> [cit. 10. 10. 2021].
- [6] Evropská komise, 2019. SDĚLENÍ KOMISE EVROPSKÉMU PARLAMENTU, EVROPSKÉ RADĚ, EVROPSKÉMU HOSPODÁŘSKÉMU A SOCIÁLNÍMU VÝBORU A VÝBORU REGIONŮ Zelená dohoda pro Evropu, [online]. Brusel, 11. 12. 2019. Dostupné na: <https://www.dataplan.info/img_upload/7bdb1584e3b8a53d337518d988763f8d/cz1_cellar_b828d165-1c22-11ea-8c1f-01aa75ed71a1.0010.02_doc_1.pdf> [cit. 10. 10. 2021].
- [7] Euroskop.cz, 2021. Zelená dohoda pro Evropu - hlavní výzva pro novou Komisi [online]. Dostupné na: <<https://www.euroskop.cz/9047/34461/clanek/zelena-dohoda-pro-evropu-hlavni-vyzva-pro-novou-komisi/>> [cit. 10. 10. 2021].
- [8] Kubíčková, H., Šafář, V., Kozel, J., Král, M., Křivánek, Z., Řezník, T., Šafář, V., Šmejkal, J., Vrobel, J., Zdražil, F., Charvát, K., Horáková, Š., Bye, B. L., Mildorf, T., 2021. E04Agri D2.4 Final Workshop - User Requirements and Gap Analysis in Different Sectors Report, March 2021, E04Agri project deliverable
- [9] Charvat, K., Horakova, S., Druml, S., Mayer, W., Safar, V., Kubickova, H., Kolitzus, D., Lopez, J., Esbrí, M. A., Catucci, A., Lucau, C., Mildorf, T., Rybokienė, S., 2020. E04Agri D5.6 White Paper on Earth Observation Data in Agriculture, E04AGRI Consortium, August 2020, E04Agri project deliverable
- [10] Bellinger, G., Castro, D., Mills, A., 2004. Data, Information, Knowledge and Wisdom [online]. Dostupné ná: <<http://www.systems-thinking.org/dikw/dikw.htm>>
- [11] Krivanek, Z., Charvat, Jr. K., Gnip, P., Vohnout, P., Charvat, K., 2012. New Tools for Data Accessibility and Decision Support. IST Africa Conference 2012.
- [12] Charvat, K., Gnip, P., Vohnout, P., Charvat jr, K., 2011. VISION FOR A FARM OF TOMORROW. IST Africa Conference 2011.
- [13] Sørensen C.G., Pesonen L., Fountas S., Suomi P., Bochtis D., Bildse P., Pedersen S.M., 2010. A user-centric approach for information modelling in arable farming. In: Computers and Electronics in Agriculture 73(1):44-55, DOI: 10.1016/j.compag.2010.04.003
- [14] Sørensen, C, Bildsøe, P. Fountas, S., Pesonen, L., Pedersen, S., Basso, B. Nash, E, Deliverable 3.1, 2009. Work package 3: Analysis and specification of knowledge based farm management. FUTUREFARM project deliverable
- [15] Charvat, K., 2018. Visionary speech about the future of digital agriculture. IOF2020 PARTNER AND STAKEHOLDER EVENTS Conference, Spain, DOI: 10.13140/RG.2.2.35832.96000.
- [16] Charvat, K., Horakova, S., Wolfert, S., Holster, H., Schmid, O., Pesonen, L., Martini, D., Mietzsch, E., Mildorf, T., 2013. Final SRA 'Common Basis for policy making for introduction of innovative approaches on data exchange in agri-food industry'. DOI: 10.13140/RG.2.1.4887.4006.
- [17] Destination Earth, 2021. Shaping Europe's digital future [online]. Dostupné na: <<https://digital-strategy.ec.europa.eu/en/policies/destination-earth>> [cit. 11. 10. 2021].
- [18] Palmer, M., 2006. Data is the new oil [online]. Dostupné na: <https://ana.blogs.com/maestros/2006/11/data_is_the_new.html> [cit. 12. 10. 2021].
- [19] Laney, D. (2001). 3D data management: Controlling data volume velocity and variety. Gartner file No. 949 [online]. Dostupné na: <<https://blogs.gartner.com/doug-laney/files/2012/01/ad949-3D-Data-Management-Controlling-Data-Volume-Velocity-and-Variety.pdf>> [cit. 12. 10. 2021].
- [20] Södergård C., Mildorf T., Berre A.J., Tsalgatidou A., Charvát K., 2021, Big Data Technologies in DataBio. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. Dostupné na https://doi.org/10.1007/978-3-030-71069-9_1
- [21] Charvat, K., Horak, P., Gnip, P., Fusco, L. Van Bemmelen, J. Betti P, J., 2006. Open Agriculture System, IST-Africa

