



Proiect SIPOCA 593 – „Sistem de monitorizare a fluxurilor de deșeuri menajere și similare în scopul îmbunătățirii mecanismelor de gestionare a instrumentului economic “Plătește Pentru Cât Arunci”” Beneficiar: Ministerul Mediului, Apelor și Pădurilor

Subactivitatea A4.1. Realizarea platformei în vederea derulării procesului de monitorizare digitală a fluxurilor de deșeuri menajere

REALIZARE PLATFORMĂ FUNCȚIONALĂ DE COLECTARE A DATELOR PRIVIND PPCA

Autori:

**Anica-Popa Ionuț-Florin Rădulescu
Maria Cristina Vrâncianu Marinela
Anica-Popa Liana-Elena Tudor Cătălin
Georgel Gorgan Vasile
Radu Lupu
Harhătă Victor-Florin Ene Marian
Olariu Ana Alexandra**

**RESPONSABIL DE PROIECT ASE:
VALENTIN LAZĂR**

Cuprins

1	Sistemul de colectare a deșeurilor "Plătești pentru cât arunci" (PPCA).....	4
1.1	Problematika deșeurilor în context global, regional și național	4
1.1.1	Situația deșeurilor la nivel global.....	5
1.1.2	Situația deșeurilor la nivelul Uniunii Europene	7
1.1.3	Situația deșeurilor în context național.....	13
1.2	Posibile abordări în managementul deșeurilor: „Orașul Inteligent”	17
1.3	Sistemul de colectare a deșeurilor PPCA (PAYT).....	23
1.3.1	Principii generale	23
1.3.2	Adoptarea sistemului PPCA în statele membre.....	27
1.3.3	Analiză comparativă a implementărilor schemei "Pay-as-you-throw"	30
1.3.4	Concluzii.....	34
2	Tehnologii suport	36
2.1	Sisteme de gestiune a bazelor de date	36
2.1.1	Sistemul de gestiune a bazelor de date MySQL.....	38
2.1.2	Sistemul de gestiune a bazelor de date Microsoft SQL Server	41
2.1.3	Asemănări și diferențe între MySQL și SQL Server	44
2.1.4	Considerații privind utilizarea MySQL și Microsoft SQL Server pentru implementarea platformei de colectare a datelor privind PPCA	49
2.2	Soluții tehnice pentru schimbul electronic de date.....	51
2.3	Arhitecturi și tehnologii pentru gestiunea și analiza datelor.....	55
2.3.1	Conceptul de Big Data.....	55
2.3.2	Baze de date, depozite de date, Data Lake: posibile alegeri pentru elaborarea arhitecturii sistemului dezvoltat	59
2.3.3	Arhitecturi bazate pe Data Lake	60

2.3.4	Etape în dezvoltarea Data Lake	64
2.3.5	Elementele unei abordări agile în construirea unui Data Lake.....	65
2.3.6	Considerente privind maturitatea unui Data Lake: Data Swamp, Data Pond și Data Puddles	66
2.3.7	Data Lakehouse.....	68
2.3.8	Abordări agile și flexibile: deschideri către tendințe privind evoluția arhitecturilor de date	68
2.3.9	Depozite de date.....	72
2.3.10	Extragerea, transformarea și încărcarea datelor	77
2.3.11	Tablouri de bord	79
2.3.12	Considerente în alegerea unui instrument pentru vizualizarea datelor analitice la nivel tactic și strategic: Power BI vs. Tableau vs. QlikView	80
3	Arhitectura informatică a sistemului propus	83
3.1	Descrierea proceselor de afaceri privind componenta de integrare a sistemelor informatice operaționale și de analiză a datelor privind colectările de deșeuri	87
3.1.1	Etapizarea procesului de afaceri principal privind centralizarea și analiza datelor...	87
3.1.2	Subprocesul de colectare date pentru perioada de raportare.....	89
3.1.3	Subprocesul de transfer de date din baza de date operațională în cea centralizatoare	91
3.1.4	Subprocesul privind procesarea transferurilor de date și generarea înregistrărilor în baza de date centralizatoare	93
3.1.5	Subprocesul privind urmărirea stării transferurilor efectuate în baza de date centralizatoare	95
3.2	Sistemul operațional de colectare a deșeurilor	97
3.3	Schimbul electronic de date între componentele sistemului	97

3.3.1 Modelul de date necesare componentei software de transfer a datelor din sistemul operațional în depozitul de date.....	115
3.3.2 Etape principale în dezvoltarea sistemului informatic de analiză a datelor colectate prin programul PPCA.....	118
3.3.3 Construirea serviciului web pentru transferul informațiilor în baza de date centralizatoare a sistemului de analiză a datelor prin programul PPCA.....	121
3.4 Sisteme de asistare a deciziei la nivel tactic și strategic pentru programul PPCA.....	128
4 Concluzii	135
Bibliografie	136

1 Sistemul de colectare a deșeurilor *”Plătești pentru cât arunci”* (PPCA)

1.1 Problematika deșeurilor în context global, regional și național

Elaborarea platformei de monitorizare a colectării deșeurilor menajere a avut ca punct de plecare o analiză amplă efectuată de echipa de implementare cu privire la problematica deșeurilor în contextul actual global, regional și național, pentru a surprinde cât mai corect situația actuală și tendințele ce s-au profilat la diverse niveluri în generarea, colectarea și evacuarea deșeurilor.

Pentru o fundamentare riguroasă din punct de vedere legislativ a sistemului proiectat, au fost studiate și prezentate succint cele mai importante reglementări legislative referitoare la gestiunea deșeurilor, la nivelul Uniunii Europene și la nivel național. În scopul identificării unei soluții tehnice optime, au fost cercetate studii de caz semnificative referitoare la sisteme fundamentate pe tehnologii moderne pentru managementul deșeurilor ca inițiative specifice așa-numitelor „orașe inteligente”.

Echipa a realizat un studiu amplu de analiză a instrumentului economic „Plătești pentru cât arunci” (PPCA) și a diverselor sale implementări în statele membre ale Uniunii Europene (UE). Plecând de la principiile care au stat la baza instituirii acestui sistem, trecând prin abordările teoretice ale posibilelor modalități de cuantificare a taxei de colectare a deșeurilor și prin schemele asociate sistemului PPCA, au fost ulterior evidențiate experiențele avute în implementarea sistemului PPCA de diverse state ale UE, cu limitele și beneficiile sale.

1.1.1 Situația deșeurilor la nivel global

Deșeurile, definite la articolul 3, alineatul (1) din Directiva 2008/98/CE reprezintă „orice substanță sau obiect pe care deținătorul îl aruncă, intenționează sau este obligat să îl arunce”, constituind o pierdere enormă de resurse atât sub formă de materiale, cât și de energie. În plus, atât gestionarea cât și eliminarea deșeurilor pot avea un impact grav asupra mediului. Depozitele de deșeuri, de exemplu, ocupă terenuri vaste, provocând poluarea aerului, apei și a solului, în timp ce incinerarea deșeurilor poate duce la emisii de poluanți atmosferici. Prin urmare, invariabil, toate politicile de gestionare a deșeurilor urmăresc reducerea impactului deșeurilor asupra mediului, sănătății umane, dar și îmbunătățirea utilizării resurselor. Scopul pe termen lung al acestor politici este reducerea cantității de deșeuri generate și, atunci când generarea de deșeuri este inevitabilă, promovarea acestora ca resursă și atingerea unor niveluri mai ridicate de reciclare și eliminare în siguranță a deșeurilor.

Un raport al Băncii Mondiale (World Bank Report, 2018) preconizează că problema deșeurilor va avea în viitorul apropiat un efect devastator asupra vieții umane și a mediului înconjurător, anticipându-se că deșeurile globale vor crește cu 70% până în 2050, dacă nu sunt puse în aplicare măsuri imediate pentru monitorizare și control. În plus, din cauza evoluției ascendente a populației globale și a urbanizării, se apreciază că volumul de deșeuri produs la nivel global va ajunge la 3,4 miliarde de tone în următoarele trei decenii.

Se estimează că generarea zilnică de deșeuri pe cap de locuitor în țările cu venituri ridicate va crește cu 19% până în 2050, în comparație cu țările cu venituri mici și medii, unde se așteaptă să crească cu aproximativ 40% sau mai mult. Cantitatea totală de deșeuri generată de țările cu venituri mici este de așteptat să crească de peste trei ori până în 2050. Regiunea Asiei de Est și Pacificului generează cea mai mare parte a deșeurilor din lume, circa 23%, iar regiunea Orientului Mijlociu și Africii de Nord produce cel mai puțin în termeni absoluți, circa 6%. Cu toate acestea, regiunile cu cea mai rapidă creștere sunt Africa Subsahariană, Asia de Sud, Orientul Mijlociu și Africa de Nord, unde, până în 2050, generarea totală de deșeuri este de așteptat să se dubleze și chiar tripleze. În aceste regiuni, mai mult de jumătate din deșeuri sunt în prezent aruncate în mod deschis, acest fapt necesitând acțiuni urgente de stopare a fenomenului, tendința de creștere a deșeurilor având deja implicații majore asupra mediului. De la publicarea raportului de climă al Grupului Interguvernamental pentru Schimbări Climatice (IPCC, 2018), dezbaterile privind amenințările actuale pentru mediu a fost mai activă decât oricând. Sunt necesare acțiuni urgente pentru a aborda provocările apărute, cum ar fi schimbările climatice, deficitul de resurse și plasticul oceanic.

Măsurile de reciclare a deșeurilor pot constitui un factor important în soluționarea problemei deficitului de resurse.

Orașele sunt elementul cheie pentru o abordare în spiritul principiilor economiei circulare a utilizării resurselor, deoarece mai mult de jumătate din populația lumii locuiește acum în orașe, iar până la jumătatea secolului, se așteaptă ca această pondere să crească la două treimi (United Nations, 2018). Deși deșeurile municipale au o pondere redusă în total deșeuri, accentul politic pus pe acestea este foarte mare din cauza caracterului lor complex, a distribuției între mai mulți generatori de deșeuri și a legăturilor cu modelele de consum (Eurostat, 2023). Atunci când sunt gestionate necorespunzător, deșeurile municipale dăunează mediului, au consecințe negative directe asupra calității vieții urbane și creează o povară financiară considerabilă asupra municipalităților (Biliewski, 2008). Un management adecvat al deșeurilor urmărește protejarea sănătății cetățenilor și izolarea substanțelor periculoase rezultate în diversele stadii din ciclurile materialelor, acordând în același timp o importanță deosebită reciclării.

