



Evropská unie
Evropský sociální fond
Operační program Zaměstnanost

Metodický materiál pro zavádění a využívání nástrojů a postupů pro mapování, analýzu a predikci kriminality včetně doporučeného vzdělávání uživatelů

Obsah:

1. Úvod.....	3
2. Situace před realizací projektu.....	4
3. Vývoj funkce predikce	4
4. Metodická doporučení získané praktickou implementací funkce predikce	5
5. Využívání funkce predikce v praxi PČR	5
6. Vzdělávání	7

Přílohy:

Predikce kriminality v rámci Policie ČR – neveřejný dokument

1. Úvod

Tento dokument byl vytvořen v rámci projektu „Mapy budoucnosti II – využití prostorových dat pro vytvoření a pilotní ověření nástrojů a postupů pro analýzu a predikci kriminality za účelem jejího předcházení a potírání“ (registrační číslo CZ.03.4.74/0.0/0.0/15_025/0010253), který je financován z operačního programu Zaměstnanost a státního rozpočtu (dále jen „projekt“).

Z pohledu zadání projektu byla problematika predikčních technologií pojata především prakticky, tzn. pro potřebu optimalizace hlídkové služby v pilotních lokalitách – tato inovační aktivita, která odpovídá definici aplikovaného výzkumu, byla jedním ze zásadních technických úkolů projektu. Výkonná skupina projektu splnila zadání i v tomto ohledu a technologií predikce kriminality, jež využívá strojového učení (dále jen „funkce predikce“), zavedla do praxe ve všech 6 pilotních lokalitách. Funkce predikce je dostupná v intranetu Policie ČR v prostředí Informačního systému MAPA¹ (dále jen IS Mapa) kde ji mohou vyzkoušet všichni uživatelé intranetu.

V pilotních lokalitách byly ustanoveny tzv. lokální týmy, které měly za úkol testovat a připomínkovat funkci predikce. Díky takto nastavené spolupráci byly do funkce predikce implementovány následující požadavky:

- hranice pro testování v rámci Prahy bude zohledňovat pouze lokalitu OŘ Praha II,
- možnost predikovat na denní i noční směnu separátně,
- predikovat na více dnů dopředu,
- predikovat pouze trestné činy nebo trestné činy včetně přestupkového jednání.

Pro každou pilotní obec/oblast bylo nasazeno separátní strojové učení využívající technologii („boostovaných“) rozhodovacích stromů, při kterých se stroj učil na historických datech kriminality a dalších souvisejících datových sadách (např. historická data o počasí, segmentace částí obcí dle využití území a podobně). Tímto postupem vznikla pro každou lokalitu sada zájmových modelů, které jsou využívány pro každodenní výpočet predikce kriminality z aktuálních dat o výskytu kriminality a predikovaného počasí.

Zásadní komplikací byla při těchto často zdlouhavých výpočtech epidemická situace v souvislosti s výskytem koronaviru SARS-CoV-2 v populaci České republiky. Kriminalita se díky tomu vyvíjela anomálně a funkce predikce následně nefungovala dle poznatků zástupců lokálních týmů dobře. Řešením bylo období od počátku výskytu koronaviru ve společnosti ze strojového učení vyloučit. Ve stručnosti lze konstatovat, že stroj se naučil na datech před výskytem koronaviru, kdy lze chování společnosti označit za „normální“. Takto vygenerované modely jsou využívány při každodenních výpočtech funkce predikce na další dny. Období s téměř nulovou kriminalitou tak nebylo strojem posuzováno jako dlouhodobý trend.

Výkonná skupina projektu se od počátku projektu do Q3 2022 soustředila primárně na optimalizaci této technologie, což bude probíhat i po ukončení projektu Q4 2022–2023. Hlavním účelem takto koncipovaného záměru rozvoje je vyvinout funkční globální model funkce predikce využitelný na celém území ČR. Ten by měl v budoucnu sloužit k automatickému navrhování kontrolních bodů v nové verzi instruktážních programových prostředků.