- 2006 Conference Proceedings Paul Cunningham and Miriam Cunningham (Eds) IIMC International Information Management Corporation, 2006 ISBN: 1-905824-01-7, (PDF) Open Agriculture System.
- [22] Safar, V., Charvat, K., Horakova, S., Orlickas, T., Rimgaila, M., Kolitzus, D., Bye, B. L., 2020. EO4Agri D2.2 Initial Workshop - User Requirements and Gap Analysis in Different Sectors Report, EO4AGRI Consortium, 29/04/2020
- [23] Auer, S. R.; Bizer, C.; Kobilarov, G.; Lehmann, J.; Cyganiak, R.; Ives, Z., 2007, DBpedia: A Nucleus for a Web of Open Data. The Semantic Web. Lecture Notes in Computer Science. 4825. p. 722. ISBN 978-3-540-76297-3. Dostupné na: doi:10.1007/978-3-540-76298-0_52.
- [24] Open Knowledge Foundation, 2021, Open Data Handbook [online], Dostupné na: <<https://opendatahandbook.org/>> [cit. 27.10.2021].
- [25] Ayris, P., Berthou, J.-Y., Bruce, R., Lindstaedt, S., Monreale, A., Mons, B., Murayama, Y., Södergård, C., Tochtermann, K., & Wilkinson, R., 2016, Realising the European open science cloud. European Union. Dostupné na: <<https://doi.org/10.2777/940154>>
- [26] Mons, B., Neylon, C., Velterop, J., Dumontier, M., da Silva Santos, L. O. B., Wilkinson, M. D., 2017, Cloudy, Increasingly FAIR; Revisiting the FAIR Data Guiding Principles for the European Open Science Cloud [online]. 1 Jan. 2017 : 49 – 56. Dostupné na: <<https://content.iospress.com/articles/information-services-and-use/isu824>> [cit. 29. 10. 2021].
- [27] European Commission, 2016, Horizon 2020 FAIR Data Management Plan (DMP) [online]. Dostupné na: <https://ec.europa.eu/research/participants/data/ref/h2020/gm/reporting/h2020-tpl-oa-data-mgt-plan_en.docx> [cit. 29. 10. 2021].
- [28] Berners-Lee, T., 2006, Linked Data - Design Issues [online]. Dostupné na: <<http://www.w3.org/DesignIssues/LinkedData.html>> [cit. 29. 10. 2021].
- [29] Simonis I., Charvát K., 2021, Standards and EO Data Platforms. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. Dostupné na: <https://doi.org/10.1007/978-3-030-71069-9_2>
- [30] Jonák, Z., 2003. Data. In: KTD: Česká terminologická databáze knihovnictví a informační vědy (TDKIV) [online]. Praha : Národní knihovna ČR, 2003. Dostupné na: <http://aleph.nkp.cz/F/?func=direct&doc_number=000000442&local_base=KTD> [cit. 20. 10. 2021].
- [31] DataBio, 2017, DataBio D6.2 – Data Management Plan. Various. Dostupné na: <https://www.databio.eu/wp-content/uploads/2017/05/DataBio_D6.2-Data-Management-Plan_v1.0_2017-06-30_CREA.pdf> [cit. 21. 10. 2021].
- [32] BIG DATA VALUE ASSOCIATION (BDVA), 2019. BDVA Strategic Research and Innovation Agenda [online]. Dostupné na: <http://www.bdva.eu/sites/default/files/BDVA_SRIA_v4_Ed1.1.pdf> [cit. 22. 10. 2021].
- [33] Iso.org, 2021. About us [online], Dostupné na: <<https://www.iso.org/about-us.html>> [cit. 25. 10. 2021].
- [34] Iso.org. 2017. ISO and agriculture. [online] Dostupné na: <<https://www.iso.org/files/live/sites/isoorg/files/store/en/PUB100412.pdf>> [cit. 25. 10. 2021].
- [35] W3C.org. 2021. World Wide Web Consortium (W3C). [online] Dostupné na: <<https://www.w3.org/>> [cit. 25. 10. 2021].
- [36] W3C.org. 2021. Agriculture Community Group. [online] Dostupné na: <<https://www.w3.org/community/agri/>> [cit. 25. 10. 2021].
- [37] OASIS Open. 2021. About Us - OASIS Open. [online] Dostupné na: <<https://www.oasis-open.org/org/>> [cit. 25. 10. 2021].
- [38] iee.org. 2021. IEEE. [online] Dostupné na: <<https://www.ieee.org/>> [cit. 25. 10. 2021].
- [39] Isobus.net. 2021. ISOBUS Data Dictionary. [online] Dostupné na: <<https://www.isobus.net/isobus/>> [cit. 25.10.2021].
- [40] Vest.agrisemantics.org. 2021. AGROXML / AGRORDF. [online] Dostupné na: <<https://vest.agrisemantics.org/content/agroxml-agrorrdf>> [cit. 25.10.2021].
- [41] Ogc.org. 2021. About OGC. [online] Dostupné na: <<https://www.ogc.org/about>> [cit. 25. 10. 2021].
- [42] Ogc.org. 2021. Agriculture DWG. [online] Dostupné na: <<https://www.ogc.org/projects/groups/agriculturedwg>> [cit. 25. 10. 2021].
- [43] Inspire.ec.europa.eu. 2021. About INSPIRE. [online] Dostupné na: <<https://inspire.ec.europa.eu/about-inspire/563>> [cit. 25. 10. 2021].
- [44] Leibovici, D. G., Santos, R., Hobona, G., Anand, S., Kamau, K., Charvat, K., Schaap, B, Jackson, M., 2021. Geospatial standards and their adoption through the example of the agriculture domain, Springer International Publishing [In press]
- [45] Montanarella, L., Wilson, P., Cox, S., Mcbratney, A., Ahamed, S., McMillan, B., Jacquier, D., Fortner, J., 2010. Developing SoilML as a global standard for the collation and transfer of soil