Un alt aspect important este cel al colectării deșeurilor. Țările cu venituri mici colectează aproximativ 48% din deșeuri în orașe, dar această proporție scade drastic la 26% în afara zonelor urbane. De exemplu, Africa Subsahariană colectează aproximativ 44% din deșeuri, în timp ce Europa, Asia Centrală și America de Nord colectează cel puțin 90% (Kaza et al., 2018). După cum se observă, sunt discrepanțe majore între zonele specificate, un volum important de deșeuri rămânând necolectat în zonele cu venituri mai modeste. Finanțarea sistemelor de management al deșeurilor constituie o provocare semnificativă, mai ales din cauza costurilor operaționale. În țările cu venituri mari, costurile de operare pentru managementul integrat al deșeurilor, inclusiv colectarea, transportul, tratarea și eliminarea, depășesc în general 100 USD pe tonă. Țările cu venituri mai mici cheltuiesc mai puțin, în termeni absoluți, pe operațiunile cu deșeuri având costuri de aproximativ 35 USD pe tonă și întâmpinând mult mai multe dificultăți în recuperarea costurilor. Gestionarea deșeurilor necesită forță de muncă, deci cheltuieli pentru salarii, iar costurile de transport sunt între 20 și 50 USD pe tonă. Recuperarea costurilor pentru serviciile privind deșeurile diferă drastic în funcție de nivelul venitului specific fiecărui stat. Taxele pe care utilizatorii trebuie să le plătească variază de la o medie de 35 USD pe an în țările cu venituri mici, la 170 USD pe an în țările cu venituri mari, recuperarea totală sau aproape totală a costurilor fiind limitată, în mare parte, doar la țările cu venituri mari. Taxele pentru colectarea deșeurilor pot fi fixe sau variabile în funcție de tipul de utilizator care este facturat. De obicei, administrațiile locale acoperă aproximativ 50% din costurile investițiilor pentru sistemele de deșeuri, iar restul provine în principal din subvențiile guvernamentale naționale și din sectorul privat. La nivel global (Figura 1), statele care au produs cea mai mare cantitate de

deșeurile municipale sunt: SUA cu 265,2 milioane tone metrice, urmată de China cu 215,2 milioane tone metrice și Germania cu 53,7 milioane tone metrice (la nivelul anului 2021).

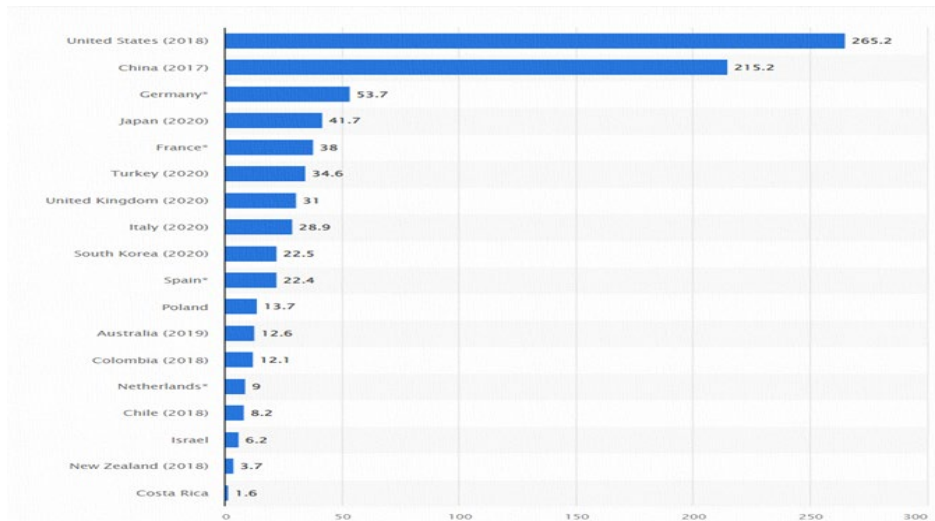


Figura 1. Situația statelor generatoare de deșeurile municipale la nivelul anului 2021

Sursa: Statista (2023)

1.1.2 Situația deșeurilor la nivelul Uniunii Europene

Politicile UE de gestionare a deșeurilor includ o serie de instrumente complementare de reglementare care vizează reducerea impactului asupra mediului. Directiva-cadru revizuită privind deșeurile (European Parliament and Council, 2018) impune statelor membre să atingă o rată de reciclare de 55% până în 2025, 60% până în 2030 și, în final, 65% până în 2035. Aceste cerințe fac parte dintr-o politică mai largă a UE în domeniul economiei circulare. Pentru a îndeplini aceste cerințe, Comisia Europeană sugerează stimulente economice ca o posibilă metodă de revigorare a reciclării (European Commission, 2011).

O consecință directă a acestei reglementări a fost aceea a creșterii cantității de deșeurile reciclate de la 87 kg pe cap de locuitor în 1995 la 257 kg pe cap de locuitor în 2021, cu o rată medie anuală de 4,3%. Ponderea totală a deșeurilor municipale reciclate pentru perioada 1995-2021 a crescut de la 19% la 49%.

Conform Agenției Europene de Mediu (European Environment Agency, 2022) taxele de depozitare a deșeurilor constituie instrumentul economic cel mai frecvent utilizat în UE pentru a îmbunătăți gestionarea deșeurilor municipale, 22 de state membre punându-le în aplicare. Aceste taxe sunt adesea folosite în combinație cu interdicțiile privind depozitarea anumitor tipuri de deșeurile.



Taxele de incinerare, un al doilea instrument economic utilizat frecvent în UE, sunt impuse doar de nouă state membre și sunt, în medie, stabilite la un nivel mult mai scăzut decât cel al taxelor de depozitare a deșeurilor. Nivelurile taxelor pentru depozitarea deșeurilor și incinerare variază foarte mult între statele membre, eficacitatea acestor taxe depinzând nu numai de nivelul lor, ci și de modul în care sunt concepute, implementate și aplicate.

Un al treilea instrument economic utilizat în UE este reprezentat de taxele de colectare a deșeurilor, frecvent concepute într-un mod care să stimuleze producătorii de deșeuri să reducă generarea de deșeuri și să își sorteze deșeurile, prin implementarea diverselor scheme ale sistemului „Plătești pentru cât arunci” (PPCA), construit pe principiul „Poluatorul plătește”.

Un al patrulea instrument major în managementul deșeurilor este reprezentat de sistemele de colectare separată a deșeurilor. Acestea necesită ca producătorii de deșeuri (cetățeni, entități publice sau private) să-și separe deșeurile la sursă în diverse materiale reziduale sau în diverse combinații de materiale care sunt apoi colectate separat pentru prelucrare ulterioară.

La nivelul Uniunii Europene, cantitatea medie de deșeuri municipale măsurate per capita a crescut în UE, din 2006 până în 2021. Totuși, tendințele variază de la țară la țară. De exemplu, deși cantitățile de deșeuri municipale per capita au crescut în majoritatea țărilor UE, acestea au scăzut în țări ca Malta, Cipru, Spania și România (Figura 2). Similar tendinței manifestate la nivel global, statele mai bogate tind să producă mai multe deșeuri per capita.

Conform datelor publicate pe site-ul Parlamentului European, în fiecare an 2,2 miliarde de tone de deșeuri sunt generate în UE. Circa 27% sunt deșeuri municipale colectate zilnic și tratate de municipalități, acestea fiind generate în principal de gospodării. Circa 39,2 % din deșeuri au fost reciclate și 32,2% depozitate la gropile de gunoi, în UE, în 2020.

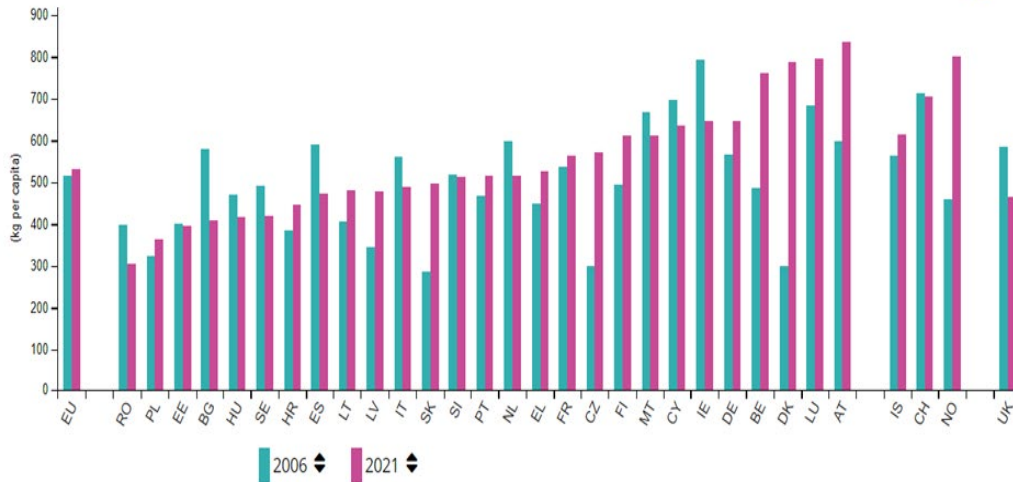


Figura 2. Situația generării de deșuri municipale per capita în Europa-anii 2006 și 2021

Sursa: Eurostat (2023a)

Datele evidențiază discrepanțe mari între țările UE în privința cantităților (Figura 3) dar și a modului în care sunt gestionate deșeurile, reglementările vizând cu precădere intensificarea reciclării și reducerea folosirii gropilor de gunoi.

