



Raport stiintific si tehnic

BIG-SMART-LOG

Etapa 1 (2019)

Parteneri

Consortiul este format din 5 parteneri provenind din 4 tari: Turcia, Grecia, Polonia si Romania.

Sigla	Partener	Tara	Website
	Ekol Lojistik, Technology Group (Turkey)	Turcia	www.ekol.com
	Istanbul Şehir University (Turkey)	Turcia	www.sehir.edu.tr
	Centre for Research and Technology Hellas, Information Technologies Institute (Greece)	Grecia	www.certh.gr
	The John Paul II Catholic University of Lublin, Faculty of Mathematics, IT and Landscape Architecture (Poland)	Polonia	www.kul.pl/institute-of-mathematics-and-computer-science,2141.html
	SC HOLISUN SRL	Romania	www.holisun.com

Sumar al proiectului

„Logistics 4.0” (*Barreto et al. (2017)*) se referă la combinația de utilizare a logisticii cu inovațiile și aplicațiile adăugate de Cyber Physical Systems (CPS). „Logistica inteligentă” este un sistem logistic, care poate spori flexibilitatea, ajustarea la schimbările pieței și va ajuta compania să fie mai aproape de nevoile clienților. Această nouă paradigmă este rezultatul utilizării sporite și îmbunătățite a internetului care permite comunicarea între mașini, dispozitive și oameni în timp real și utilizarea a ceea ce este cunoscut sub numele de „Digitalizare avansată”. Concurența provocatoare pe piața globală obligă companiile de transport să dezvolte soluții inteligente pentru gestionarea mai eficientă a logisticii și a operațiunilor lanțului de aprovizionare. În procesul de căutare a acestor soluții, un număr emergent de companii îmbracă oportunitățile pe care le oferă conceptul de *BigData*.

Acest proiect își propune să abordeze provocarea re-proiectării întregului sistem de telemetrie și date produse prin tehnologia de urmărire a vehiculelor activată prin satelit (IoT-Internet of Things) către sisteme inteligente care comunică cu alte vehicule, care schimbă între ele informații despre trafic, vreme și condiții de drum, etc. Sistemul își propune integrarea datelor de la diferiți senzori cu privire la vehiculul supravegheat și sarcina acestuia, informații pe baza cărora actualizările de rută și locație sunt setate în mod automat și autonom pentru fiecare vehicul. Acest sistem va agrega inclusiv comenzi de lucru ale clienților, servicii de logistică ale terților (transportatori, expeditori, operatori de logistică și terminale), sisteme open source și social media.

Scopul proiectului este de a proiecta un model de procesare bazat pe auto-învățare îmbunătățită semantic, care să surprindă ritmul infrastructurii existente prin analiza datelor operaționale în timp real. Proiectul utilizează datele învățate pentru a influența pozitiv utilizarea infrastructurii și resurselor existente, pentru a crește rezistența sistemului și calitatea serviciilor și reducerea emisiilor de gaze cu efect de seră, consumul de combustibil, timpul inactiv, congestionarea traficului, pierderile de venituri, livrările târzii, reclamațiile clienților prin utilizarea tehnologiilor *Big Data*. Pe lângă asta, învățarea automată, învățarea profundă, inteligența artificială și algoritmi de *Data Science* vor oferi predicții și prescripții pentru a maximiza eficiența operațională, protecția mediului și siguranța. Rezultatul proiectului este construirea unui cadru reproductibil semantic, care poate fi aplicat diferitelor organizații pe mai multe medii.

Activități

Planul de activități este următorul:

Etapa	Perioada
Modelarea arhitecturii platformei de analiza	01/06/2019 - 31/12/2019
Dezvoltarea platformei și a algoritmilor de analiza a datelor	01/01/2020 - 31/12/2020
Validarea platformei de analiza IoT	01/01/2021 - 31/05/2021

În cadrul etapei 1 (aferea anului 2019), sunt prevazute doua activitati:

- Activitatea 1.1: Modelarea ontologiei cadru pentru analiza IoT
- Activitatea 1.2: Proiectarea arhitecturii platformei IoT

Activitatea 1.1 - Modelarea ontologiei cadru pentru analiza IoT

Specificațiile generale au fost obținute de la aplicantul principala – EKOL, care este și beneficiarul precum și business case providerul.