¹ Rozkaz policejního prezidenta 79/2016, kterým se upravuje zkušební provoz informačního systému MAPA

2. Situace před realizací projektu

V rámci vnitřní sítě Policie ČR již před realizací projektu byl provozován Informační systém Mapa (dále jen „IS Mapa“, zachycen na obrázku 1) zobrazující nápad trestné činnosti a přestupků registrovaných Policií ČR. IS Mapa je určen pro podporu výkonu policejních činností a není dostupný mimo vnitřní síť Policie ČR či zobrazovací zařízení Policie ČR využívaných v terénu. IS Mapa zobrazuje nápad trestné činnosti v reálném čase a dále poskytuje uživatelům další funkce potřebné k výkonu jejich práce např. dopravní informace, měření vzdáleností, ploch. Umožňuje také přesměrování do dalších zájmových systémů Policie ČR nebo např. do katastru nemovitostí. Zásadním aspektem pro zvolení tohoto informačního systému jakožto publikační platformy výsledků projektu v rámci Policie ČR v oblasti predikce kriminality je především jeho rozšířenost a jeho popularita. Tento informační systém, který v minulosti vyvíjela podstatná část technického týmu projektu, denně využije více jak 3000 uživatelů policejní vnitřní sítě. IS Mapa je uživatelům vnitřní sítě dostupný na konstantní a známé adrese.

3. Vývoj funkce predikce

Pro stávající IS Mapa zobrazující výskyt registrovaných skutků Policií ČR byla v rámci projektu vyvinuta a zprovozněna funkce predikce² (nový nástroj označený jako „predikce“). Funkce byla vytvořena na základě širokého spektra prostorových dat nashromážděných Policií ČR, které byly dále zpracovány (extrakce, strojové učení) a využívány pro výpočty modelů predikce v jednotlivých městech (dále jen „pilotní lokality“). Tento nástroj uživateli v pilotních lokalitách (Praha, Turnov, Liberec, Jablonec nad Nisou, Kladno, Kolín) predikuje buňky/lokality (čtverec o hraně 100 metrů) předem stanovené sítě, ve kterých je pro následující den nejvyšší riziko majetkového trestného činu. Tyto buňky jsou každý den brzy ráno (5:00am) počítány z aktuálních dat na základě modelu, který byl před spuštěním pilotního provozu vytvořen pomocí technologie strojového učení.

Funkce predikce byla vytvořena jako algoritmus nad méně či více historickými daty za minulá období, který nevytváří programátor ani statistik. Pro stávající výpočty jsou použité tzv. boostované rozhodovací stromy (knihovna XGBoost). Pilotní lokalita je tedy vždy rozdělena na čtverce, které jsou popsány sadou signálů. To vytváří sadu pravidel pro model rozhodovacího stromu, které umožní vyhodnocovat další podmínky a určovat pravděpodobnost výskytu deliktu. Učení (tvorba pravidel a rozhodování) může být v závislosti na výkonu užitých HW zdrojů poměrně časově náročné, neboť se musí při tvorbě pravidel opakovaně procházet velká historická data. Strojové učení bylo po dobu celého projektu spouštěno manuálně na základě rozboru dosavadních výsledků takto produkovaných modelů. Vlastní predikce, která je založena na využití modelů (sady pravidel) pro daný den či nejbližší budoucnost je již relativně rychlá.

Pro výpočty jsou využívána data z databáze Policie ČR. Jedná se o kriminální signály (krádeže vloupáním, krádeže, jiná majetková trestná činnost, přestupky proti majetku podle paragrafu 50) včetně informace o době a místě, kdy a kde se delikt stal. Dále se využívají geografické signály (např. silnice, různé budovy – např. hotely, restaurace, herny) a ostatní signály (např. pracovní den, svátek, počasí, den v týdnu). Model dokáže po naučení určit důležitost jednotlivých signálů použitých k učení. Díky tomu můžeme některé signály vyřadit a jiným se věnovat a zpřesňovat je. Podrobnosti k doposud využitým postupům v rámci projektu jsou uvedeny v příloze³.

² Funkce predikce není sdílena s veřejností a je určena pouze pro vnitřní využití v rámci Policie ČR.

³ Obsah přílohy nebude sdílen s veřejností a je určen pouze pro vnitřní využití v rámci Policie ČR.