Situația generării de deșuri în Europa la nivelul anului 2020

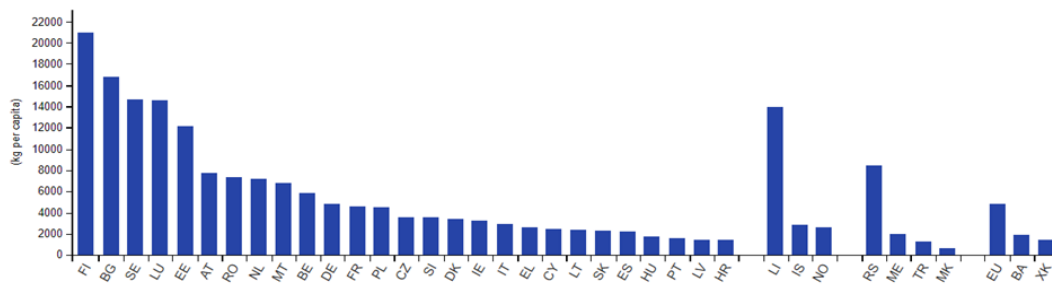


Figura 3. Situația generării de deșuri în Europa la nivelul anului 2020

Sursa: Eurostat (2023b).

O persoană din Uniunea Europeană a produs, în medie, o cantitate de 505 kilograme de deșuri municipale în 2020, cea mai scăzută cantitate fiind înregistrată de România, cu 287 kilograme per capita. Cantitatea de deșuri municipale generată de o persoană din UE a crescut în 2020 cu patru kilograme față de 2019 și cu 38 kilograme față de 1995. În 2020, în patru state membre ale UE, au fost înregistrate peste

600 de kilograme de deșuri municipale per capita, respectiv Danemarca și Luxemburg (ambele cu 845 kg per capita), Malta (643 kg per capita) și Germania (632 kg per capita).

În perioada 1995-2021, totalul deșeurilor municipale depozitate în UE a scăzut cu 55%, de la 286 kg/cap de locuitor în 1995 la 121 kg/cap de locuitor în 2021, respectiv o scădere medie anuală de 3,0 %. Rata depozitării deșeurilor în UE a scăzut de la 61% în 1995 la mai puțin de 23% în 2021. În termeni absoluți, deșeurile municipale per capita au înregistrat cele mai mari niveluri în Austria, Luxemburg, Danemarca și Belgia, în timp ce nivelurile cele mai scăzute au fost înregistrate în Spania, Letonia, Croația și Suedia.

Practicile de gestionare a deșeurilor variază de la o țară la alta în cadrul UE. UE dorește să promoveze cât mai mult posibil prevenirea generării deșeurilor și reutilizarea produselor. Dacă acest lucru nu este posibil, se preferă reciclarea/compostarea, urmată de utilizarea deșeurilor în generarea de energie. Cea mai dăunătoare opțiune pentru mediu rămâne însă aruncarea deșeurilor sau depozitarea acestora în gropile de gunoi, deși este și una dintre cele mai ieftine posibilități. Trebuie remarcat faptul că, în UE, chiar dacă nivelul deșeurilor generate pe cap de locuitor a crescut, modul în care se gestionează deșeurile s-a îmbunătățit prin reciclare și compostare și prin scăderea numărului de depozite pentru deșuri. Din păcate, România era în 2020 în coada clasamentului european la capitolul gestiunea deșeurilor (Figura 4).

Un alt aspect important este constituit de managementul deșeurilor municipale biodegradabile. Directiva 31/1999 privind depozitele de deșuri prevede obligativitatea reducerii de către statele membre ale UE a cantității de deșuri municipale biodegradabile care ajung în depozite la circa 75 % până la 16 iulie 2006, 50 % până la 16 iulie 2009, 35 % până la 16 iulie 2016 și la 10% până în 2035. Reducerea a fost calculată pe baza cantității totale de deșuri municipale biodegradabile produse în 1995. Această inițiativă a determinat țările să adopte diferite strategii pentru a evita trimiterea fracției organice a deșeurilor municipale la groapa de gunoi respectiv compostarea, incinerarea și pretraterea.

Colectarea separată eficientă a deșeurilor biologice, componenta dominantă a deșeurilor municipale, este esențială pentru a obține rate ridicate de reciclare: statele membre cu cele mai bune performanțe au sisteme de colectare a deșeurilor biologice foarte convenabile, spre deosebire de cele care au performanțe mai slabe. Toate cele cinci state membre ale UE cu cele mai mari rate de reciclare – Germania, Austria, Slovenia, Țările de Jos și Luxemburg – folosesc fie o taxă, fie interdicția depozitelor de gunoi, fie o combinație a acestora. De asemenea, o mare parte a populației lor beneficiază de facilități pentru colectarea separată a deșeurilor biologice și au sisteme de plată PPCA.

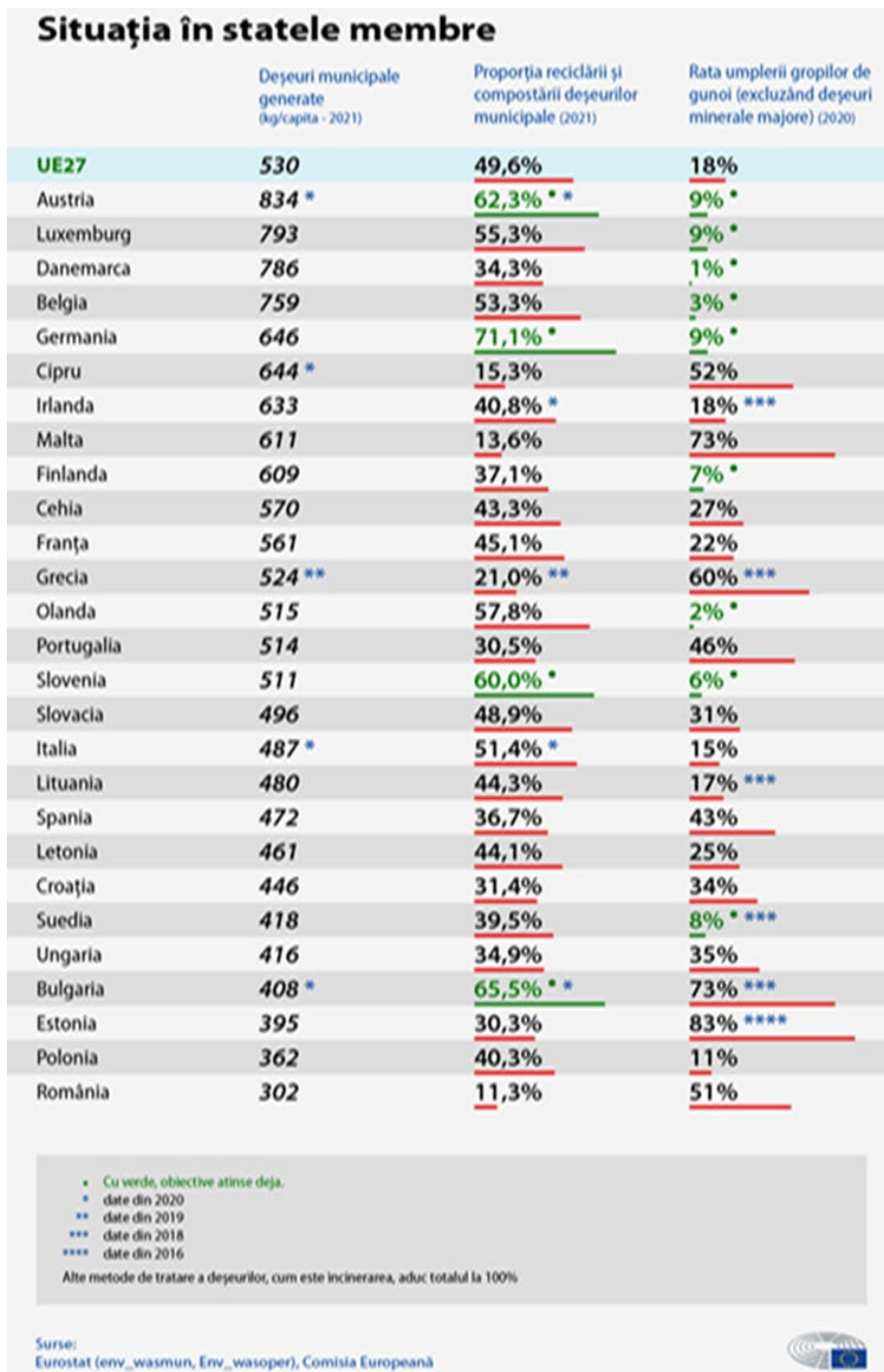


Figura 4. Infografic parțial privind situația deșeurilor în statele membre UE

Sursa:Parlamentul European(2023)

Regimul deșeurilor este gestionat în UE prin intermediul următoarelor reglementări:

- Directiva 2008/98/CE privind deșeurile și de abrogare a anumitor directive ce instituie un cadru legislativ pentru tratarea deșeurilor în UE. Cadrul este conceput pentru a proteja mediul și sănătatea populației accentuând importanța tehnicilor adecvate de gestionare, valorificare și reciclare a deșeurilor pentru a reduce presiunea asupra resurselor și a îmbunătăți utilizarea acestora. Directiva introduce obiective de reciclare și valorificare, care trebuiau atinse până în 2020 în ceea ce privește deșeurile menajere (50 %) și deșeurile provenind din activități de construcție și demolări (70 %);
- Directiva de modificare (UE) 2018/851, în cadrul unui pachet de măsuri privind economia circulară modifică Directiva 2008/98/CE. Stabilește cerințe minime de exploatare pentru scheme de răspundere extinsă a producătorilor. Consolidează regulile de prevenire a deșeurilor. De asemenea, stabilește noi obiective de reciclare a deșeurilor municipale: până în 2025, vor trebui reciclate deșeuri municipale în proporție de cel puțin 55 % din greutatea acestora. Acest obiectiv va crește la 60 % până în 2030 și 65 % până în 2035. Statele membre trebuie să instituie până la 1 ianuarie 2025 colectarea separată a textilelor și a deșeurilor periculoase generate de gospodării și să asigure faptul că, până la 31 decembrie 2023, biodeșeurile sunt colectate separat sau reciclate la sursă (de exemplu, prin compostare). Directiva evidențiază exemple de stimulente pentru aplicarea ierarhiei deșeurilor, cum ar fi taxele de eliminare a deșeurilor prin depozitare și taxele de incinerare, precum și schemele de plată în funcție de cantitatea de deșeuri generată;
- Directiva 850/2018 privind depozitele de deșeuri;
- Directiva 852/2018 a Parlamentului European și a Consiliului din 30 mai 2018 de modificare a Directivei 94/62/CE privind ambalajele și deșeurile de ambalaje;
- Regulamentul (UE) 2023/1542 al Parlamentului European și al Consiliului din 12 iulie 2023 privind bateriile și deșeurile de baterii, de modificare a Directivei 2008/98/CE și a Regulamentului (UE) 2019/1020 și de abrogare a Directivei 2006/66/CE;
- Directiva 904/2019 privind reducerea impactului anumitor produse din plastic asupra mediului (plasticul de unică folosință și nu numai) transpusă prin Ordonanța nr. 6/2021 privind reducerea impactului anumitor produse din plastic asupra mediului;
- Directiva 849/2018 de modificare a Directivei 2000/53/CE privind vehiculele scoase din uz, a Directivei 2006/66/CE privind bateriile și acumulatorii și deșeurile de baterii și acumulatori și a Directivei 2012/19/UE privind deșeurile de echipamente electrice și electronice;
- Directiva 19/2012 privind deșeurile de echipamente electrice și electronice;

- Directiva 66/2006 privind deșeurile din baterii și acumulatori,
- Directiva 20/2005 de modificare a Directivei 94/62/CE privind ambalajele și deșeurile de ambalaje.

Implementarea, aplicarea și executarea corectă a legislației UE privind deșeurile constituie priorități ale politicii de mediu în UE, după cum se menționează pe site-ul oficial al Comisiei Europene. Toate statele membre UE trebuie să raporteze Comisiei Europene cu privire la punerea în aplicare a legislației UE privind deșeurile, inclusiv cu privire la atingerea obiectivelor de colectare, reutilizare, reciclare și valorificare a deșeurilor în fiecare an sau din doi în doi ani.

În 2023, Comisia a publicat un raport (European Commission, 2023) care identifică statele membre care riscă să nu atingă obiectivele de reciclare a deșeurilor municipale și de ambalaje pentru 2025 și obiectivul de depozitare a deșeurilor pentru 2035.

Nouă state membre erau pe cale să atingă principalele obiective de reciclare pentru deșeurile municipale (provenite de la gospodării și întreprinderi) și deșeurile de ambalaje pentru 2025: Austria, Belgia, Cehia, Danemarca, Germania, Italia, Luxemburg, Țările de Jos și Slovenia.

Optsprezece state membre riscau să rateze una sau mai multe ținte: Bulgaria, Croația, Cipru, Estonia, Finlanda, Franța, Grecia, Ungaria, Irlanda, Lituania, Letonia, Malta, Polonia, Portugalia, România, Slovacia, Spania și Suedia. Comisia Europeană a prezentat recomandări și pune la dispoziție sprijin financiar și tehnic acestor state membre.

Directiva-cadru modificată privind deșeurile, Directiva privind ambalajele și deșeurile de ambalaje și Directiva privind depozitele de deșeuri au eliminat obligația de a raporta modurile de implementare. Acum, Comisia Europeană, sprijinită de Agenția Europeană de Mediu, are obligația să publice rapoarte de avertizare timpurie cu trei ani înainte de anii țintă de reciclare și depozitare.

1.1.3 Situația deșeurilor în context național

Toate rapoartele și statisticile ultimilor ani, publicate pe site-ul Parlamentului European concluzionează că una dintre problemele majore de mediu ale României o constituie managementul deșeurilor. Concret, deșeurile produse în marile orașe, dar și în micile comunități sunt depozitate de regulă în gropi de gunoi fără a fi sortate, reciclate și refolosite. Astfel, în momentul aderării la Uniunea Europeană, în România erau 2686 de gropi comunale de gunoi care ar fi trebuit să fie închise până în anul 2009. România ar fi trebuit să construiască 30 de facilități de management al deșeurilor concomitent



închizând 150 spații de depozitare a deșeurilor municipale și circa 1500 de spații de depozitare ilegale până în 2015 pentru a se conforma standardelor UE, lucru care nu s-a întâmplat.

De asemenea, este necesar ca România să își intensifice eforturile pentru atingerea obiectivelor UE în materie de reciclare. Conform Raportului de țară al României pentru anul 2022 (Comisia Europeană, 2022) progresele înregistrate sunt limitate, neexistând nicio creștere reală a gradului de reciclare și compostare. Potrivit datelor Eurostat, la nivelul anului 2020, gradul de reciclare în România era de numai 11%, media europeană fiind de 45%. România depozita 70% din deșeurile produse (272 kg/cap de locuitor), în timp ce media europeană era de 24%. În ceea ce privește reciclarea și compostarea deșeurilor, România este ultima în clasamentul european, cu 14%, media UE fiind de 46%. În 2021, România a adoptat 41 de planuri județene de gestionare a deșeurilor și planul de gestionare a deșeurilor pentru municipiul București și urmează să își revizuiască planul național de gestionare a deșeurilor. O altă inițiativă în domeniu este marcată de Planul național de redresare și reziliență (PNRR) al României ce cuprinde un plan de acțiune al strategiei naționale privind economia circulară și prevede investiții în domeniul deșeurilor. În prezent, este evident faptul că România trebuie să renunțe la practica depozitării deșeurilor și a tratării mecano-biologice (TMB) și să dezvolte o infrastructură de deșeuri cu o poziție superioară în ierarhia deșeurilor. Dată fiind situația, România ar putea ajunge să plătească amenzi pentru depozitele sale ilegale și neconforme de deșeuri.

Datele provizorii publicate de Eurostat pentru perioada 2020-2021 indică faptul că România a generat 302 kg de deșeuri municipale/cap de locuitor, comparativ cu media europeană de 503 kg/cap locuitor, în condițiile în care în UE cantitatea totală de deșeuri a crescut și deșeurile municipale depozitate la gropile de gunoi au scăzut. Problema majoră a României rămâne însă reciclarea și faptul că multe dintre gropile de gunoi sunt ilegale, neconforme sau supraaglomerate.

Dincolo de reciclarea deșeurilor menajere, România are o problemă reală cu importul de deșeuri și arderea acestora. Afacerile cu deșeuri, în special cu cele importate, au ajuns să reprezinte o afacere foarte profitabilă, estimată la aproape jumătate de miliard de euro. Potrivit datelor raportate de Registrul Comerțului și Ministerul Finanțelor, există peste 3300 de firme de profil ce au depășit 425 milioane de euro în 2020, în creștere cu 175 milioane de euro față de 2016. Se estimează că aceste afaceri ar depăși 1 miliard de euro, luând în calcul și evaziunea din domeniu. Sunt raportate numeroase afaceri care au profitat de lacunele legislative și au adus în România cantități uriașe de deșeuri, intrate ca produse second-hand.

Comisia Europeană a avertizat România de mai multe ori cu procedura de infringement, restricționându-i accesul la fondurile europene. De asemenea, există pe probleme de mediu și management al deșeurilor câteva procese împotriva României la Curtea Europeană de Justiție. În ciuda unor progrese modeste înregistrate până în anul 2020, România nu a reușit să atingă niciunul dintre obiectivele asumate prin Tratatul de Aderare la Uniunea Europeană referitoare la sistemul de management al deșeurilor. În 2020, Comisia Europeană a deschis împotriva României 40 de cazuri noi de constatare a neîndeplinirii obligațiilor la Curtea Europeană de Justiție, potrivit Raportului anual privind monitorizarea aplicării dreptului Uniunii Europene. Garda de Mediu a emis avertizări potrivit cărora, în curând, România nu va mai avea unde să depoziteze gunoiul, gropile fiind arhipline, deșeurile fiind nesortate corect și gradul de reciclare infim. Deși a avut posibilitatea de a accesa bani europeni, România nu a reușit acest lucru, cei care fac afaceri cu deșeuri preferând să plătească amenzile, decât să se conformeze cerințelor europene în domeniu.

Legislația deșeurilor, elaborată la nivel strategic prin directivele europene, este transpusă în România prin circa 80 de acte normative privind gestionarea deșeurilor. Dintre acestea, amintim:

- Legea nr. 17 din 6 ianuarie 2023 pentru aprobarea Ordonanței de urgență a Guvernului nr. 92/2021 privind regimul deșeurilor. Legea are ca obiect de reglementare transpunerea în legislația națională a Directivei 2008/98/CE a Parlamentului European și a Consiliului din 19 noiembrie 2008 privind deșeurile precum și abrogarea anumitor directive;
- Legea Nr. 132/2010 privind colectarea selectivă a deșeurilor în instituțiile publice;
- Legea 101/2006 privind serviciul de salubritate a localităților, actualizat prin Legea 99/2014, OUG 58/2016, OUG 172/2020, OUG 133/2022, OUG 30/2023;
- OM 82/2015 privind aprobarea Regulamentului-cadru al serviciului de salubritate a localităților modificat de Ordinul nr. 520/2018;
- Regulamentul-cadru al serviciului de salubritate a localităților din 09.03.2015;
- Legea Nr. 249/2015 privind gestionarea ambalajelor și deșeurilor de ambalaje;
- OM 794/2012 privind procedura de raportare a datelor referitoare la ambalaje și deșeuri de ambalaje;
- OUG 74/2018 pentru modificarea și completarea privind regimul deșeurilor, a Legii 249/2015 privind modalitatea de gestionare a ambalajelor și a deșeurilor de ambalaje și a OUG 196/2005

privind Fondul pentru mediu, aprobată prin Legea 31/2019 (modificată de OUG 92/2021 privind regimul deșeurilor);

- OUG 6/2021 privind reducerea impactului anumitor produse din plastic asupra mediului;
- OM 1271/2018 privind înregistrarea colectorilor la UAT-uri (unități administrativ teritoriale);
- OM 710/2022 pentru aprobarea Ghidului de finanțare a Programului privind ecologizarea zonelor afectate de deșeuri;
- OM 1647/2022 pentru aprobarea Normelor metodologice privind controlul transferurilor de deșeuri;
- OM 1736/2022 privind aprobarea Procedurii de înregistrare, raportare și declarare a operatorilor economici care introduc deșeuri în țară pentru a fi valorificate;
- OM 1595/2020 pentru aprobarea Instrucțiunilor de utilizare a aplicației informatice Sistemul informatic de asigurare a trasabilității deșeurilor (SIATD), în vederea monitorizării și verificării corectitudinii tranzacțiilor cu deșeuri de ambalaje în sistemul răspunderii extinse a producătorului;
- OUG 155/2020 pentru elaborarea Planului național de redresare și reziliență (PNRR);
- OUG 124/2021 pentru gestionarea fondurilor europene alocate prin Planul Național de Redresare și Reziliență (PNRR);
- HG 1172/2022 pentru aprobarea Strategiei naționale privind economia circulară.