Cerințele sunt rezumate în continuare:

- O arhitectură generică ce facilitează operația a diferite scenarii în domeniu logistic prin
 - Abstractizarea procesărilor componentelor Big Data precum și componentelor logicii de business (ex: tehnicile statistice Bayes);
 - Sprijinirea aplicării diferitelor tehnici de analiză în arhitecturile distribuite.
- Încorporarea, procesarea și prelucrarea datelor din diferite surse, formate și modele (atât fluxuri de date în timp real cât și colecții pre-existente de date). Exemple indicative de surse de date sunt social media, sistemele de informații de management/gestiune precum și date de circulație liberă.
- Sisteme de asigurare a confidențialității, când mai multe entități implicate în cadrul unui lanț de furnizare nu sunt dispuse să-și dezvăluie datele
- Algoritmi din mediul online, cu cerințe de operare în timp real, ce își fac apariția în industria logistică
- Recuperarea sau restabilirea datelor lipsă sau corupte:
 - Algoritmi distribuiți
 - Algoritmi de împărțire pe cozi (*distributed queueing*)
 - Sisteme și funcționalități de tip *MapReduce*.

- Transformări de date reutilizabile care primesc ca intrări date de tip static, dinamic sau fluxuri
- Localizarea datelor pe bază geografică:
 - *Edge Computing*
 - Scalare orizontală
- Vizualizări
 - Rețeaua (sub formă de graf)
 - Analiza și interpretarea datelor
- Analize prescriptive:
 - Siguranța
 - Rutare dinamică

Ontologia este o conceptualizare formală a unei cunoștințe particulare despre lume, prin reprezentarea explicită a unor concepte de bază, relații precum și reguli de deducție despre ele însăși (Noy (2004)). Ontologiile pe domeniu pot fi folosite pentru a furniza suport de cunoștințe în procesele cognitive de baza precum și inter-relațiile dintre acestea.

Câteva motive pentru care ar fi utilă dezvoltarea unei ontologii este descrisă de către Noy & McGuinness (2001): pentru a partaja o înțelegere comună a structurii informației între oameni și/sau agenți software, pentru a permite reutilizarea cunoștințelor dintr-un anumit domeniu, de a face ipoteze de domeniu explicite, de a separa cunoștințele de domeniu de cunoștințele operaționale. Dezvoltarea unei ontologii presupune crearea unui set de date împreună cu structura acestora pentru uzul altor aplicații. Metodele de rezolvare a problemelor, aplicațiile independente de domeniu precum și agenții software folosesc ontologiile precum și bazele de cunoștințe extrase din acestea ca date.

Ontologia dezvoltată în cadrul proiectului permite atât integrarea cunoștințelor și a modului de lucru al companiei, cât și modelarea constrângerilor de lucru (funcționale sau de exploatare) și chiar modelarea modului de rezolvare a diferitelor probleme apărute pe perioada de exploatare a sistemului. Spre deosebire de un sistem bazat pe o bază de date convențională, o ontologie reține nu doar date bine stabilite și legături rigide între acestea, ci și cunoștințe noi, acumulate de sistem într-un mod dinamic, pe parcursul exploatării. Astfel, sistemul este în stare să extragă concepte din aceste cunoștințe, atât pentru a completa eventualele goluri (prin deducție), cât și pentru a rezolva probleme pentru care nu are o rezolvare explicit introdusă în faza de proiectare.

Alăturarea ontologiei unui sistem de inteligență artificială capabilă atât să extragă date din ontologie cât și să

introducă unele noi, aduce sistemul la un nou nivel de dezvoltare, oferindu-i puterea atât de-a reacționa prompt la evenimente neprevăzute, cât și de a-și extinde cunoștințele și a-și evalua în timp calitatea acțiunilor sau precizia rezultatelor.

Pentru ameliorarea resurselor folosite și a timpului de răspuns, ontologia este împărțită în mai multe module.