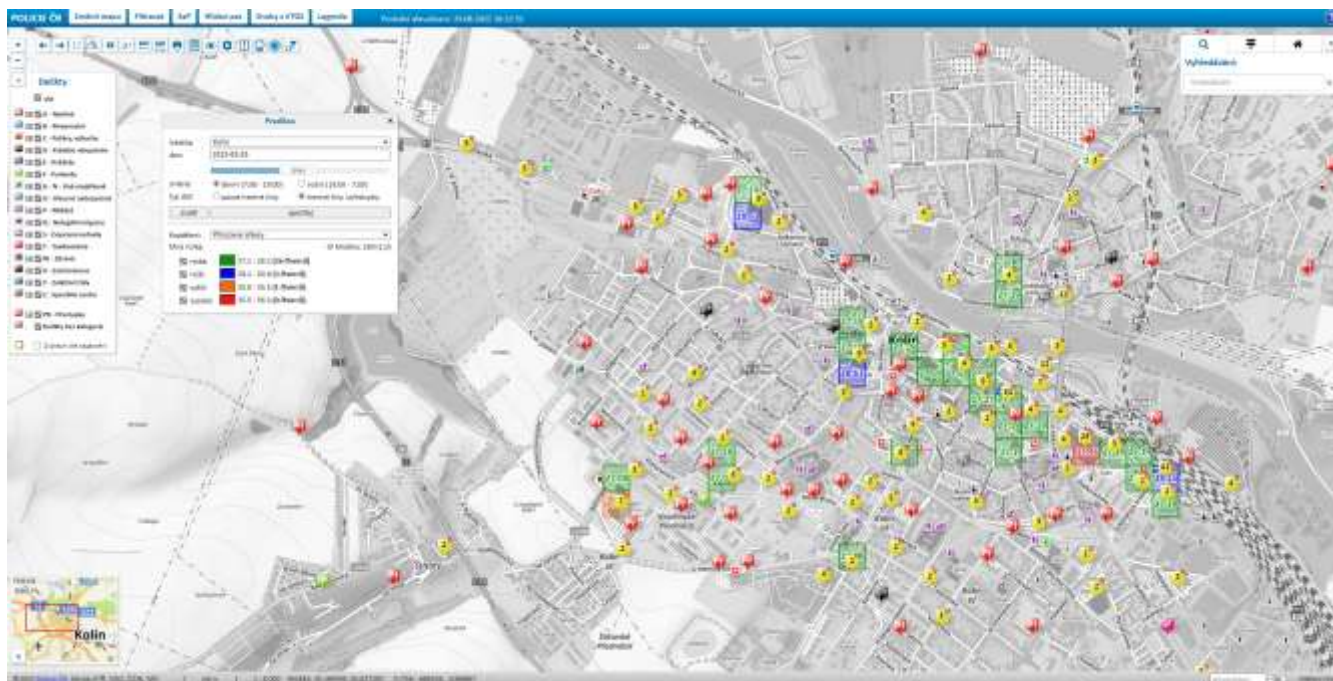
4. Metodická doporučení získané praktickou implementací funkce predikce

Obecně lze formulovat následující metodická doporučení:

- Úspěch a možnosti využití technologií při realizaci funkce predikce je přímo úměrný podrobnosti a kvalitě dat na vstupu. Je vhodné poskytnout pro potřeby strojového učení data v maximálním možném objemu a podrobnosti.
- Zdroje možných vstupů je nutné kontinuálně vyhodnocovat a rozvíjet.
- Řízení přístupu ke vstupním či výstupním informacím funkce predikce by mělo vždy být v moci organizace. Organizace by měla mít k dispozici vždy personální zdroje, které znají a přímo ovlivňují detailní architekturu celého řešení.
- Data pro strojové učení omezovat na období, kdy fungování společnosti není omezeno restrikcemi v oblasti volného pohybu, produkce či využití služeb.
- Vývoj a implementace technologií funkce predikce v prostředí kterékoli organizace představuje nezanedbatelné nároky na zdroje (lidské, technické → finanční). Je velmi důležité, aby tyto aktivity vždy směřovaly k alespoň částečné návratnosti, tzn. účel a způsob praktického využití byl přesně definován. V případě Policie ČR bylo stanoveným využitím optimalizace hlídkové činnosti.
- Prostorová data o bodech zájmu jsou pro predikci kriminality velmi důležitá i přesto, že nejsou součástí státem garantovaných datových sad.
- Pracovní postupy v technické rovině pravidelně podrobovat supervizi výzkumným subjektům či autoritám, které nejsou do projektu přímo zapojeny.
- Navyšovat povědomí a vzdělanost odborné komunity uživatelů o možnostech a využití nejen funkce predikce.

