

OCCAPI – ODPRTA INTELIGENTNA KOMUNIKACIJSKA PLATFORMA

Marko Bajec¹, Dejan Lavbič¹, Tomaž Vidonja², Andrej Kos³

¹Fakulteta za računalništvo in informatiko, Laboratorij za podatkovne tehnologije

{marko.bajec, [dejan.lavbic](mailto:dejan.lavbic@fri.uni-lj.si)}@fri.uni-lj.si

²Zavod ICT, Kompetenčni center KC OPCOMM

tomaz.vidonja@ict-slovenia.net

³Fakulteta za elektrotehniko, Laboratorij za telekomunikacije

andrej.kos@fe.uni-lj.si

Abstract

OPEN INTELLIGENT COMMUNICATION PLATFORM

In this presentation we will describe a high-level architecture of an open intelligent communication platform that is being developed within the OPCOMM competence centre. The purpose of the platform is twofold: (a) to support fusion of large amounts of data, irrespective of their source or structure (e.g. sensors, intelligent devices, metering systems, cameras, web sources,...) and (b) to provide users or devices with semantically analysed and enriched data according to their needs and context. The platform will implement open interfaces to gain interoperability and will have security and access control mechanisms built into the system. Using the platform users will be able to access semantically enriched data and receive warnings about specific circumstances or events that will be recognised by the system. The platform itself can serve as a support for the realisation of advanced event and/or data-oriented services, including for example the analysis and optimisation of services and networks, and the identification of contextual data about user's presence. Based on its modularity, the platform can be extended to support analysis of large amounts of data of any kind and purpose.

Ključne besede

Komunikacijska platforma, inteligentna platforma, odprti vmesniki, inteligentna analiza podatkov, fuzija podatkov

Key words

Communication Platform; Intelligent Platform; Open Interfaces, Intelligent Data Analysis; Data Fusion