Acestea conceptualizează următoarele entități (v. Figura 1):

- **Device:** integrează diferitele platforme fizice utilizate, particularitățile acestora precum și diferitele conversii necesare și limitări de care sistemul trebuie să țină cont în comunicarea cu acesta, la un nivel scăzut, aproape de hardware.
- **Platform:** date despre diferitele componente software ce compun sistemul, modul de interacțiune dintre acestea și protocoalele folosite (*Seydoux et al. (2016)*).
- **Units and measurments:** sistemele și unitățile de măsură precum și modul de conversie între acestea. Cuprinde și standardele ce vor fi folosite de sistem pentru a persista datele (salvarea lor în fișiere/baze de date).
- **Location:** cuprinde sistemul de poziționare globală, diferitele standarde de modelare a coordonatelor precum și conversia de precizie între acestea. Acest modul integrează și modul de analiză al datelor despre locație venite din surse exterioare (de exemplu Google Maps, date despre trafic, CNADNR).
- **Observation and actuation:** integrează diferitele observații ale sistemului referitoare la entitățile periferice. Modulul integrează datele telemetrice și posibilitatea de-a trimite informații.
- **UX Services:** Serviciile de interfață (UX) disponibile. Sistemul va putea hotărî ce date sunt disponibile diferitelor clase de utilizatori, în funcție de drepturile atribuite acestora în modulul aferent, va determina informațiile disponibile în funcție de sistemele disponibile (senzori, module hard, limitări ale acestora) precum și acțiunile manuale ce pot fi efectuate de utilizatori.
- **User:** clasele de utilizatori din sistem, include atât utilizatorii umani cât și entitățile software din cadrul sistemului.

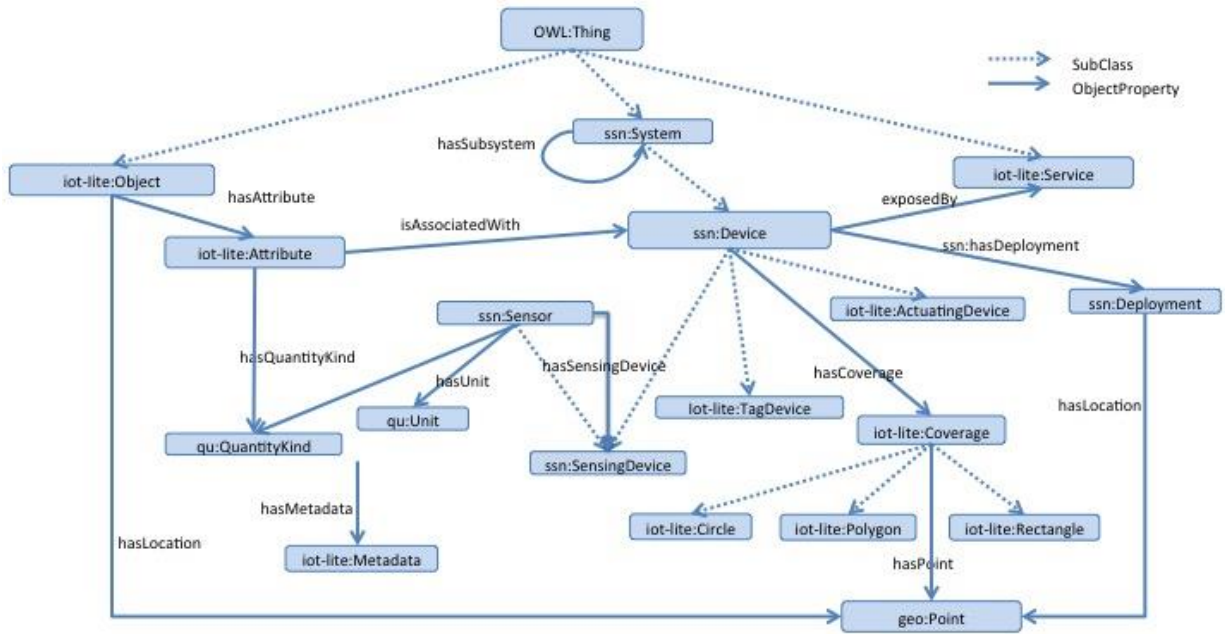


Figura 1. Structura globala a ontologiei

Modulele sunt interconectate la nivel logic, fiecare dintre ele conține informațiile necesare conexiunilor între modulele vecine aferente. Interconectarea, transferul de date și comunicarea internă se fac prin modele de date nestructurate, cum ar fi *JSON*.

Ontologia conține, de asemenea, informații despre starea sistemului la un moment dat. Aceasta s-a implementat utilizându-se formatul standard dezvoltat de industrie, și anume *object-subject-predicate*. Astfel, sistemul poate fi, de exemplu, interogat referitor la starea unui vehicul, la locația acestuia, însă ontologia permite și interogări de nivel înalt, de exemplu dacă vehiculul este în pericol de a nu-și putea efectua sarcina, dacă operatorul acestuia îl exploatează în condiții optime/de siguranță, poate oferi sugestii referitoare la îmbunătățirea modului de îndeplinire a sarcinilor pentru operatori, etc.