5. Využívání funkce predikce v praxi PČR

Současné výsledky projektu jsou zpřístupněny v interním systému Policie ČR pro oblast pilotních lokalit. Funkce predikce je sdílена s uživateli z pilotních lokalit i v rámci obecní sekce map kriminality pro veřejnost. Obě tyto situace jsou dokumentovány na následujících snímcích.



Obrázek 1 - ukázka využití funkce predikce ve vnitřní síti Policie ČR



Obrázek 2 - ukázka využití funkce predikce v obecních sekci map kriminality pro veřejnost

Zásadním požadavkem z praktického hlediska byla maximální jednoduchost ovládání pro uživatele funkce predikce. Uživatel by neměl, popřípadě nemusí znát složitost technologie, již využívá. Pro svou práci v přímém výkonu služby, to nepotřebuje. Samotná skutečnost, že uživatel prostým využitím funkce predikce v mapě v jeden okamžik využívá až 16 modelů (4 situace x 4 dny) vygenerovaných na základě obrovského množství dat a velmi výkonného HW bez toho aniž by musel disponovat specifickými znalostmi, lze označit za úspěšný předpoklad pro další využití této technologie v praxi Policie ČR. Pro zobrazení nejrizikovějších míst v mapě v rámci pilotní lokality stačí nastavit (nebo ponechat v defaultním nastavení) 5 parametrů.

Vzhledem k plánovanému využití, jímž je optimalizace hlídkové činnosti Policie ČR v rámci přímého výkonu služby, příkládá Policie ČR i v současné praxi velký důraz na oprávněný požadavek ohledně sdílení těchto dat se zájmovými partnery na úseku vnitřní bezpečnosti a veřejného pořádku. Z hlediska metodického doporučení, neexistuje žádný relevantní důvod, proč by Policie ČR neměla výsledná prostorová data anonymizovaných čtverců rizikovitosti sdílet se svými partnery (obecní policie) v zájmových lokalitách. Taktické hledisko tento metodický postup jen utvrzuje.

Dalším velmi důležitým metodickým poznatkem získaným současným testováním funkce predikce v přímém výkonu služby je ještě více zvýšit úsilí při dohledávání a následné konsolidaci dat bodů zájmu a jejich zdrojů. Praktické výsledky ukazují, že přítomnost restauračního zařízení, či zastávky MHD v blízkém okolí predikovaného místa kriminality je často jedním z nejzásadnějších signálů využívaných při rozhodování na základě výsledků strojového učení u funkce predikce (viz příloha tohoto dokumentu). Hypotéza byla funkcí predikce prakticky ověřena. Situace v této oblasti, není nijak řešena ze strany kompetentních institucí. Z tohoto důvodu je nutné provádět činnost sběru a následné správy přímo v rámci Policie ČR a to i přesto, že tato organizace by měla být pouze konzumentem bezplatných a garantovaných prostorových dat o bodech zájmu na území ČR a okolních států.

Současná praxe

Pro budoucí praxi Policie ČR, vzbuzuje funkce predikce zásadní očekávání z pohledu zjednodušení plánování hlídkové činnosti. Výstupy projektu jsou nyní implementovány do nově vyvíjeného IS Instruktaž, kde budou automaticky generované rizikové čtverce v příslušné lokalitě sloužit jako kontrolní body hlídek.

6. Vzdělávání

Funkce predikce bude v rámci ČR (rozuměj u Policie ČR a obecní policie) využívána v míře přímo úměrné vynaloženému úsilí na vzdělávání a osvětu v této oblasti. I přesto, že samotné použití funkce predikce v zájmových aplikacích je uživatelsky velmi jednoduché, musí uživatel mít alespoň obecné povědomí o možnostech a limitech tohoto nástroje. Z tohoto důvodu byl v rámci projektu navržen, realizován a následně i optimalizován naprosto konkrétní vzdělávací rámec pro uživatele z řad Policie ČR i partnerských organizací.

Výsledkem tak byla sekce přímo zaměřená na využití funkce predikce v rámci prezenčně navštěvovaných jednodenních kurzů, které byly v rámci projektu realizovány. V průběhu této aktivity projektu tak byla s více než 160 konkrétními uživateli probrána problematika predikce kriminality, která je díky projektu v současnosti v rámci ČR dostupná zainteresovaným pracovníkům zabývajících se prevencí, odhalováním či snižováním důsledků kriminálního jednání (dále jen „odborná komunita uživatelů“) ve vybraných lokalitách. Uživatelé byli seznámeni s tím, jak a kde mohou tuto inovativní technologii v rámci své práce využít. I po ukončení projektu budou probíhat další takto orientovaná školení v rámci Policie ČR, která budou začleněna do rozličných vzdělávacích kurzů.