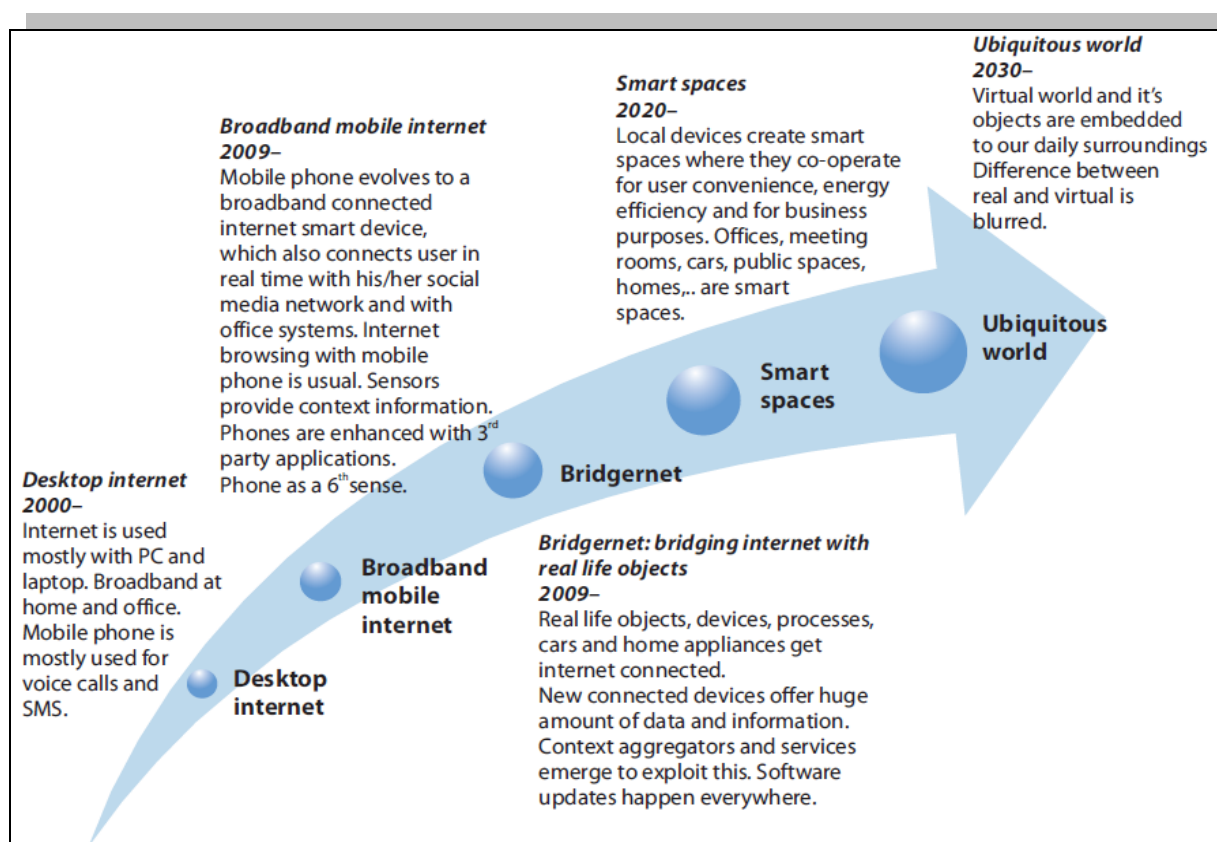
1. KOMPETENČNI CENTER OPCOMM

Globalni trendi na področju IKT prehajajo iz splošnih principov v mnogo bolj fokusirane in uporabniško naravnane. Glede na stanje in trende pa v tem trenutku v ospredje stopajo specifične tehnologije, ki bodo temelj za izdelke in storitve v obdobju med letoma 2015 in 2020 (slika 1).

Kompetenčni center OPCOMM (odprta komunikacijska platforma za integracijo storitev, v nadaljevanju KC OPCOMM) sestavlja konzorcij partnerjev, ki skupaj zagotavljamo kritično maso znanja za razvoj novih tehnologij in izdelkov, storitev in procesov na področju naprednih uporabniških platform in vmesnikov.

KC OPCOMM ima za cilj razviti znanja, tehnologije in procese, ki bodo ponudili rešitve za tri ključne izzive in probleme družbe:

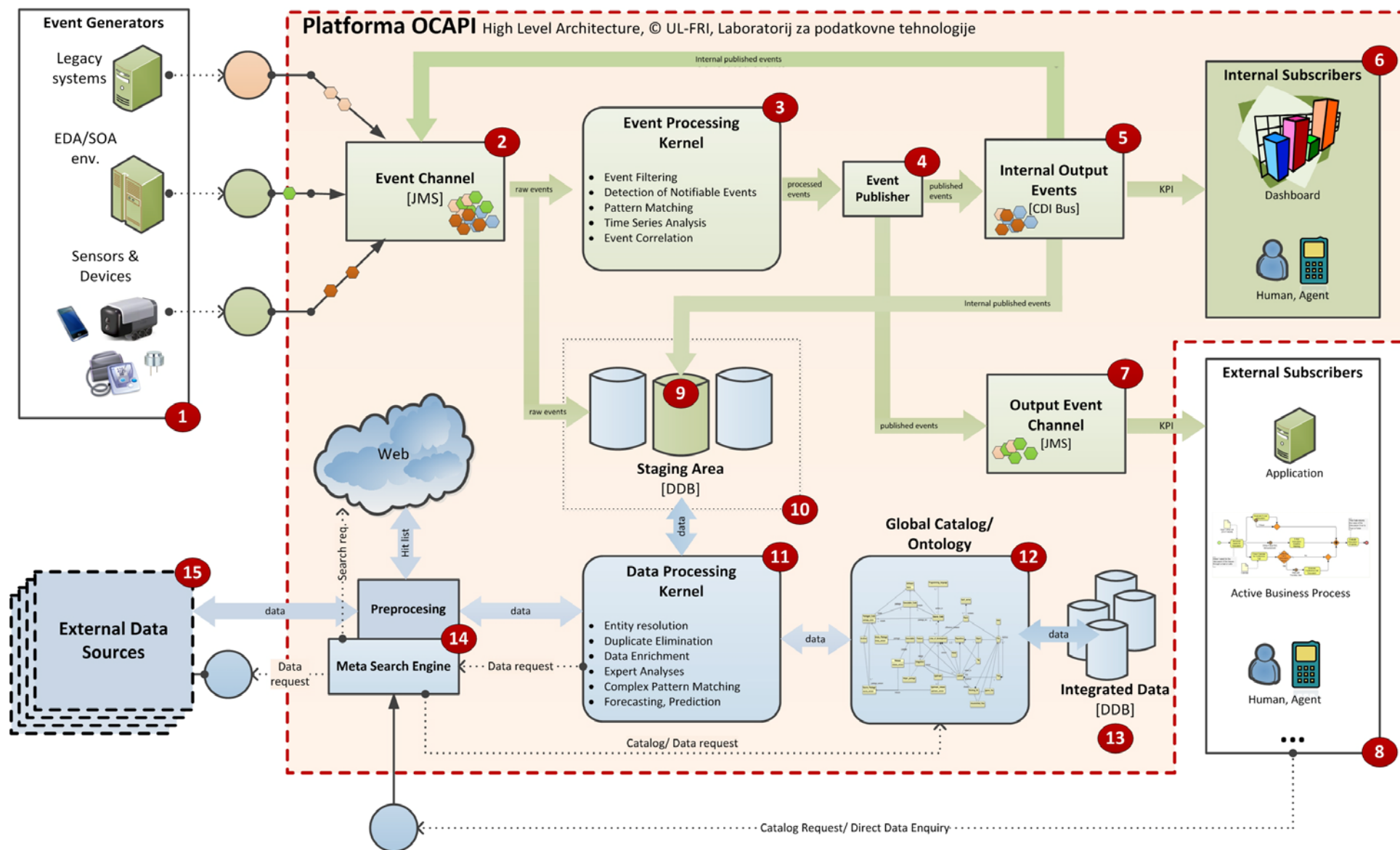
- Prepočasno širjenje širokopasovnega dostopa in interneta in s tem posledično zaostajanje pri izkoriščanju prednosti, ki jih ponuja Internet bodočnosti (ang. Future Internet), kar je posledica predrage opreme in neekonomičnih poslovnih modelov, ki so v uporabi pri gradnji omrežij.
- Preobilica podatkov, informacij ter vsebin in nezmožnost njihove učinkovite uporabe v celovitih aplikacijah za posamezna področja življenja (pametna mesta, pametna logistika, pametno zdravstvo, pametna energija, pametne zgradbe, pametni dom,...).
- Porast števila pametnih naprav. Do leta 2020 jih bo v Internet povezanih več deset milijard in komunikacija med samimi napravami bo predstavljala preko 95% vsega prometa in pretoka informacij po sodobnih komunikacijskih omrežjih.