Conform Raportului de țară pentru anul 2022 (Comisia Europeană, 2022), România pune integral în aplicare legislația Uniunii în domeniul deșeurilor, dar nu reușește să atingă obiectivele referitoare la reciclare, separarea deșeurilor și crearea unor sisteme separate pentru biodeșeuri.

Ce ar trebui să facă România pentru eficientizarea managementului deșeurilor? În mod cert, o soluție ar putea fi utilizarea unor tehnologii moderne și inovatoare care să deschidă calea pentru o mai bună monitorizare și pentru asigurarea unui control al deșeurilor mai eficace în comparație cu abordarea tradițională a colectării deșeurilor menajere. Abordări eficiente și exemple de "bune practici" în managementul deșeurilor sunt adeseori documentate în literatura de specialitate prin soluțiile implementate de așa-numitele "orașe inteligente" (în engleză, "Smart City").

1.2 Posibile abordări în managementul deșeurilor: „Orașul Inteligent”

Folosirea tehnologiei pentru o serie de elemente de administrare a zonelor urbane, asociază sistemele de monitorizare și control din managementul deșeurilor „orașului inteligent”. Un oraș inteligent reprezintă un concept de dezvoltare urbană care integrează tehnologii și sisteme moderne pentru administrarea eficientă și securizată a resurselor unui oraș, în vederea îmbunătățirii calității vieții cetățenilor, dezvoltării comunității și protejării mediului. Tehnologiile moderne integrate în administrarea orașelor inteligente vizează: sistemele de transport, sistemele energetice, sistemele de sănătate, rețelele de distribuție a apei, sistemele de management al deșeurilor, școli, biblioteci, agenții publice și alte servicii ale comunității.

În ceea ce privește managementul deșeurilor, conceptul înglobează circuitul complet al deșeurilor, de la colectarea separată la nivel de comunități prin sisteme inteligente și eficiente, cum sunt platformele subterane de colectare a deșeurilor, la transportul acestora cu vehicule „prietenos” cu mediul, complet electrice, până la crearea de stații de transfer, stații de sortare și de tratare mecano-biologică a deșeurilor (Radu, 2022). Într-un „Smart City”, toate aceste elemente sunt interconectate, eficientizând întregul proces, favorizând dezvoltarea economiei circulare.

În orașele inteligente, se presupune că identificarea și implementarea unor soluții concrete în lupta împotriva poluării mediului constituie obiective asumate benevol, măsurile de salubritate fiind vitale. Monitorizarea și controlul în timp real al containerelor de deșeuri amplasate în locații strategice și eliminarea acestor deșeuri la o destinație finală este esențială potrivit cercetărilor recente (Giffinger, 2017).

Un management eficient al deșeurilor presupune însă luarea în considerare și a altor aspecte operaționale și strategice: optimizarea traseelor vehiculelor de colectare, realizarea unor analize care să identifice eventuale trenduri în structura deșeurilor colectate (în scopul reciclării), identificarea perioadelor de vârf în colectarea deșeurilor pentru dimensionarea corectă a flotei de vehicule, monitorizarea nivelului de umplere a containerelor pentru a evita ridicarea ineficientă a deșeurilor etc. Toate aceste informații au ca obiectiv principal reducerea costurilor municipalității. Fără colectarea și eliminarea corespunzătoare a deșeurilor, pubelele/containerelor specifice prezintă risc de supraumplere și contaminare a împrejurimilor, ceea ce poate cauza probleme majore mediului. Reducerea nivelului poluării și asigurarea unui grad de salubritate corespunzător impun existența unui proces de monitorizare digitală a fluxurilor de deșeuri menajere.

În acest fel, tehnologia modernă devine o componentă esențială în abordarea provocărilor tot mai mari ale gestionării deșeurilor în orașele inteligente. Acest lucru este influențat și de cererea tot mai mare de soluții eficiente, ecologice și durabile de gestionare a deșeurilor. O tehnologie cheie în orașele inteligente este Internetul Lucrurilor (IoT) a cărui utilizare în sistemele de management al deșeurilor a oferit soluții notabile la această problemă.

Modalitatea principală prin care IoT își face simțită prezența este încorporarea de dispozitive și senzori inteligenți în peisajul urban, însărcinați cu monitorizarea și colectarea datelor privind deșeurile. Având o abordare centrată pe date, acest sistem permite gestionarea mai eficientă a colectării și eliminării deșeurilor, logisticii transportului și eforturilor de reciclare. Sistemul de management al deșeurilor poate utiliza tehnologii de urmărire și partajare a datelor integrate, optimizând astfel procesele specifice ciclului de viață al deșeurilor, cum ar fi colectarea, recuperarea și prevenirea generării de deșeuri.

Un sistem de management al deșeurilor (Figura 5) bazat pe IoT cuprinde coșuri inteligente de gunoi echipate cu senzori care le monitorizează nivelul de umplere, permițând gestionarea și analiza deșeurilor în timp real. Datele referitoare la aceste coșuri pot fi apoi transmise wireless către un hub central, facilitând colectarea promptă a deșeurilor și prevenind debordarea. Senzorii instalați în aceste sisteme măsoară și greutatea și nivelul deșeurilor din interiorul coșului de gunoi, comunicând aceste informații sistemelor de management al deșeurilor. Acest amalgam tehnologic, cu blockchain care oferă caracteristici de trasabilitate, transparență și audit într-o manieră descentralizată și sigură, se dovedește esențial în îmbunătățirea managementului deșeurilor în orașele inteligente.

Tehnologia de rețea cu rază lungă de acțiune (LoRaWAN) este o tehnologie cheie pentru arhitecturile IoT. Această tehnologie poate suporta diferite tipuri de servicii în cadrul sistemului inteligent de management al deșeurilor, inclusiv coșuri inteligente și containere de depozitare care interacționează cu utilizatorii precum și unități de supraveghere video echipate cu capacități de învățare automată ce pot fi utilizate, de exemplu, pentru detectarea unui eventual incendiu.

Tehnologia coșului inteligent folosește inteligența artificială pentru a redirecționa deșeurile în mod eficient, cu un minim consum de forță de muncă și cu cel mai redus grad de congestione a traficului. Un coș inteligent poate fi realizat prin combinarea mai multor tehnologii, precum Inteligența Artificială a Lucrurilor (AloT), viziunea computerizată, robotica, etichete RFID, comunicarea în câmp apropiat (NFC), senzori, rețele de senzori fără fir, compactoare, actuatoare, sisteme inteligente de monitorizare și tablouri de bord. Astfel, se pot primi direct informații despre materialele reciclabile prin integrarea coșului inteligent cu un tablou de bord pentru monitorizarea în timp real. Companiile care depozitează deșeurile

Își pot eficientiza astfel fluxurile de lucru și pot reduce timpul de verificare a coșurilor de gunoi prin utilizarea indicațiilor de monitorizare. O platformă inteligentă de analiză a coșurilor de gunoi permite unităților să efectueze auditul deșeurilor, să evalueze ratele de deviere și să estimeze potențialele economii rezultate din reducerea costurilor. Prin utilizarea coșurilor de gunoi inteligente, utilizatorii pot maximiza potențialul de reciclare pentru diverse deșeuri materiale. Ocazional, se întâmplă ca oamenii să nu își trimită deșeurile în coșurile adecvate sau să nu le sorteze corect pentru reciclare. Se poate utiliza un coș de gunoi inteligent care sortează automat materialele reciclabile în mai multe compartimente, utilizând identificarea obiectelor cu ajutorul inteligenței artificiale, într-un efort de a elimina sortarea eronată a materialelor reciclabile (exemplu: coșul inteligent Bin-e). După sortare, deșeurile sunt comprimate, iar sistemul monitorizează cât de plin este fiecare coș. Coșurile inteligente elimină erorile umane din procedura inițială de sortare, accelerând și simplificând procesarea materialelor pentru instalațiile de reciclare. Acest lucru poate crește foarte mult productivitatea lucrătorilor și poate reduce cheltuielile de gestionare a deșeurilor cu până la 63% (Dey, 2023).