Activitatea 1.2 - Proiectarea arhitecturii platformei IoT

Arhitectura generală a sistemului este prezentată în Figura 2.

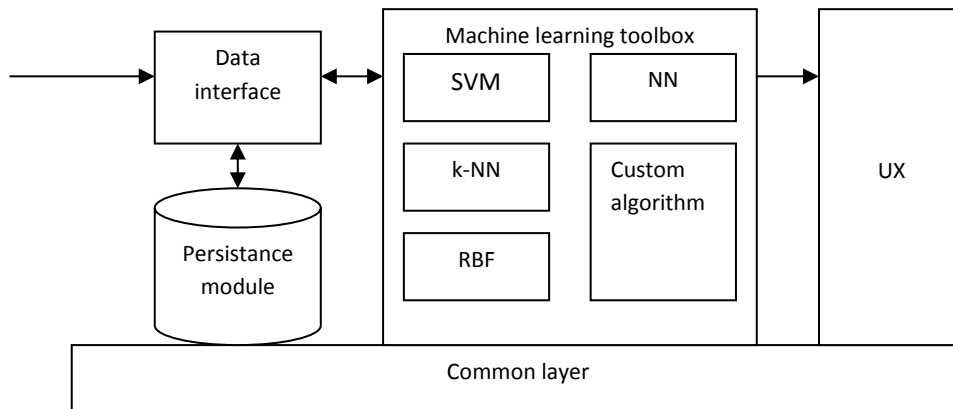


Figura 1. Arhitectura generală a sistemului de inteligență artificială cu învățare automată

Rezultatul acestui proiect va produce o platformă reutilizabilă pentru medii dinamice, cu constrângeri personalizabile și surse de date parametrizate. Platforma poate fi implementată pe diferite lanțuri logistice din diferite culturi, organizații sau intervale de timp (Matei et al. (2017)).

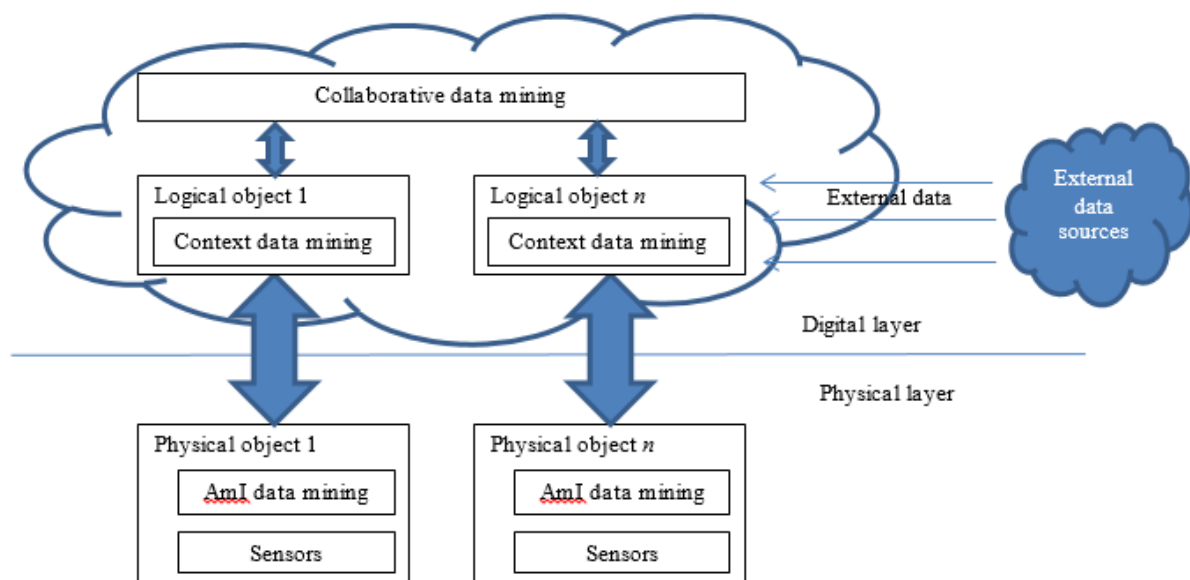


Figura 2. Arhitectura multistrat a sistemului de procesare a informațiilor

Arhitectura globală propusă în figurile 2 și 3 iau în calcul limitările fizice ale sistemelor IoT și a fluxului de date dintre acestea și sistemul cloud. De aceea sunt dezvoltate mai multe straturi (*layers*) ce contribuie activ la abordarea de tip Data Mining de extragere a informațiilor din datele disponibile (*Matei, Anton et al. (2017)*). Această abordare este foarte potrivită atunci când există deja platforme de colectate și analiză complexă a datelor în cloud, așa cum este situația de față.