Slika 1: Evolucija IKT (Vir: Roadmap for Communication Technologies, Services and Business Models 2010, 2015 and Beyond, Tekes Review, 275/2010).

2. INTELIGENTNA KOMUNIKACIJSKA PLATFORMA

Eden od pomembnejših ciljev KC-OPCOMM je razvoj inteligentne komunikacijske platforme, ki jo označujemo s kratico *Occapi*. Osnovne elemente platforme Occapi smo prototipno že razvili in demonstrirali njeno uporabnost na različnih scenarijih. Specifični elementi pa so še v razvoju. V nadaljevanju predstavljamo visoko-nivojsko arhitekturo platforme, predvidene uporabnike platforme ter nekaj zanimivejših scenarijev njene uporabe.



Slika 2: Visokonivojska arhitektura platforme Occapi

Platforma je sestavljena iz več komponent, razdeljenih v dve skupini:

- a) Komponente, ki zagotavljajo zajem in procesiranje dogodkov, generiranih v opazovanem okolju, obveščanje o posebnih dogodkih ter spremljanje ključnih indikatorjev opazovanega sistema;
- b) Komponente, ki omogočajo spremljanje in nadzor opazovanega sistema v daljšem časovnem obdobju, integracijo s poljubnimi javnimi in zasebnimi podatkovnimi viri ter napredno analitiko tako zajetih podatkov.

Visokonivojsko arhitekturo, iz katere je razvodna predvsem jedrna funkcionalnost platforme, prikazuje slika 2. Sledi kratek opis posameznih komponent.

2.1 Generatorji dogodkov

Na sliki glejte oznako 1 (ang. Event Generators).

Osnovni vhod v platformo so dogodki, ki se kreirajo v opazovanem okolju. Generator dogodkov so lahko:

- a) Senzorji in naprave, ki posredujejo merjene podatke v platformo,
- b) SOA/EDA okolja, to je aplikacije, ki temeljijo na storitveno in dogodkovno usmerjenih arhitekturah ter
- c) Zastareli sistemi, ki dogajanje sistema beležijo v različne dnevnik.