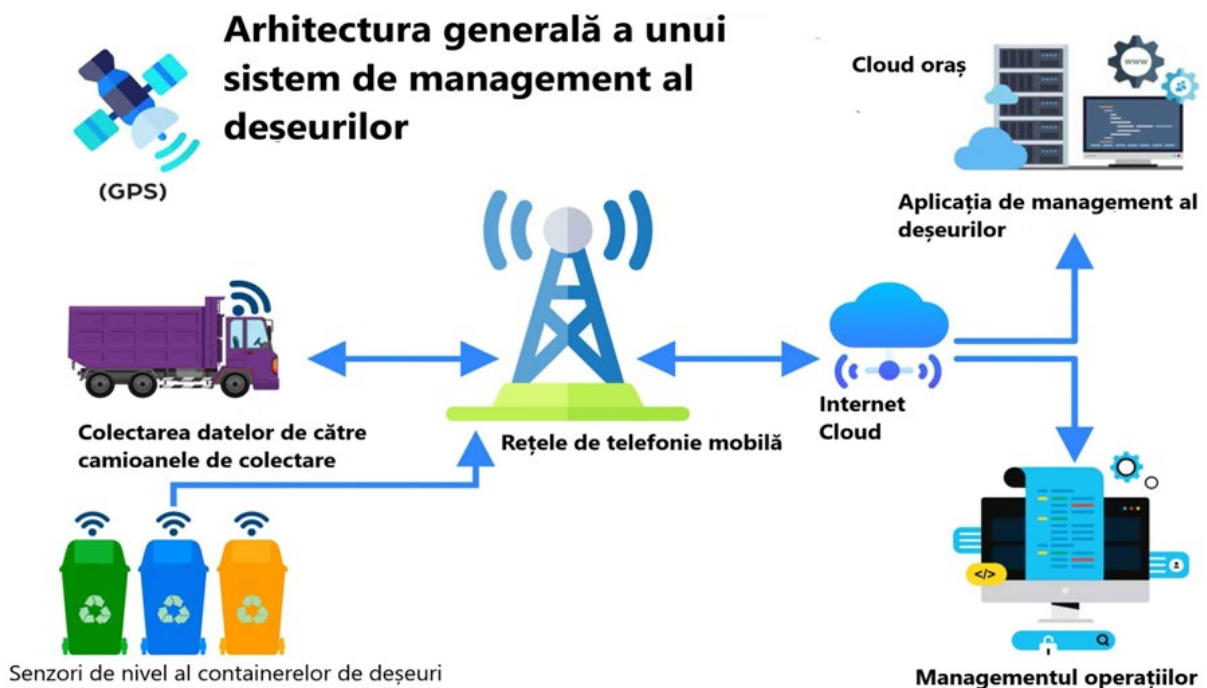


Figura 5. Arhitectura generală a unui sistem de management al deșeurilor

Sursa: <<https://mycity360.co.il/smart-waste-management-system/>>

Tehnologiile de sortare a deșeurilor bazate pe inteligența artificială utilizează echipamente și software care folosesc analiza imaginilor pentru a sorta materialele reciclabile cu o precizie și rată de recuperare mult mai mari decât cele specifice sistemelor convenționale. Există o multitudine de companii (exemplu: compania care, pentru a automatiza procesele de reciclare, folosesc AI și roboți. În prezent, în întreaga lume sunt sute de instalații de sortare a deșeurilor care utilizează această tehnologie. Împiedicând ca materialele reciclabile să ajungă în gropile de gunoi și făcându-le mai ușor de procesat și reutilizat, o eventuală extindere a utilizării tehnologiilor de sortare a deșeurilor bazate pe inteligența artificială va reduce volumul deșeurilor și va aduce beneficii mediului. Un exemplu de "bune practici" în acest sens este orașul San Francisco care, în anul 2009, a adoptat o ordonanță prin care toate întreprinderile și toți rezidenții erau obligați să sorteze deșeurile în trei categorii: materiale reciclabile, deșeuri compostabile și deșeuri pentru groapa de gunoi, devenind astfel primul oraș american care făcea compostarea obligatorie. În prezent, orașul își redirecționează aproximativ 80% din deșeurile din gropile de gunoi în fiecare an, lăudându-se cu una dintre cele mai mari rate de reciclare din SUA. Acest lucru a devenit posibil prin parteneriatul cu Recology, o companie de colectare a deșeurilor, care a cheltuit 20 de milioane de dolari pentru a-și moderniza instalațiile și a achiziționa roboți pentru sortarea eficientă și precisă a materialelor reciclabile.

Tehnologiile bazate pe senzori de nivel al deșeurilor permit utilizatorilor să vadă instantaneu cât de plin este fiecare coș de gunoi, aceștia putând să planifice din timp colectarea și să întreprindă acțiuni bazate pe datele obținute de la senzori. Angajații implicați în colectarea deșeurilor pot aranja modul în care se desfășoară activitatea, concentrându-se doar pe locațiile coșurilor de gunoi pline, cu ajutorul unei platforme de monitorizare live. Costul managementului deșeurilor pentru administrațiile locale poate fi foarte ridicat, colectarea deșeurilor fiind, uneori, cel mai costisitor articol bugetar pentru comunitățile din țările mai puțin bogate. De asemenea, costul managementului deșeurilor crește la nivel global, ceea ce are un impact deosebit de negativ asupra zonelor cu venituri mici. Atunci când sunt combinați cu o platformă de monitorizare a nivelului de umplere, senzorii de nivel al deșeurilor oferă soluții la problemele financiare prin reducerea cheltuielilor de colectare a gunoierii cu până la 50%, deoarece dacă scade numărul activităților de colectare a deșeurilor, vor scădea și costurile pentru timpul șoferului, combustibil și întreținerea vehiculului.

Este de așteptat ca sistemele de management al deșeurilor, în contextul extinderii numărului de proiecte specifice orașelor inteligente, să evolueze în mod constant. Piața globală de Smart City reacționează într-un ritm rapid, cu o medie anuală a creșterii previzionată a deveni 13,8% până în 2026,

când se estimează că va atinge 873,7 miliarde de dolari, comparativ cu 457 de miliarde la sfârșitul anului 2021. Abordările Smart City în marile orașe ale lumii sunt o prezență constantă, potrivit clasamentului realizat de Nexus Integra în 2022, primele 10 orașe inteligente din lume fiind: Londra, New York, Paris, Tokyo, Reykjavík, Copenhaga, Berlin, Amsterdam, Singapore și Dubai.

În ceea ce privește implementările concrete ale conceptului Smart City în România, s-a depășit faza proiectelor pilot, fără a avansa însă la proiecte consistente de Smart City. Potrivit raportului Vegacomp Consulting (2022), România este cu 2-3 ani în urmă față de media europeană, și a pierdut startul, deoarece multe dintre proiectele-pilot nu au fost utilizate sau nu au fost continuate de administrațiile locale sau centrale. Exemple în acest sens sunt proiectele-pilot abandonate din Alba Iulia, dar și proiectele implementate de Telekom România în 2016 – 2017 în sectorul 4 și în Constanța. Potrivit Vegacomp Consulting (2022) topul orașelor inteligente era condus, în anul 2022, de Cluj-Napoca cu 63 de proiecte, Iași cu 56 de proiecte și primăria municipiului București cu 54 de proiecte. În ceea ce privește inițiativele smart, primăriile de sector și primăria generală a municipiului București ocupă, cumulativ, primul loc la nivel național, cu 115 proiecte.

Tabel 1-Topul SMART CITY al orașelor din România

Nr.	Oraș	Total proiecte	Smart Economy	Smart Mobility	Smart Environment	Smart People	Smart Living	Smart Governance
		1001	130	322	59	35	217	238
1	Cluj Napoca	63	5	23	8	2	13	12
2	Iasi	56	4	18	13	2	9	10
3	Bucuresti	54	1	30	2	0	15	6
4	Alba Iulia	49	14	9	2	3	14	7
5	Bucuresti Sector 4	35	2	10	1	1	10	11
	Sibiu	35	5	13	3	2	3	9
6	Oradea	26	1	9	1	0	10	5
	Timisoara	26	5	9	1	1	6	4
7	Arad	25	1	5	1	3	7	8
	Slanic Moldova	25	6	5	3	1	4	6
8	Bistrita	20	5	9	0	1	0	5

Sursa: Vegacomp Consulting (2022)

Literatura de specialitate abundă în studii privind soluțiile inteligente și demonstrează beneficiile incontestabile rezultate din apelarea la tehnologie în gestiunea deșeurilor municipale. Un exemplu de astfel de sistem destinat comunităților urbane este BigBelly, containerul inteligent de colectare separată folosit deja în orașe din SUA, Europa sau Asia. Bigbelly este un sistem inteligent bazat pe energie solară ce utilizează tehnologia senzorilor pentru a compacta deșeurile, o stație Bigbelly procesând de 6 până la 8

ori mai multe deșeuri decât o pubele standard. Stația notifică echipajele atunci când este gata pentru colectare printr-o platformă online numită CLEAN (Smart City Magazine, 2016). Bigbelly reduce volumul colectărilor de deșeuri cu peste 80% și facilitează creșterea eficienței economice în gestionarea deșeurilor. Zonele publice sunt menținute mai curate și mai sigure, deoarece stațiile Bigbelly sunt complet sigilate, dăunătorii neavând acces la deșeuri alimentare. În România, un astfel de sistem a fost implementat inițial în sectorul 3 al capitalei, în urma proiectului-pilot inițiat de GreenPoint Management și Coca Cola HBC. Acest proiect are, la ani distanță de la plasarea primelor stații, rezultate notabile ce continuă să se vadă în randamentul procesului de colectare separată, în diminuarea costurilor și în reducerea amprenteii de carbon pe raza sectorului 3. Ulterior, pubele BigBelly au fost plasate și în orașele Iași și Brașov.

În principiu, containerele inteligente pot monitoriza și detecta următoarele elemente: depășirea nivelului de umplere, ora exactă a colectării deșeurilor, deșeurile care nu sunt colectate în containerele dedicate, rutele de colectare și comunicarea directă cu șoferul.

Un alt proiect-pilot Smart City pe problema managementului deșeurilor a fost derulat în municipiul Oradea, în anul 2022. Sistemul de salubritate din Oradea a fost dotat cu cinci țarcuri de deșeuri și un sistem de acces și monitorizare care a folosit inteligența artificială. Proiectul pilot a avut scopul de a stimula colectarea separată a deșeurilor și, totodată, de a ajuta asociațiile să distribuie echitabil costul facturii de salubritate. Fiecare sac prezenta un cod unic RFID, care conținea informații despre apartamentul căruia i-a fost repartizat și tipul de deșeu pe care trebuia să-l conțină, fiind necesară scanarea codului RFID pentru a obține acces în țarc. Printr-o aplicație specială, folosită de asociația de proprietari și de poliția locală, s-au putut semnala situațiile în care regulile au fost încălcate.