Současně možný způsob využití dosavadních výsledků funkce predikce je podrobně dokumentován v interaktivním e-learningovém kurzu vytvořeném v rámci projektu⁴. E-learningový kurz popisuje funkci predikce a její využití, součástí je dokonce i verbálně komentovaný videotutoriál, který velmi podrobně popisuje možné způsoby využití této technologie v praxi Policie ČR.

Dlouhodobá a stálá uživatelská znalost možností a přínosů těchto technologií (dále jen „znalost“) u odborné komunity uživatelů však nebude zabezpečena bez prosazení do praxe přímého výkonu služby Policie ČR a v rámci odborné komunity uživatelů. Tato cílová skupina uživatelů nebude ochotna měnit své pracovní návyky, dokud nebude přesvědčena o tom, že jí zamýšlená změna přinese benefity (úsporu času, úsporu zdrojů). Dosavadní zkušenosti, zpětná vazba a výsledky evaluace ukazují, že cílová skupina je však schopna nové technologie využívat i bez direktivních nařízení, pokud je dostatečně informovaná a znalá. Přímé a pravidelné využití funkce predikce je přímo úměrné povědomí o těchto možnostech na základních útvarech Policie ČR.

Pro efektivní zacílení dlouhodobých vzdělávacích aktivit v rámci organizace je doporučeno:

- Vyčlenit z dostupných stávajících personálních kapacit Policie ČR pracovníky na vzdělávání v oblasti geografického zabezpečení. V tomto směru doplnit, popřípadě upravit kompetentním pracovníkům náplň práce.
- Zařadit do vzdělávacího procesu Policie ČR povinné krátké presenční vzdělávací bloky zaměřené na problematiku přímé geografické podpory rozhodování následujícím způsobem:
 - Základní odborná příprava příslušníků Policie ČR
(1 vyučující den, tj. 8 vyučujících hodin), kompletní náplň školení s důrazem na základní orientaci ve službách Geografického Informačního Systému Policie ČR (dále jen „GIS PČR“) a zjednodušení pracovních postupů.
 - Vzdělávací kurzy určené pro vedoucí obvodních oddělení
(1/2 vyučujícího dne, tj. 4 vyučující hodiny), základní náplň školení s důrazem na aplikace poskytované GIS PČR využitelné v přímém výkonu služby.
 - Vzdělávací kurzy určené pro velitele bezpečnostních opatření
(1/2 vyučujícího dne, tj. 4 vyučující hodiny), základní náplň školení s důrazem na efektivní přípravu na bezpečnostní opatření.
 - Vzdělávací kurzy určené pro analytické pracovníky Policie ČR

⁴ E-learningové kurzy projektu <https://evaluace.upvsp.cz/>

(1 vyučující den, tj. 8 vyučujících hodin), kompletní náplň školení s důrazem na analytické propojení služeb a mapového obsahu.

- Zařadit do vzdělávacího procesu Policie ČR prakticky zaměřené e-learningové kurzy/Sylabusy poskytující příklady a cvičení k výše uvedeným presenčním kurzům.
- Pravidelně vzdělávat zástupce odborné komunity uživatelů z jiných organizací (např. městské policie).
- Připravit a publikovat prakticky zaměřené e-learningové kurzy/Sylabusy poskytující příklady a cvičení k výše uvedeným presenčním kurzům pro zástupce z jiných organizací.
- Zvýšit úroveň informovanosti odborné komunity například častější účastí pracovníků Policie ČR v odborných diskuzích, webinářích či konferencích v oblasti geoinformatiky, big data, predikčních analýz a strojového učení.