Vmesniki za komunikacijo s platformo so različni. Za senzorje in naprave obstajajo standardi in mnoge naprave danes že omogočajo avtomatsko oddajanje merjenih podatkov na standardizirane vmesnike. Prav tako so vmesniki za SOA in EDA okolja relativno enostavni. Za zastarele sisteme pa so vmesniki kompleksnejši, saj prevzemajo vlogo generatorja. Dogodke razberejo iz posredovanega dnevnika, jih generirajo in posredujejo v platformo. Vmesniki za zastarele sisteme so prilagojeni vsakemu sistemu posebej.

2.2 Vhodna vrsta

Na sliki glejte oznako 2 (ang. Event Channel, JMS).

Dogodki, poslani v platformo, se zapišejo v vhodno vrsto (Enterprise JMS). Iz vhodne vrste se dogodki posredujejo na dva kanala: (a) v komponento za procesiranje dogodkov ter (b) v podatkovno bazo. Komunikacija je asinhrona.

2.3 Procesiranje dogodkov

Na sliki glejte oznako 3 (ang. Event Processing Kernel).

Dogodki se v platformi procesirajo. Komponenta za procesiranje omogoča:

- *Filtriranje dogodkov* (ang. Event Filtering): komponenta lahko v okviru filtriranja dogodka razvrsti po v naprej določenih skupinah, združuje, doda meta podatke, eliminira dvojnike itn. Primer: pri elektronskem cestninjenju se zgodi, da nek dogodek (ang. ticket) pride v sistem dvakrat. V okviru filtriranja komponenta za procesiranje dogodkov podvojeni dogodek izloči.

- *Detekcija pomembnih dogodkov* (ang. Identification of notifiable Events): pomembna funkcionalnost, ki jo komponenta za procesiranje dogodkov zagotavlja, je razpoznavna t.i. pomembnih dogodkov, to je takih, ki nas zanimajo v okviru sistema, ter generiranje izhodov (na sliki označeno z zeleno puščico in oznako »processed events«). Izhodi, ki jih komponenta pripravi, so prav tako dogodki, ki s seboj prinašajo vse potrebne podatke za objavo, bodisi na nadzorni plošči, v obliki alarmov ali v zunanji vrsti, namenjeni zunanjim sistemom za nadaljnjo obravnavo. Procesirane dogodke, ki jih komponenta kreira in posreduje komponenti za objavo dogodkov (ang. Event Publisher), imenujemo tudi KPI (ang. Key Performance Indicator).
- *Identifikacija vzorcev* (ang. Pattern Matching): komponenta za procesiranje dogodkov omogoča tudi identifikacijo vzorcev, ki nakazujejo verjetnost na obstoj ali pojavu določene zelene ali neželene situacije. Vzorci, ki jih komponenta išče, morajo biti definirani. Primer: v okviru elektronskega spremljanja prometa v večjih mestih je napovedovanje zastojev ter preusmerjanje prometa pomembna in uporabna funkcionalnost. Zastoj, ki nastane zaradi povečanja prometa, se da napovedati, saj se ne zgodijo naenkrat, temveč postopoma, pri čemer je dogajanje v času razvoja zastoja predvidljivo (nižja pretočnost prometa na določeni točki kot običajno, velik pritok vozil glede na poznano distribucijo obremenitev posameznih odsekov itn).
- *Analiza časovnih vrst* (ang. Time Series Analysis): pri procesiranju dogodkov pogosto ne zadošča izolirana obravnava posameznega dogodka, temveč moramo procesirati množico dogodkov, ki so se pojavili v nekem časovnem intervalu. Komponenta za procesiranje dogodkov omogoča analizo časovnih vrst dogodkov, pri čemer je velikost časovnega okna določena z vhodnim parametrom in je odvisna od zmogljivosti strojne opreme. Primer: v okviru elektronskega cestninjenja želimo beležiti povprečno hitrost vozil na določenem odseku za določeno časovno obdobje.
- *Koralacija med dogodki* (ang. Event Correlation): ne glede na okolje, ki generira dogodke, se večinoma izkaže, da dogodki niso med seboj neodvisni, temveč da obstajajo mnoge korelacije. Torej, da lahko na osnovi poznavanja trenutnega dogodka z veliko verjetnostjo napovemo pojavu drugega dogodka ali skupine dogodkov. Analiza dogodkov in računanje korelacije presega zmožnosti komponente za procesiranje dogodkov v realnem času in se izvaja v okviru analitike. Identificirane korelacije pa predstavljajo pomemben vhod za funkcionalnost, ki omogoča analizo časovnih vrst ter identifikacijo (novih) vzorcev.