- 1) Datorită resurselor limitate, atât de calcul cât și de conectivitate, este dificil sau poate chiar imposibil de transmis absolut toate datele către platforma cloud. De aceea este foarte importantă o primă prelucrare, locală, a datelor (așa numitul *Edge Computing*) chiar în cadrul microsistemului IoT instalat pe vehicul. Acesta va avea sarcina de-a face o primă integrare a datelor, extrăgând informațiile importante din acestea pe termene scurte (de ex o oră). Acestea vor fi salvate temporar local, urmând ca la realizarea conexiunii de internet să se sincronizeze cu platforma cloud.
- 2) Pe nivelul cloud a platformei, se pot executa operații de analiză și data mining mai complexe. Limitele sunt impuse doar de datele disponibile, furnizate de echipamentele Edge instalate pe vehicule, deoarece virtual nu există nici o limitare sau restricție legată de puterea de calcul și resursele folosite în cadrul platformei cloud. Cu toate acestea, un factor important este dat de limitările de timp, în special în faza de învățare a algoritmilor.
- 3) Dacă există date suplimentare legate de ambientul în care se execută sarcinile, se poate extrage un context al sarcinii (*Avram et al. (2019)*). Aceasta se realizează pe baza datelor Aml precum și a datelor suplimentare externe, furnizate de terți (de exemplu prognoze meteo, date referitoare la trafic, etc). Aceasta este util pentru a obține o imagine de ansamblu clară a circumstanțelor în care se obține Aml precum și pentru a testa optimalitatea executării unor sarcini. Sistemul poate evalua aceste aspecte și alerta personalul responsabil în caz de nevoie.
- 4) Dacă în cadrul ontologiei din cloud există mai multe obiecte digitale (reprezentarea virtuală a unor sisteme fizice), așa cum se întâmplă de obicei, se poate obține o imagine de ansamblu globală asupra populației de obiecte, împreună cu relațiile dintre ele (*Anton et al. (2019)*). Algoritmii pot utiliza abordări de colaborare pentru a-și crește precizia și pentru a crește gradul de cunoaștere și înțelegere a sistemului fizic, fără a fi nevoie de intervenția unui operator.

Fluxul de date și informații din cadrul IoT

Fluxul de date și informații are un flux ascendent, de la senzorii obiectului fizic către platforma cloud ce conține toate copiile digitale ale obiectelor fizice, așa cum se arată în Figura 3 (*Matei, Anton et al. (2017)*):

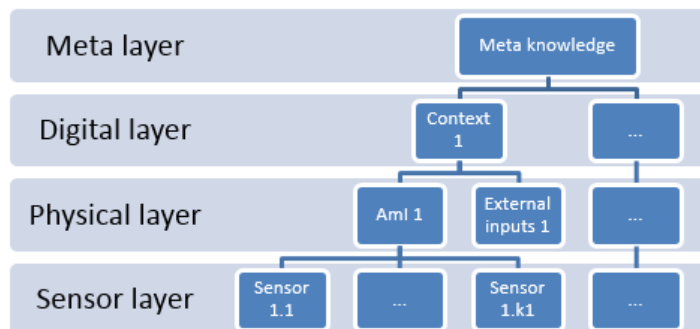


Figura 3. Fluxul de date în IoT

1) Pentru fiecare obiect, datele de la senzori sunt agregate ca inteligență ambientală (Aml), o tehnologie care va deveni încorporată în mediul nostru natural, prezentată ori de câte ori avem nevoie, activată prin interacțiuni simple și fără efort, adaptabilă la utilizatori și sensibili la context și autonom. Datele de la senzori pot fi prelevate pe timp sau pe evenimente. Prima abordare generează semnificativ mai multe date decât cea din urmă, dar poate fi mai exactă în cazul analizei. Cu toate acestea, aceasta din urmă se declanșează atunci când are loc o schimbare importantă a intrărilor și oferă suficiente date pentru statistici detaliate.