- data and information. *Geophysical Research Abstracts*, vol. 12., European Geosciences Union
- [46] Pourabdollah, A., Leibovici, D.G., Simms, D.M., Tempel, P., Hallett, S.H., Jackson, M.J., 2012. Towards a standard for soil and terrain data exchange: SoTerML. In: *Computers & Geosciences* 45, pp.270–283. DOI:10.1016/j.cageo.2011.11.026
- [47] Ritchie, A., 2016. OGC Soil Data Interoperability Experiment. Open Geospatial Consortium, No. OGC 16-088r1. Dostupné na: <https://portal.opengeospatial.org/files/?artifact_id=69896>
- [48] Hoffmann, C., Schulz, S., Eberhardt, E., Grosse, M., Stein, S., Specka, X., Svoboda, N., Heinrich, U., 2019. Data Standards for Soil- and Agricultural Research. BonaRes Data Centre. DOI:10.20387/BonaRes-ARM4-66M2
- [49] Batjes, N.H., van den Bosch, H., 2018. How well is the WDC-Soils serving its user groups?. *SciDataCon-IDW 2018*, Gaborone, Botswana. DOI:10.17027/6F38-HN42
- [50] Kempen, B., Yigini, Y., Viatkin, K., de Jesus, J.M., de Sousa, L.M., van den Bosch, R., Vargas, R., 2019. The Global Soil Information System (GloSIS)–concept and design. In: *Geophysical Research Abstracts (Vol. 21)*, European Geosciences Union
- [51] Helbig, C., Dransch, D., Böttinger, M., Devey, C., Haas, A., Hlawitschka, M., Kuenzer, C., Rink, K., Schäfer-Neth, C., Scheuermann, G., Kwasnitschka, T., Unger, A., 2017. Challenges and strategies for the visual exploration of complex environmental data. In: *International Journal of Digital Earth*, 10 pp.1070–1076. DOI:10.1080/17538947.2017.1327618
- [52] Guan, W.W., Bol, P.K., Lewis, B.G., Bertrand, M., Berman, M.L., Blossom, J.C., 2012. WorldMap – a geospatial framework for collaborative research. In: *Annals of GIS*, 18 pp.121–134. DOI: 10.1080/19475683.2012.668559
- [53] Čerba, O., Kepka, M., Macura, J., Kellar, J., 2017. SMART POINTS OF INTEREST. In: *Symposium GIS Ostrava 2017 Geoinformatika v pohybu – konferenční sborník, VŠB - Technická univerzita Ostrava*
Eds.: Tomáš Inspektor, Jiří Horák, Jan Růžička, ISBN 978-80-248-4029-1, ISSN 1213-239X
- [54] Bizer, C., Heath, T., Berners-Lee, T., 2009. Linked data-the story so far. *Semantic services, interoperability and web applications: emerging concepts*, In: *International journal on Semantic Web and information systems*, 205-227. doi: 10.4018/jswis.2009081901
- [55] Heath, T., & Bizer, C., 2011, *Linked data: Evolving the web into a global data space. Synthesis lectures on the semantic web: theory and technology*, 1(1), 1-136, <https://doi.org/10.2200/S00334ED1V01Y201102WBE001>