Návrh obsahu vzdělávání zaměřeného na problematiku přímé geografické podpory rozhodování v rámci Policie ČR:

ZÁKLADNÍ ODBORNÁ PŘÍPRAVA PŘÍSLUŠNÍKŮ POLICIE ČR

(1 vyučující den, tj. 8 vyučujících hodin), kompletní náplň školení s důrazem na základní orientaci ve službách GIS PČR a zjednodušení pracovních postupů:

- Úvod do GIS u Policie ČR (historie GIS PČR – vývoj pracoviště a služeb).
- IS čerpající data GIS – popis dostupných služeb (základní policejní mapa, obsah mapové vrstvy, tencí klienti GEO a MAPA, vektorová policejní mapa).
- Rozcestník GIS PČR.
- Zdrojová data jak v rámci PČR tak i mimo organizaci (RÚIAN, ČÚZK, MO a další), aktuálnost dat, propojení s policejními databázemi.
- Mapy kriminality – nastavení a popis spolupráce s obcemi a městskými částmi, propojení databází přestupků. Mapa kriminality pro veřejnost, mapa kriminality pro obce.
- Predikce kriminality – popis služby, metodika základní práce a ovládání modulu.
- Mapové produkty a aplikace GIS PČR ukázka základních úloh GIS – porovnání postupů pro pokročilého uživatele X zjednodušené postupy v dostupných aplikacích v rámci intranetu PČR).
- Mobilní bezpečná platforma.
- Další aplikace GIS.

VZDĚLÁVACÍ KURZY URČENÉ PRO VEDOUCÍ OBVODNÍCH ODDĚLENÍ

(1/2 vyučujícího dne, tj. 4 vyučující hodiny), základní náplň školení s důrazem na aplikace poskytované GIS PČR využitelné v přímém výkonu služby:

- Úvod do GIS u Policie ČR (základní představení).
- Informační systémy čerpající data GIS – popis dostupných služeb (základní policejní mapa, obsah mapové vrstvy, tencí klienti GEO a MAPA, vektorová policejní mapa).
- Rozcestník GIS PČR.
- Mapy kriminality pro veřejnost/obce – nastavení a popis spolupráce s obcemi a městskými částmi, propojení databází přestupků.
- Predikce kriminality – základní popis služby, metodika základní práce a ovládání modulu, využití při plánování hlídkové služby.

- Mobilní bezpečná platforma.

VZDĚLÁVACÍ KURZY URČENÉ PRO VELITELE BEZPEČNOSTNÍCH OPATŘENÍ

(1/2 vyučujícího dne, tj. 4 vyučující hodiny), základní náplň školení s důrazem na efektivní přípravu na bezpečnostní opatření:

- Úvod do GIS u Policie ČR (základní představení).
- IS čerpající data GIS – popis dostupných služeb (základní policejní mapa, obsah mapové vrstvy, tencí klienti GEO a MAPA, vektorová policejní mapa).
- Rozcestník GIS PČR.
- Aplikace GEO dostupná – seznámení s mapovými službami a jejich možnostmi, přidávání služeb, problematika vytváření a požadavky na vstupní data.
- Predikce kriminality – popis služby, základní metodika práce a ovládání modulu, využití při plánování bezpečnostních opatření.
- Mobilní bezpečná platforma.

VZDĚLÁVACÍ KURZY URČENÉ PRO ANALYTICKÉ PRACOVNÍKY POLICIE ČR

(1 vyučující den, tj. 8 vyučujících hodin), kompletní náplň školení s důrazem na analytické propojení služeb a mapového obsahu:

- Úvod do GIS u PČR (základní představení a služby).
- Úvod do desktopové aplikace Quantum GIS (dále jen QGIS) – základní analytické úlohy.
- Informační systémy čerpající data GIS – popis dostupných služeb (základní policejní mapa, obsah mapové vrstvy, tencí klienti GEO a MAPA, vektorová policejní mapa).
- Rozcestník GIS PČR.
- Zdrojová data jak v rámci Policie ČR, tak i mimo organizaci (RÚIAN, ČÚZK), aktuálnost dat, propojení s policejními databázemi.
- Mapy kriminality – nastavení a popis spolupráce s obcemi a městskými částmi, propojení databází přestupků. Mapa kriminality pro veřejnost, mapa kriminality pro obce – základní informace.
- Predikce kriminality – popis služby, základní metodika práce a ovládání modulu.
- Mapové produkty a aplikace GIS PČR ukázka základních úloh GIS – porovnání postupů pro pokročilého uživatele (QGIS) X zjednodušené postupy v dostupných aplikacích v rámci intranetu PČR).
- Mobilní bezpečná platforma – základní informace.
- Další aplikace GIS.