2.4 Objava dogodkov

Na sliki glejte oznako 4 (ang. Event Publisher).

Izhod iz komponente za procesiranje dogodkov so t.i. generirani dogodki ali izhodni dogodki, ki predstavljajo pomembne podatke za različne prejemnike (alarmi, indikatorji). Takšni dogodki se posredujejo komponenti za objavo dogodkov, ki poskrbi, da so o nekem dogodku obveščeni vsi zainteresirani prejemniki.

2.5 Interna in eksterna izhodna vrsta

Na sliki glejte oznako 5 (ang. Internal Output Events) ter 7 (ang. Output Event Channel)

Izhodni dogodki se zapišejo v interno ter eksterno izhodno vrsto, kjer so na voljo internim in eksternim naročnikom za asinhrono obravnavo.

2.6 Interni naročniki

Na sliki glejte oznako 6 (ang. Internal Subscribers)

Platforma ločuje med notranjo in zunanjo uporabo procesiranih dogodkov. Za potrebe notranjega nazora platforme je na voljo generična nadzorna plošča, ki omogoča spremljanje dogajanja v platformi s pomočjo t.i. KPI (ang. Key Performance Indicator). KPI so lahko enostavni (posamezen dogodek, ki ga želimo spremljati, npr. trenutno število vozil na določenem odseku) ali kompleksni, to je sestavljeni iz več dogodkov (trenutni delež posameznih vrst vozil v prometu). Uporabnik lahko sam izbere KPI, ki jih želi spremljati, ter določi vrsto in način prikaza (tip grafa, frekvenca osveževanje, oznake itn.). Poseben primer KPI, ki ga lahko sistem posreduje tudi na elektronski naslov ali prek SMS, je alarm. Nadzorna plošča ima za potrebe prikaza alarmov posebno mesto.

2.7 Eksterni naročniki

Na sliki glejte oznako 8 (ang. External Subscribers)

Eksterni naročniki so o dogodkih obveščeni prek zunanje oziroma eksterne izhodne vrste. Platforma ne omejuje sistemov oziroma zunanjih naročnikov na dogodke. Ti so lahko npr. zunanje aplikacije, SOA procesi, mobilne naprave itn.

2.8 Začasna hramba podatkov

Na sliki glejte oznako 9 (ang. Stageing Area)

Za vzpostavitev zanesljive platforme je permanentno shranjevanje dogodkov potrebno. Problem nastane, ko je frekvenca dogodkov tako velika, da postane shranjevanje v podatkovno bazo ozko grlo. V okviru platforme je zato predvideno distribuirano shranjevanje podatkov, pri čemer je nivo distribucije odvisen od intenzivnosti nastajanja dogodkov.

V platformi imamo dve vrsti podatkovnih baz: eno namenjeno začasemu shranjevanju ter »čiščenju« in povezovanju podatkov ter drugo namenjeno hranjenju prečiščenih in skladnih podatkov (glej Integrirana podatkovna baza).

2.9 Integrirana podatkovna baza

Na sliki glejte oznako 13 (ang. Integrated Data)

V integrirani podatkovni bazi se hranijo »prečiščeni« podatki, skladni z globalnim katalogom (glej tudi oznako 12). Baza je namenjena direktnemu poizvedovanju ter poglobljenim analizam (analitika). Integrirana podatkovna baza je distribuirana in nameščena na klasičnem podatkovnem strežniku (podatkovne baze tipa noSQL se ne kažejo primerne).

2.10 Procesiranje podatkov

Na sliki glejte oznako 11 (ang. Data Processing Kernel).