- [56] Twitter, 2021. The Linked Open Data Cloud [online]. Dostupné na: <https://twitter.com/lod_cloud> [cit. 26. 10. 2021].
- [57] DataBio, 2019. DataBio D7.5 – Exploitation Plan v2. DataBio Consortium [cit. 21. 10. 2021].
- [58] Šafář, V., Charvát, K., Horáková, Š., Orlickas, T., Rimgaila, M., Kolitzus, D., Bye, B. L., Kaas, O., Kepka Vichrová, M., Křivánek, Z., Vrobel, J., Kvapil, J., Křivánková, K., Leitner, F., Zadražil, F., Kozhukh, D., Šnevajs, H., Kubíčková, H., 2020. D2.3 Mid-Term Workshop - User Requirements and Gap Analysis in Different Sectors Report, EO4Agri consortium, May 2020
- [59] Charvat, K., Horakova, S., Druml, S., Mayer, W., Kolitzus, D., Esbrí, M. A., Clergue, L., 2020. Collaboration Framework for Global and EU Initiatives, EO4AGRI Consortium
- [60] Zinke-Wehlmann, C., Kirschenbaum, A., Palma, R., Brahma, S., Charvát, K., Charvát Jr, K., Reznik, T., 2021. Linked Data and Metadata. In *Big Data in Bioeconomy* (pp. 79-90). Springer, Cham. Dostupné na: <https://doi.org/10.1007/978-3-030-71069-9_7>
- [61] Charvat, K., Řezník, T., Kafka, S., Palma, R., Šplíchal, M., Kozuch, D., Beere, A., Coene, Y., 2017. User Friendly and Non-Disturbing Metadata Integration. Joint conference „Geomatics in projects“ 2017, DOI:10.13140/RG.2.2.29395.89128.
- [62] Charvat, K., Jr, K. C., Reznik, T., Lukas, V., Jedlicka, K., Palma, R., Berzins, R., 2018. Advanced visualisation of big data for agriculture as part of databio development. In: *IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium* (pp. 415-418). IEEE.
- [63] GitHub, 2021. GitHub – Wirelessinfo/FOODIE-data-model: FOODIE data model [online]. Dostupné z: <<https://github.com/Wirelessinfo/FOODIE-data-model>> [cit. 26. 10. 2021].
- [64] GitHub, 2021. GitHub – FOODIE-cloud/ontology [online]. Dostupné na: <<https://github.com/FOODIE-cloud/ontology>> [cit. 26. 10. 2021].
- [65] Řezník, T., Charvát, K., Charvát, K. Jr., Horáková, Š., Křivánek, Z., Kepka, M., 2016. INSPIRE, COPERNICUS AND GEOSS in Precision Farming: Economical and Ecological Synergies. DOI: 10.13140/RG.2.2.30302.23364.
- [66] Reznik, T., Lukas, V., Charvat, K., Horakova, S., Kepka, M., 2016. Foodie data models for precision agriculture. In: *13th International Conference on Precision Agriculture July 31 – August 4, 2016, St. Louis, Missouri, USA* (No. August 2016, p. 195).
- [67] Inspire.ec.europa.eu, 2021. Seznam kategorií využití území dle INSPIRE – Hierarchical INSPIRE Land Use Classification System (HILUCS) [online]. Dostupné na: <<http://inspire.ec.europa.eu/codelist/HILUCSValue/>> [cit. 27. 10. 2021].