Un alt aspect important în managementul deșeurilor ține de importanța datelor colectate. Sistemele unui oraș-inteligent dispun de capacități de analiză a datelor menite să susțină procesele decizionale. Ca în orice alt domeniu, pentru a gestiona eficient problema deșeurilor sunt necesare cunoștințe adecvate. Pentru a crea cunoștințe semnificative, utile în fundamentarea deciziilor, este necesară analiza unui volum mare de date, care să acopere diversele aspecte ale activității. Asigurarea avantajului competitiv este obținută prin înțelegerea semnificației datelor și prin capacitatea de a anticipa evoluția faptelor plecând de la aceste date.

Analiza avansată a datelor și tablourile de bord în timp real solicită un volum important de date pe care organizațiile îl colectează și stochează. Potrivit Forbes, la nivelul anului 2018, 90% din datele lumii fuseseră generate în perioada 2016-2018, iar conform Forumului Economic Mondial se așteaptă să fie generați peste 463 de exabytes pe zi până în anul 2025.

În contextul generării unor volume foarte mari de date, dificil de gestionat, se impune utilizarea unor tehnologii specifice Big Data, termen sinonim cu extragerea, manipularea și analiza unor seturi de date care sunt prea mari pentru a fi prelucrate în mod obișnuit, prin mijloacele tradiționale. Aceste date pot fi extrase pentru a obține informații și pot fi utilizate ulterior în proiecte de învățare automată, modelare predictivă și alte aplicații avansate de analiză.

Soluția tehnică pe care am avut-o în vedere în construirea unei platforme dedicate procesului de monitorizare digitală a colectării deșeurilor poate fi asociată conceptului de Smart City fiind fundamentată pe conceptul de Big Data și pe tehnologiile specifice. Platforma-prototip realizată de echipă a urmărit în primă fază identificarea celui mai potrivit model de arhitectură Big Data care să permită importul datelor referitoare la colectarea deșeurilor menajere din bazele de date ale unităților administrativ-teritoriale. În acest scop, echipa de implementare a realizat un așa-numit Data Lake și ulterior, un Data Warehouse pentru ca aceste date să poată fi utilizate în generarea unor analize și crearea unor tablouri de bord care să susțină procesul decizional.

Obiectivul principal al acestei platforme l-a constituit îmbunătățirea mecanismelor de gestionare a instrumentului economic "Plătește pentru cât arunci" (în engleză, „Pay-as-you-throw”), ale cărui principii vor fi prezentate în cele ce urmează.

1.3 Sistemul de colectare a deșeurilor PPCA (PAYT)

1.3.1 Principii generale

În Uniunea Europeană, dezbaterile privind sistemele de plată PPCA s-a bazat pe aplicarea a două orientări combinate ale politicilor comunitare și internaționale de mediu: principiul „poluatorul plătește” (PPP- Polluter Pays Principle) și principiul responsabilității partajate. În acest sens, Directiva 75/442/CE, modificată prin Directiva 91/156/CE în 1991, în articolul 15, a stabilit că, în conformitate cu principiul „poluatorul plătește”, costurile de eliminare a deșeurilor trebuie să fie acoperite de către deținătorul deșeurilor manipulate de un colector de deșeuri, de deținătorii anteriori sau de către producătorul produsului din care derivă deșeurile. Prin urmare, considerentele directivei 91/156/CE introduc ideea că o parte din costuri nu este acoperită din încasările de tratare a deșeurilor, ci trebuie plătită în conformitate cu principiul „poluatorul plătește”. Aceeași formulare a fost apoi transpusă în articolul 15 din Directiva 2006/12/CE. Deci, sistemul „pay-as-you-throw” este un instrument economic care aplică principiul

„poluatorul plătește” prin taxarea locuitorilor în funcție de cantitatea reală de deșeuri pe care o generează. Principiile PPCA se bazează pe trei factori principali (Bilitewski, 2008):

- Identificarea producătorului de deșeuri;
- Calculul cantității de deșeuri generate;
- Prețul unitar pentru taxarea individuală în conformitate cu serviciile solicitate sau furnizate.

Atunci când este combinată cu o infrastructură bine dezvoltată și un nivel bun de conștientizare, sistemul PPCA se dovedește performant prin creșterea cantității de materiale reciclate. Mai multe studii au relevat faptul că sistemul PPCA prezintă un efect semnificativ de reducere a deșeurilor (Reichenbach, 2008; De Jaeger și Eyckmans, 2015).

Într-un sistem PPCA care funcționează optim, taxele pentru deșeuri sunt formate dintr-o componentă fixă a taxei plus o componentă variabilă care reflectă structura costurilor de gestionare a deșeurilor. Taxa este utilizată pentru a alina stimulentele pentru utilizatori (adică taxă mai mică atunci când se produc mai puține deșeuri) și pentru colectorii de deșeuri (venituri stabile din componenta fixă).

Deși Directiva 851/2018/UE prevede că sistemele de plată PPCA taxează „producătorii de deșeuri pe baza cantității efective de deșeuri generate”, nu impune ca taxa să fie în întregime calculată pe această bază. A calcula „pe baza cantității efective de deșeuri generate” nu înseamnă că taxa trebuie să fie strict și exclusiv proporțională cu cantitatea reală de deșeuri generate și nici că fiecare fracție de deșeuri colectată trebuie să fie măsurată și introdusă în calcul (Foundation Operate, 2019). Trebuie reținut faptul că implementarea schemei PPCA nu este unitară la nivelul UE și nici chiar în interiorul unui stat.

O abordare realistă a taxei include, potrivit Foundation Operate (2019):

- O parte fixă, care nu depinde de cantitatea de deșeuri măsurată, ci de alți parametri ce evidențiază practicile de generare a deșeurilor pentru anumite categorii de utilizatori, respectiv tipul de proprietate pe care utilizatorii o ocupă, suprafața și utilizarea acesteia, numărul de persoane care ocupă imobilul (pentru gospodăria), volumul containerelor puse la dispoziție utilizatorilor etc. Dacă taxa vizează și finanțarea costurilor legate de deșeurile externe, este evident că acestea vor trebui luate în considerare în determinarea părții fixe.
- O parte variabilă, raportată la cantitatea efectivă de deșeuri livrată.

Costurile pentru crearea și întreținerea instalațiilor de gestionare a deșeurilor (reciclare – reutilizare și eliminare) din diversele scheme de colectare a deșeurilor nu depind direct de cantitățile reale de deșeuri generate. Chiar și în situația în care scad considerabil cantitățile de deșeuri generate, facilitățile

construite ar avea nevoie de finanțare în continuare. Un astfel de sistem are avantaje evidente pentru indivizi chiar și în situațiile în care aceștia nu generează cantitatea minimă de deșuri care trebuie reciclate sau eliminate sau nu generează deloc deșuri. Presupunând că aceștia, în mod excepțional, ar genera deșuri, știu că le pot livra serviciului de eliminare a deșeurilor. Această posibilitate justifică o contribuție fixă pentru acoperirea costurilor.

Abordarea „minimalistă” a taxei respectă toate cerințele prevăzute în directiva 851/2018/UE: în primul rând, se referă la o schemă de plată PPCA în ceea ce privește deșeurile mixte (și eventual, alte fracții semnificative); în al doilea rând, oferă stimulente masive pentru colectarea separată (care, la minimum, nu implică nicio taxă suplimentară) și, în paralel, oferă reduceri la deșeurile mixte.

Pentru a concepe o taxă cu o astfel de structură, este fundamental să se determine:

- componentele de cost care trebuie acoperite din veniturile provenite din partea fixă a taxei;
- factorii/parametrii (a,b,...m) care pot fi utilizați ca referință pentru calcularea părții fixe și în ce măsură aceștia influențează calculul acesteia.

Din această perspectivă, rezultă că:

Taxa = partea fixă + partea variabilă

Partea fixă = f (a, b ... m)

Partea variabilă = p₁·q₁ + p₂·q₂ + ... p_n·q_n,

unde p_n reprezintă prețul unitar, iar q_n reprezintă cantitățile de deșuri colectate.

Chiar dacă pentru o gospodărie nu există cantități de deșuri livrate pentru colectare (q₁ = q₂ = q_n = 0), taxa nu este egală cu zero, ci se achită la nivelul cotei fixe.

După cum am specificat anterior, dacă nu este convenabilă măsurarea și calcularea fiecărui tip de deșuri livrat de utilizatorul individual, este posibil să se evalueze doar câteva categorii semnificative de deșuri și, cel mult, să se măsoare cantitatea de deșuri mixte generate de utilizatorul individual. În acest caz, taxa ar fi formulată după următorul binom:

Taxa= f (a, b ... m)+p_iq_i,

unde p_i, respectiv q_i reprezintă prețul unitar, respectiv cantitatea de deșuri mixte livrate.

Din punct de vedere practic, sistemul PPCA poate fi implementat sub diferite forme, numite scheme:

- scheme bazate pe volum, în care taxele de deșuri sunt percepute în funcție de dimensiunea containerelor golite;

- scheme bazate pe saci, în care taxele pentru deșeuri sunt percepute în funcție de numărul de saci de deșeuri utilizați, colectarea realizându-se doar pentru deșeurile eliminate în saci specifici preplățiți;
- scheme bazate pe greutate în care taxele pentru deșeuri sunt percepute în funcție de greutatea deșeurilor colectate într-un anumit container;
- scheme bazate pe frecvența colectării în care se percep taxe pentru deșeuri în funcție de frecvența cu care un container este lăsat afară pentru colectare. Această abordare poate fi combinată cu scheme bazate pe volum și greutate.
- scheme bazate pe un mix al schemelor specificate anterior.

Sistemele PPCA sunt proiectate în diverse moduri. Agenția Europeană de Mediu (European Environment Agency, 2022) clasifică sistemele PPCA aplicate în statele membre ale UE în sisteme elementare (de bază) și sisteme avansate, cele din urmă fiind considerate mai eficiente în influențarea comportamentului producătorilor de deșeuri decât sistemele elementare, prin faptul că oferă un stimulent economic mai puternic pentru sortarea deșeurilor.