Komponenta za procesiranje podatkov omogoča:

- *Razpoznavo entitet* (ang. Entity Resolution): podatki/dogodki prihajajo v sistem iz različnih virov. Komponenta za procesiranje podatkov je sposobna razpoznati podatke, ki predstavljajo primerke konceptov, ki so opredeljeni v globalnem katalogu oziroma

ontologiji. Podatki, ki jih komponenta ne razpozna, se ignorirajo oziroma ostanejo v začasni hrampi za nadaljnjo obravnavo.

- *Eliminacija duplikatov* (ang. Duplicate Elimination): ker se podatki/dogodki prenašajo v sistem iz različnih virov, se zgodi, da se v začasni hrampi znajdejo tudi duplikati. Komponenta za procesiranje podatkov zazna duplikate in jih eliminira.
- *Obogatitev podatkov* (ang. Data Enrichment): obogatitev podatkov je zelo uporabna, vendar tudi kompleksna funkcionalnost komponente za procesiranje podatkov. Vsaka entiteta, ki jo komponenta razpozna, se preveri, ali vsebuje vse podatke/attribute, ki so za tak tip entitete predvideni v globalnem katalogu oziroma ontologiji. Če določeni podatki manjkajo, jih komponenta lahko poskusi poiskati v alternativnih virih. Med alternativnimi viri se najpogosteje uporablja splet z javno dostopni podatki, lahko pa tudi drugi poljubni strukturirani ali nestrukturirani podatkovni viri. Med alternativne vire sodijo tudi podatki, ki jih prispevajo druge platforme. Za izbiro najbolj perspektivnih alternativnih virov si komponenta pomaga z meta-iskalnikom (ang. Meta-Search Engine, oznaka 14), ki služi kot vmesnik med poizvedbo in alternativnimi viri. Primer: v sistem se vpišejo/poizvedujejo podatki o vozilu. V globalnem katalogu je entitetni tip vozilo povezan z entitetnim tipom lokacija. V ontologiji je zapisano, da je lokacija podatek, ki ga je moč pridobiti prek storitve A, ki je na voljo na spletu. Storitve implementira zunanja platforma v sklopu svojega API (glej ang. External Data Sources, oznaka 15). Komponenta podatke, ki se vpisujejo oziroma po katerih poizvedujemo, prebere iz integrirane podatkovne baze, manjkajoče podatke pa dopolni s podatki, ki jih najde v alternativnih virih.
- *Ekspertne analize* (ang. Expert Analysis): ekspertne analize so domensko specifične in v generični platformi v naprej niso realizirane. Platforma OCAPI ponuja vtič, prek katerega je moč vključiti tudi zunanje komponente, ki realizirajo kompleksne analize.
- *Identifikacija kompleksnih vzorcev* (ang. Complex Pattern Matching): komponenta za procesiranje podatkov omogoča identifikacijo kompleksnih vzorcev, ki so za uporabnika s kakršnegakoli razloga pomembni. Vzorci, ki jih komponenta išče, morajo biti definirani.
- *Predikcijo in napovedovanje* (ang. Prediction, Forecasting): komponenta za procesiranje podatkov vsebuje tudi funkcije za enostavno napovedovanje in predvidevanje. Za kompleksnejše poizvedbe (OLAP) je boljše uporabiti zunanje sisteme s področja poslovne inteligence.

2.11 Neposreden dostop do podatkov

Na sliki glejte usmerjeno povezavo iz 8 do 14 (ang. Catalog Request/Direct Data Query).

Zunanjim naročnikom platforma omogoča tudi neposredno poizvedovanje po podatkih, ki sodijo v kontekst platforme. To so podatki, ki jih platforma hrani, ter podatki, do katerih platforma dostopa prek zunanjih virov (tudi zunanjih platform). V globalnem katalogu platforme so zapisani meta podatki, ki povedo, kjer se posamezne podatkovne kategorije nahajajo ter v primeru, da gre za podatke zunanjih ponudnikov ali podatke svetovnega spleta, prek katerih storitev so dosegljivi. Zunanji naročniki lahko prek storitev, ki jih platforma ponuja, pridobijo globalni katalog podatkov, iz katerega vidijo, katere podatke platforma posredno ali neposredno ponuja. Po podatkih lahko poizvedujejo z navadnim SQL jezikom.