- [68] Szif.cz, 2021. LPIS - Státní zemědělský intervenční fond [online]. Dostupné na: <<https://www.szif.cz/cs/lpis>> [cit. 27. 10. 2021].
- [69] Copernicus Programme, 2021. Urban Atlas - Copernicus Land Monitoring System [online]. Dostupné na: <<https://land.copernicus.eu/local/urban-atlas>> [cit. 27. 10. 2021].
- [70] Copernicus Programme, 2021. The CORINE Land Cover (CLC) [online]. Dostupné na: <<https://land.copernicus.eu/pan-european/corine-land-cover>> [cit. 27. 10. 2021].
- [71] EUROSTAT, 2021. Land use and land cover survey (LUCAS) [online]. Dostupné na: <https://ec.europa.eu/eurostat/statistics-explained/index.php?title=LUCAS_-_Land_use_and_land_cover_survey> [cit. 27. 10. 2021].
- [72] Správa územních registrů, 2021. Registr územní identifikace, adres a nemovitostí (RÚIAN) [online]. Dostupné na: <<https://www.szrcr.cz/cs/registr-uzemni-identifikace-adres-a-nemovitosti>> [cit. 27. 10. 2021].
- [73] European Soil Data Centre (ESDAC), 2021. Soil erosion by wind [online]. Dostupné na: <https://esdac.jrc.ec.europa.eu/content/Soil_erosion_by_wind> [cit. 27. 10. 2021].
- [74] European Soil Data Centre (ESDAC), 2021. Soil erosion by water (RUSLE2015) [online]. Dostupné na: <<https://esdac.jrc.ec.europa.eu/content/soil-erosion-water-rusle2015>> [cit. 27. 10. 2021].
- [75] European Soil Data Centre (ESDAC), 2021. European Soil Database v.2 [online]. Dostupné na: <<https://esdac.jrc.ec.europa.eu/content/european-soil-database-v20-vector-and-attribute-data>> [cit. 27. 10. 2021].
- [76] Wageningen University & Research, 2021. ISRIC — World Soil Information [online]. Dostupné na: <<https://www.isric.org>> [cit. 27. 10. 2021].
- [77] PostgreSQL. 2021. PostgreSQL. [online] Dostupné na: <<https://www.postgresql.org/>> [cit. 25. 10. 2021].
- [78] Copernicus Land Monitoring Service. EU-DEM. [online]. Dostupné na: <<https://www.eea.europa.eu/data-and-maps/data/copernicus-land-monitoring-service-eu-dem>> [cit. 25. 10. 2021].
- [79] SDI4Apps, Pavel Vlach, 2015. Smart POI data set – SPOI [online]. Dostupné na: <<https://kgm.zcu.cz/spoi/>> [cit. 23. 10. 2021].
- [80] Čerba, O., Charvát, K., Mildorf, T., Bērziņš, R., Vlach, P., Musilová, B., 2016, SDI4Apps Points of Interest Knowledge Base. In: Progress in Cartography (pp. 229-237). Springer International Publishing
- [81] Čerba, O., 2016. SPOI – SDI4Apps: Points of Interest [online]. Dostupné na: <<https://kgm.zcu.cz/spoi/doc/SPOI.pdf>> [cit. 23. 10. 2021].
- [82] Harris, S., Seaborne, A., Prud'hommeaux, E., 2013. SPARQL 1.1 query language. [online] W3C recommendation, 21(10). Dostupné na: <<https://www.w3.org/TR/sparql11-query/>> [cit. 29. 10. 2021].
- [83] INSPIRE Hackathon 2017, 2017: 3D-OLU Pilot [online]. Dostupné na: <<http://ng.hslayers.org/examples/3d-olu>> [cit. 29. 10. 2021].
- [84] Schowengerdt, Robert A. (2007). Remote sensing: models and methods for image processing (3rd ed.). Academic Press. p. 2. ISBN 978-0-12-369407-2
- [85] Chuvieco, E., 2020. Fundamentals of Satellite Remote Sensing: An Environmental Approach, Third Edition (2nd ed.). CRC Press, Inc., Boca Raton, FL, USA, ISBN 9781138583832
- [86] Esbrí, M.Á., 2021. Remote Sensing. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. https://doi.org/10.1007/978-3-030-71069-9_4
- [87] National Aeronautics and Space Administration (NASA), 2021. Landsat Science [online]. Dostupné na: <https://landsat.gsfc.nasa.gov/?page_id=2281> [cit. 29. 10. 2021].
- [88] European Space Agency (ESA), 2018. Copernicus DIAS: Data and Information Access Services [online]. Dostupné na: <<https://eo4society.esa.int/2018/04/20/copernicus-dias-data-and-information-access-services/>> [cit. 29. 10. 2021].
- [89] Amazon inc., 2021. Amazon Web Service (AWS) [online]. Dostupné na: <<https://aws.amazon.com/earth/>> [cit. 29. 10. 2021].
- [90] Gomes, V.C.F., Queiroz, G.R., Ferreira, K.R., 2020. An Overview of Platforms for Big Earth Observation Data Management and Analysis. In: Remote Sensing 2020, 12, 1253
- [91] National Aeronautics and Space Administration (NASA), 2021. Earth Observing System Data and Information System (EOSDIS) [online]. Dostupné na: <<https://earthdata.nasa.gov/eosdis>> [cit. 29. 10. 2021].
- [92] Google Inc., 2021. Google Earth Engine [online]. Dostupné na: <<https://developers.google.com/earth-engine>> [cit. 29. 10. 2021].
- [93] Open Data Cube (ODC), 2021. An Open Source Geospatial Data Management & Analysis Platform – Open Data Cube [online]. Dostupné na: <<https://www.opendatacube.org/>> [cit. 29. 10. 2021].
- [94] Euro Data Cube Consortium, 2021. Euro Data Cube [online]. Dostupné na: <<https://eurodatacube.com/>> [cit. 29. 10. 2021].
- [95] Sinergise Laboratory for geographical information systems, Ltd., 2021. Sentinel Hub [online]. Dostupné na: <<https://www.sentinel-hub.com/>> [cit. 29. 10. 2021].