2) Din stratul fizic, datele sunt transmise în sus către stratul digital (virtual), unde Aml și alte date externe obținute din surse terțe, cum ar fi vremea, prezența / absența anumitor circumstanțe etc. sunt agregate în contexte printr-un proces numit extragerea contextului (Avram et al. (2019)).

3) Cu toate acestea, stratul digital conține o colecție de obiecte, de aceea derivăm un al patrulea strat, numit meta-strat, care permite perfecționarea suplimentară a datelor prin combinarea contextelor tuturor obiectelor. Spre deosebire de cea de-a doua etapă, de extragere a contextului, această etapă folosește alte contexte de obiecte (agenți), care încapsulează deja mai multe Aml și date externe (Anton et al. (2019)).

Concluzii

Această etapă, cu o durată de derulare de 6 luni, a pregătit terenul pentru etapele 2 și 3. În această etapă au fost desfășurate 2 activități:

- **Activitatea 1.1:** Modelarea ontologiei cadru pentru analiza IoT
 - o În cadrul acestei activități a fost modelată ontologia cadru ce va fi utilizată pentru analiza datelor emise de dispozitivele din cadrul IoT, analiza acestora și extragerea informațiilor pertinente, utile, din aceste date. Abordarea pentru realizarea ontologiei a fost una modulară, proiectată pentru scalabilitate și flexibilitate, având posibilitatea de-a integra componente diverse, cu limitări multiple și diferite particularități, sau chiar componente

inexistente astăzi pe piață. Ontologia va susține activ procesul de învățare al algoritmilor, permițând extragerea rapidă și exactă a informațiilor din ea precum și oferind un cadru complet de salvare cât mai completă și corectă a datelor și informațiilor despre sistem (luat ca întreg).

- **Activitatea 1.2:** Proiectarea arhitecturii platformei IoT
 - A fost dezvoltată arhitectura sistemului destinat colectării și analizării datelor. Aceasta se bazează pe două concepte inovative, larg răspândite azi, și anume Cloud Computing și Edge Computing. Fiecare dintre cele două contribuie decisiv la utilitatea sistemului, la optimizarea funcționalităților, optimizarea utilității și utilizării acestuia, prin rezolvarea unor probleme generale încă prezente astăzi peste tot (lipsa conexiunilor permanente la rețeaua internet, vitezele mici de transfer, componente hardware limitate, cost ridicat al acestora).

Referințe

- Matei et al. (2017)* Matei, Oliviu, et al. "Multi-layered data mining architecture in the context of Internet of Things." 2017 IEEE 15th International Conference on Industrial Informatics (INDIN). IEEE, 2017.
- Matei, Anton et al. (2017)* Matei, O., et al. "Multi-layered architecture for soil moisture prediction in agriculture 4.0." Proceedings of International Conference on Computers and Industrial Engineering, CIE. Vol. 2. No. 1. 2017.
- Seydoux et al. (2016)* Seydoux, Nicolas, et al. "IoT-O, a core-domain IoT ontology to represent connected devices networks." European Knowledge Acquisition Workshop. Springer, Cham, 2016.
- Noy & McGuinness (2001)* Noy, Natalya F., and Deborah L. McGuinness. "Ontology development 101: A guide to creating your first ontology." (2001).
- Anton et al. (2019)* Anton, Carmen Ana, Oliviu Matei, and Anca Avram. "Collaborative Data Mining in Agriculture for Prediction of Soil Moisture and Temperature." Computer Science On-line Conference. Springer, Cham, 2019.
- Avram et al. (2019)* Avram, Anca, et al. "Context-Aware Data Mining vs Classical Data Mining: Case Study on Predicting Soil Moisture." International Workshop on Soft Computing Models in Industrial and Environmental Applications. Springer, Cham, 2019.
- Barreto et al. (2017)* Barreto, Luis, Antonio Amaral, and Teresa Pereira. "Industry 4.0 implications in logistics: an overview." Procedia Manufacturing 13 (2017): 1245-1252.
- Noy (2004)* Noy, Natalya F. "Semantic integration: a survey of ontology-based approaches." ACM Sigmod Record 33.4 (2004): 65-70.