Pomembna komponenta, ki omogoča distribuirano poizvedovanje in je v smislu aplikativne arhitekture en nivo višje kot globalni koordinator (globalni koordinator se v porazdeljenih PB uporablja za distribucijo in koordinacijo poizvedbe), je meta iskalnik (ang. Meta Search Engine). Meta iskalnik poizvedbo pregleda (sintaktična in semantična analiza poizvedbe). Iz globalnega kataloga ugotovi, kateri podatki so dostopni neposredno ter kateri posredno prek zunanjih virov oziroma spleta. Upoštevajoč te podatke razčleni poizvedbo v naslednje tri dele:

- Poizvedba A, ki se posreduje globalnemu koordinatorju interne porazdeljene PB – podatki se vrnejo v strukturi, ki ustreza ontologiji platforme;
- Poizvedba B, ki za pridobitev podatkov kliče storitev zunanjih sistemov – podatki se vrnejo v strukturi, ki ustreza ontologiji platforme, in
- Poizvedba C, ki za pridobitev podatkov angažira spletne iskalnike. V tem primeru se rezultati vrnejo kot seznam zadetkov, ki jih v končno obliko pretvori komponenta za predprocesiranje (ang. Preprocessing, glejte komponento 14).

3. UPORABNIKI PLATFORME

Tipični uporabniki platforme se delijo v naslednje skupine:

- *Ponudniki podatkov*: ponudniki podatkov so pasivni uporabniki, ki niso zainteresirani za uporabo funkcionalnosti platforme, temveč v platformo le posredujejo svoje podatke. Naprej jih delimo na posredne in neposredne ponudnike.
 - Neposredni ponudniki v platformo »prispevajo« dogodke, ki se generirajo v njihovem okolju. Tak primer bi lahko bili vsi tisti, ki zajemajo podatke v okolju, na primer v senzorskih omrežjih, kot npr. IJS, ki bi prek svoje senzorske mreže posredoval dogodke (meritve senzorjev) v platformo.
 - Druga vrsta ponudnikov so tisti, ki podatke v platformo posredujejo posredno, torej prek povezav med platformami (glej točko 2.11).
- *Uporabniki podatkov*: med uporabnike podatkov štejemo tiste uporabnike, ki so zainteresirani za direkten dostop do podatkov platforme, bodisi neposredno s poizvedbami ali v kombinaciji z zunanjimi OLAP sistemi.
- *Uporabniki platforme*: uporabniki platforme se od uporabnikov podatkov razlikujejo v tem, da ne uporabljajo le podatkovnega dela platforme, temveč tudi dogodkovni. Naprej se delijo na posredne in neposredne uporabnike platforme.
 - Neposredni uporabniki platforme spremljajo dogajanje v platformi prek nadzorne plošče in alarmov, ki jih platforma proži.
 - Posredni uporabniki pa platformo uporabljajo le kot prožilec svojih internih sistemov.

Uporabniki lahko seveda nastopijo tudi v več vlogah (npr. ponudniki podatkov in uporabniki platforme).

4. SCENARIJI UPORABE

Platforma Occapi je uporabna povsod, kjer imamo opravka z velikimi količinami dogodkov (podatkov), še posebej, če bi želeli na njih reagirati v realnem ali skoraj realnem času. Med scenarije uporabe zagotovo sodijo: pametna mesta, pametna logistika, pametno zdravstvo, pametna energija, pametne zgradbe, pametni domovi, detekcija nepravilnosti in zaračunavanje v realnem času, elektronsko cestninjenje itn.

5. ZAHVALA

Avtorji prispevka se zahvaljujemo vsem partnerjem, ki sodelujejo v okviru izvajanja programa kompetenčnega centra KC-OPCOMM (Zavod Tehnološka mreža ICT, Alpineon d.o.o., Cosylab d.d., Innova IT d.o.o., Institut Jožef Stefan, Špica International d.o.o., Globtel d.o.o., Univerza v Ljubljani, Fakulteta za elektrotehniko, Univerza v Ljubljani, Fakulteta za računalništvo in informatiko) ter Ministrstvu za visokošolstvo znanost in tehnologijo za sofinanciranje.

6. ZAKLJUČEK

V prispevku je podana kratka predstavitev inteligentne komunikacijske platforme Occapi, ki jo razvijamo v okviru programa kompetenčnega centra OPCOMM. V prispevku smo prikazali predvsem jedrne funkcionalnosti platforme.