- [96] Maxar Technologies, 2021. The GBDX Platform [online]. Dostupné na: <<https://gbdxdocs.digitalglobe.com/docs/about-the-gbdx-platform>> [cit. 29. 10. 2021].
- [97] Planet Labs Inc., 2021. Planet Homepage [online]. Dostupné na: <<https://www.planet.com/>> [cit. 29. 10. 2021].
- [98] European Commission, 2021. European Open Science Cloud [online]. Dostupné na: <<https://digital-strategy.ec.europa.eu/en/policies/open-science-cloud>> [cit. 29. 10. 2021].
- [99] Kepka, Michal a kol. Využití senzorů v zemědělství. Zemědělský svaz ČR - Česká technologická platforma pro zemědělství, Praha. 2021.
- [100] Taylor, K., Griffith, C., Lefort, L., Gaire, R., Compton, M., Wark, T., Lamb, D., Falzon, G., Trotter, M., 2013. Farming the Web of Things. In: IEEE Intelligent Systems, vol. 28, no. 6, pp. 12-19, Nov.-Dec. 2013, doi: 10.1109/MIS.2013.102.
- [101] Rogotis S., Fournier F., Charvát K., Kepka M., 2021. Sensor Data. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. https://doi.org/10.1007/978-3-030-71069-9_3
- [102] European Citizen Science Association (ECSA), 2015. Deset principů občanské vědy [online]. Dostupné na: <https://ecsa.citizen-science.net/wp-content/uploads/2020/02/ecsa_ten_principles_of_cs_czech.pdf> [cit. 28. 10. 2021].
- [103] Irwin, A., 1995. Citizen Science: A Study of People, Expertise and Sustainable Development. Routledge. 212 Pages. ISBN 9780415130103.
- [104] Gura, T., 2013. Citizen science: amateur experts. In: Nature. 496 (7444): 259-261. doi:10.1038/nj7444-259a
- [105] Executive Agency for SMEs (EASME), 2016. Have you heard about the concept of Citizens' Observatories? [online]. Archivováno na: <<https://wayback.archive-it.org/12090/20170111094444/https://ec.europa.eu/easme/en/news/have-you-heard-about-concept-citizens-observatories>> [cit. 30. 10. 2021].
- [106] Newman, G., Wiggins, A., Crall, A., Graham, E., Newman, S., Crowston, K., 2012. The future of citizen science: emerging technologies and shifting paradigms. In: Frontiers in Ecology and the Environment, 10(6), 298-304.
- [107] Robinson, J. A., Kocman, D., Horvat, M., Bartonova, A., 2018. End-user feedback on a low-cost portable air quality sensor system—Are we there yet?. In: Sensors, 18(11), 3768
- [108] Charvat, K., Bye, B. L., Mildorf, T., Berre, A. J., Jedlicka, K., 2018. Open data, VGI and citizen observatories INSPIRE hackathon. In: International Journal of Spatial Data Infrastructures Research, 13, 109-130.
- [109] Goodchild, M.F., 2007. Citizens as sensors: the world of volunteered geography. In: GeoJournal 69 (4): 211-221.
- [110] Elwood, S., 2008. Volunteered Geographic Information: Future Research Directions Motivated by Critical, Participatory, and Feminist GIS. In: GeoJournal 72 (3&4): 173-183
- [111] Graham, M., 2010. Neogeography and the Palimpsests of Place. In: Tijdschrift voor Economische en Sociale Geografie 101 (4): 422-436
- [112] Charvát K., Kepka M., 2021. Crowdsourced Data. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. https://doi.org/10.1007/978-3-030-71069-9_5
- [113] Tableau Software, LLC., 2021. What Is Data Visualization? Definition, Examples, And Learning Resources [online]. Dostupné na: <<https://www.tableau.com/learn/articles/data-visualization>> [cit. 28. 10. 2021].
- [114] Rubens, P., 2017. Big Data Visualization [online]. Dostupné na: <<https://www.datamation.com/big-data/big-data-visualization/>> [cit. 30. 10. 2021].
- [115] Esbrí M.Á., Klien E., Charvát K., Zinke-Wehlmann C., Hitado J., Södergård C., 2021. Big Data Visualisation. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. https://doi.org/10.1007/978-3-030-71069-9_13
- [116] Jedlička, K., Charvát, K., 2018. Visualisation of Big Data in Agriculture and Rural Development. 2018 IST-Africa Week Conference (IST-Africa).
- [117] SmartAgriHubs, 2020. Flagship Innovation Experiment 20 Smart GroundWater and Weather Sensors project [online]. Dostupné na: <<https://www.smartagrihubs.eu/flagship-innovation-experiment/20-FIE-smart-ground-water-sensors>> [cit. 30. 10. 2021].
- [118] Ježek, J., Jedlička, K., Mildorf, T., Kellar, J., Beran, D., 2017. Design and evaluation of WebGL-based heat map visualization for big point data. In The Rise of Big Spatial Data (s. 13-26). Springer, Cham.
- [119] Stijn, D., Watson, R., 2006. Extending google docs to collaborate on research papers. Toowoomba, Queensland, AU: The University of Southern Queensland, Australia 23 (2006).
- [120] Sarah Hsueh-Jui, L., Lan, Y., 2016. Social constructivist approach to web-based EFL learning: Collaboration, motivation, and perception on the use of Google docs. In: Educational Technology & Society 19.1 (2016): 171-186.

- [121] Young-Wook, J., Lim, Y., Kim, M., 2017. Possibilities and limitations of online document tools for design collaboration: The case of Google Docs. In: Proceedings of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing. 2017.
- [122] Charvát, K., Bergheim, R., Bērziņš, R., Zadražil, F., Langovskis, D., Vrobel, J., Horakova, S., 2021. Map Whiteboard Cloud Solution for Collaborative Editing of Geographic Information. In: Cloud Computing and Data Science, 36-55.
- [123] Järvinen, P., Siltanen, P., & Kirschenbaum, A. (2021). Data Analytics and Machine Learning. In Big Data in Bioeconomy (pp. 129-146). Springer, Cham.
- [124] Alpaydin, E. 2020. Introduction to machine learning. In: MIT press, 2020, Cambridge, MA, USA.
- [125] Wong, P.C., 1999. Guest editor's introduction: Visual data mining. In: IEEE Computer Graphics and Applications, vol. 19, č. 5, s. 20-21.
- [126] Ferreira de Oliveira, M.C., Levkowitz, H., 2003. From visual data exploration to visual data mining: A survey. In: Visualization and Computer Graphics, IEEE Transactions on, vol. 9, č. 3, s. 378-394.
- [127] Amar, R., Eagan, J., Stasko, J., 2005. Low-level components of analytic activity in information visualization. In IEEE Symposium on Information Visualization, 2005. INFOVIS 2005. (s. 111-117). IEEE.
- [128] Arel, I., Rose, D. C., Karnowski, T. P., 2010. Deep machine learning-a new frontier in artificial intelligence research. In: IEEE computational intelligence magazine, 5(4), 13-18.
- [129] OGC, 2019, OGC earth observation applications pilot: call for participation [online], Dostupné na <https://portal.ogc.org/files/?artifact_id=90733> [cit. 7. 11. 2021].
- [130] DataBio, 2018. DataBio D1.1 – Agriculture Pilot Definition. Various. 30.06.2018. Dostupné na: <https://www.databio.eu/wp-content/uploads/2017/05/DataBio_D1.1-Agriculture-Pilot-Definition_v1.1_2018-04-26_LESPRO.pdf> [cit. 27.10.2021].
- [131] Drake, M., 2019. A Comparison of NoSQL Database Management Systems and Models [online]. , Dostupné na: <<https://www.digitalocean.com/community/tutorials/a-comparison-of-nosql-database-management-systems-and-models>> [cit. 7. 11. 2021].
- [132] Charvát K., Lukas V., Charvát K., Horáková Š., 2021. Delineation of Management Zones Using Satellite Imageries. In: Södergård C., Mildorf T., Habyarimana E., Berre A.J., Fernandes J.A., Zinke-Wehlmann C. (eds) Big Data in Bioeconomy. Springer, Cham. https://doi.org/10.1007/978-3-030-71069-9_18
- [133] Mezera, J., Lukas, V., Elbl, J., Kintl, A., Smutný, V., 2019. Evaluation of variable rate application of fertilizers by proximal crop sensing and yield mapping. 19th International Multidisciplinary Scientific Geoconference, SGEM 2019.
- [134] Euxdat.eu, 2021. About EUXDAT. [online]. Dostupné na: <<https://www.euxdat.eu/about-euxdat/>> [cit. 27.10.2021].
- [135] Euxdat.eu, 2021. Objectives – EUXDAT. [online]. Dostupné na: <<https://www.euxdat.eu/objectives/>> [cit. 27.10.2021].
- [136] Nieto, F. J. , 2017. European e-Infrastructure for Extreme Data Analytics in Sustainable Development (EUXDAT) [online]. Project proposal. Dostupný na: <https://cordis.europa.eu/project/id/777549>. [cit. 27. 10. 2021]
- [137] EUXDAT, 2020, EUXDAT D5.6 Pilots Implementation v2. Various. Dostupné na: <https://www.euxdat.eu/wp-content/uploads/2020/10/EUXDAT_D5.6-Pilots-Implementation-v2_FINAL.pdf> [cit. 27.10.2021].
- [138] Doolin, K., 2019. Digitising and transforming European industry and services: digital innovation hubs and platforms (DEMETER). Grant Agreement No. 857202.
- [139] DEMETER, 2021. DEMETER Project web page [online]. Dostupné na: <<https://h2020-demeter.eu/>> [cit. 27.10.2021].
- [140] DEMETER, 2021. DEMETER PILOT 2.3 Data Brokerage Service and Decision Support System for Farm Management [online]. Dostupné na: <<https://h2020-demeter.eu/pilots-overview/pilot-cluster-two/data-brokerage-service-and-decision-support-system-for-farm-management/>> [cit. 27.10.2021].
- [141] InnoVar, 2021. INNOVAR Project home page [online]. Dostupné na: <<https://www.h2020innovar.eu/>> [cit. 27.10.2021].
- [142] Plan4All z.s., 2021. AgriHub made in CzechoSlovakia (AgriHub CZ&SK) [online]. Dostupné na: <<https://www.plan4all.eu/projects/agrihub-czsk/>> [cit. 27.10.2021].



VELKÁ DATA V ZEMĚDĚLSTVÍ

Autoři:

RNDr. Karel Charvát	(Plan4all z.s., WIRELESSINFO)
Ing. Pavel Hájek, Ph.D.	(Plan4all z.s., Západočeská univerzita v Plzni)
Ing. Michal Kepka, Ph.D.	(Plan4all z.s., Západočeská univerzita v Plzni)
Ing. Martina Kepka Vichrová, Ph.D.	(Plan4all z.s., Západočeská univerzita v Plzni)
Bc. Dmitrij Kozhukh	(Plan4all z.s.)
Mgr. Jiří Kvapil	(Plan4all z.s., LESPROJEKT-SLUŽBY s.r.o.)
Ing. Hana Kubíčková	(Plan4all z.s., WIRELESSINFO)
Ing. Petr Uhlíř	(Plan4all z.s., LESPROJEKT-SLUŽBY s.r.o.)
Dr. Ing. Jaroslav Šmejkal	(Plan4all z.s., LESPROJEKT-SLUŽBY s.r.o.)
Ing. Antonín Kubíček	(Plan4all z.s., LESPROJEKT-SLUŽBY s.r.o.)
Bc. Heřman Šnevajs	(Plan4all z.s., WIRELESSINFO)
Stein Runar Bergheim, MSc.	(Plan4all z.s., Asplan Viak Digital Services)
Ing. Karel Charvát, Jr.	(Plan4all z.s., LESPROJEKT-SLUŽBY s.r.o.)
Ing. Šárka Horáková	(Plan4all z.s., WIRELESSINFO)

Vydavatel:

Zemědělský svaz ČR - Česká technologická platforma pro zemědělství

Grafika:

Pavla Brus Ortová

Tiskárna:

SYNERGIE: 4U s.r.o.

Vydání: první

Rok vydání: 2021

Náklad: 1000 výtisků

ISBN: 978-80-905151-5-4 (tisk)
978-80-905151-6-1 (pdf)

Za obsahovou a jazykovou správnost díla odpovídá autor.