Sistemele elementare PPCA sunt sisteme bazate pe volum care iau în calcul în stabilirea taxei de colectare dimensiunea containerului și uneori și frecvența de colectare. În cazul acesta, gospodăriile pot alege numărul sau dimensiunea containerelor pentru deșeurile municipale mixte atunci când acest lucru este convenit în contractul de servicii. În UE, sistemul PPCA este un instrument economic utilizat în mod obișnuit. Majoritatea statelor membre au deja implementat un sistem PPCA pentru cel puțin o parte a populației și au introdus o legislație care impune utilizarea/dezvoltarea sistemelor PPCA sau care permite municipalităților să introducă astfel de sisteme. Paisprezece state membre utilizează o combinație de sisteme PPCA avansate și elementare, iar alte șase utilizează numai sisteme PPCA elementare. Niciun stat membru nu utilizează doar sisteme avansate de PPCA. Trei dintre statele membre care nu utilizează în prezent un sistem PPCA au planuri ferme de a implementa unul. Toate sistemele PPCA necesită identificarea utilizatorilor individuali, de exemplu prin adoptarea de cipuri electronice pe containerele de deșeuri și prin măsurarea fluxurilor de deșeuri colectate la nivel de utilizator individual. În plus, sistemele PPCA necesită definirea unui preț unitar al colectării deșeurilor care să conducă în mod eficient la modificarea comportamentului rezidenților către reducerea generării de deșeuri și o separare mai bună pe tipuri de deșeuri.

Tabel 2. Situația acoperirii cu sisteme PPCA pentru gospodării în EU-27, 2022

Tipul de sistem PPCA				
		Mix Avansat/Elementar	Elementar	Fără PPCA
Grad acoperire populație prin sistem PPCA	Înalt	Austria Belgia Croația Irlanda Slovenia	Finlanda Ungaria Suedia	
	Mediu	Danemarca Lituania Luxemburg România	Estonia	
	Scăzut	Cehia Germania Franța Țările de Jos Slovacia	Spania Italia	
Planuri de implementare a sistemului PPCA				Cipru Grecia Malta
Fără planuri de implementare sistem PPCA				Letonia Polonia Portugalia

Sursa: European Environment Agency (2023)

Indiferent de schema PPCA adoptată, este esențial ca sistemul să fie completat de o infrastructură de colectare eficientă și ușor de utilizat pentru fracțiile colectate separat. Implicarea rezidenților în înțelegerea corectă a caracteristicilor schemei PPCA este absolut necesară pentru a evita aruncarea ilegală de deșuri sau transferul deșeurilor în alte teritorii nedeservite de o schemă PPCA, deoarece au fost raportate și astfel de activități ca efect negativ al implementării acestui sistem (Berglund, 2005; Thogersen, 2003). Stimulentele economice pot submina morala și motivația intrinsecă a unui individ și pot încuraja "turismul deșeurilor" (mutarea deșeurilor în comunitățile învecinate) sau depozitarea ilegală. În plus, cantitățile crescute de contaminanți din materialele reciclabile reprezintă o potențială amenințare. În cele din urmă, se poate ajunge la situația în care pot apărea costuri suplimentare atât pe partea investițiilor, cât și pe partea operațională.

1.3.2 Adoptarea sistemului PPCA în statele membre

Sistemul PPCA nu este adoptat într-o viziune unitară la nivelul statelor membre ale UE. Vom evidenția în cele ce urmează cele mai notabile experiențe de implementare a schemelor Pay-as-you-throw

(PPCA), aplicate în unele state membre, schemele PPCA fiind concepute ca instrumente economice pentru minimizarea cantității de deșeurii municipale generate. La nivelul anului 2022, 20 de state membre UE aplicau diverse scheme PPCA (Figura nr. 6) pentru gestionarea deșeurilor municipale, deși numai unele dintre ele aplică sistematic PPCA în toate municipiile. Procentul de populație implicată/deservită de sistemele PPCA este extrem de variabil și poate fi doar estimat. Acesta poate varia de la o cotă limitată a populației (de exemplu, Spania aplică sistem PPCA numai în Catalonia), la 20% în municipalitățile din Țările De Jos și Italia, la 40% din populația din Luxemburg și apoi la aplicarea la scară națională (exemplu: Austria). O analiză a modului în care este înregistrată colectarea de deșeurii, evidențiază faptul că sistemul PPCA este în principal aplicat prin scheme bazate pe volumul deșeurilor și/sau frecvența colectării deșeurilor de la utilizatori, schemele care implică înregistrarea greutății deșeurilor fiind în minoritate. Mai multe state membre adoptă o abordare mixtă care utilizează moduri diferite de înregistrare a colectării de deșeurii. În funcție de statul membru, atât baza de impozitare, cât și suma de bani percepută de la utilizatori sunt variabile, ceea ce înseamnă că metodele de aplicare nu sunt comparabile a priori. Dintre cele 27 de state membre ale UE, doar 20 implementează diverse scheme PPCA în managementul deșeurilor (Figura 6).

Niciunul dintre cele cinci state membre cu cele mai scăzute rate de reciclare, între 11 și 27%, respectiv Malta, România, Cipru, Grecia și Portugalia, nu aplică o taxă sau o interdicție bine concepută pentru depozitarea deșeurilor. În majoritatea acestor state, sistemele de colectare separată a bio-deșeurilor trebuie îmbunătățite. Doar Malta a instituit foarte recent taxe concepute pentru a descuraja depozitarea deșeurilor. Cele cinci state membre cu cele mai mari rate de depozitare a deșeurilor nu au instalații de incinerare.

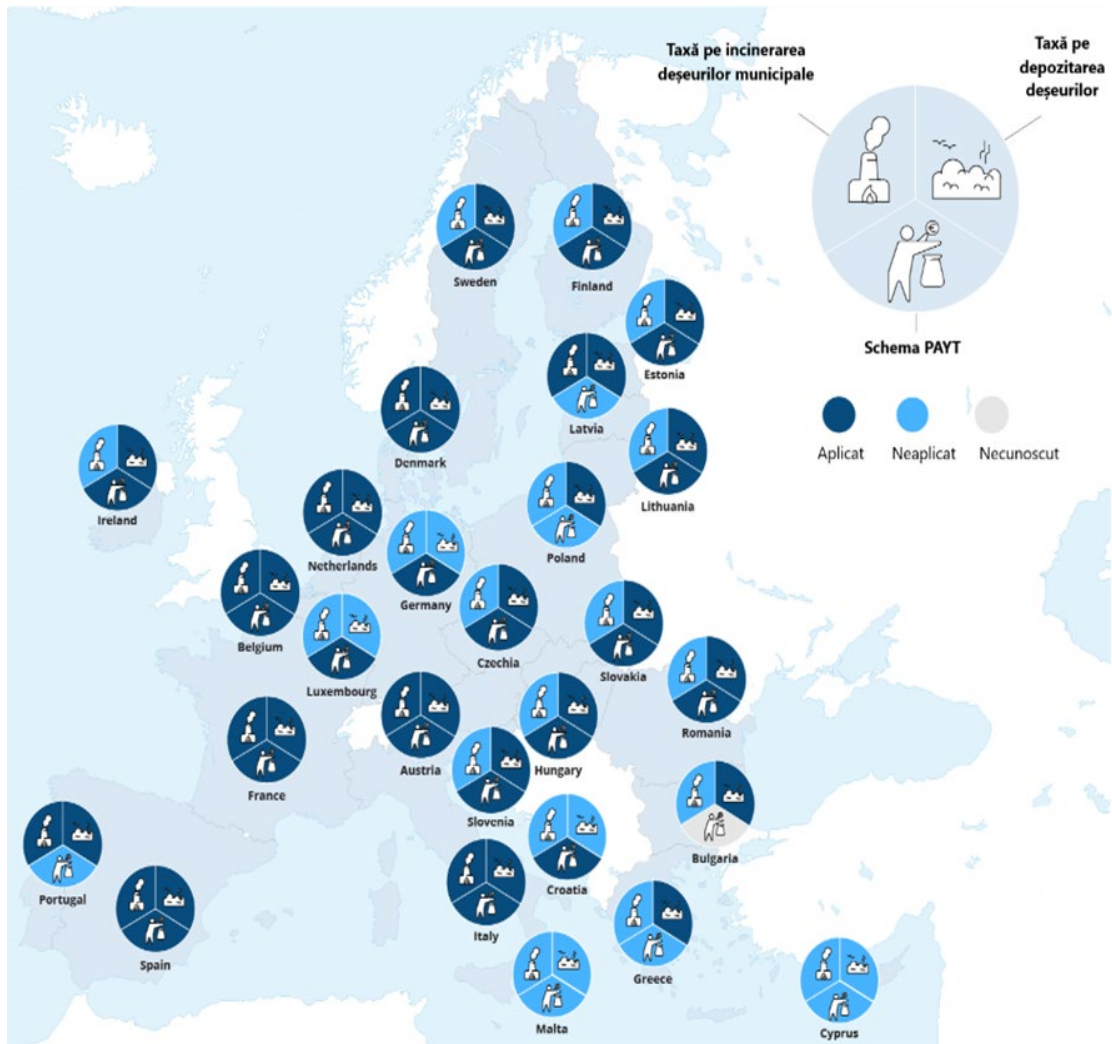


Figura 6. Instrumente cheie pentru managementul deșeurilor municipale aplicate de statele UE

Sursa: European Environment Agency (2023)

Analizele efectuate de Agenția Europeană de Mediu (2023) au relevat faptul că un stat membru poate obține performanțe bune în gestionarea deșeurilor fără utilizarea mai multor instrumente (de exemplu, Italia), dar e valabil și contrarul (de exemplu, Suedia). Acest lucru se poate datora faptului că în unele state membre sunt utilizate alte instrumente care stimulează performanța. De exemplu, Italia a stabilit obiective pentru fiecare municipalitate privind colectarea separată a deșeurilor municipale. Capacitățile mari de incinerare existente ar putea, de asemenea, concura cu eforturile de a crește reciclarea. În plus, eficacitatea instrumentelor depinde de proiectarea, implementarea, calendarul și aplicarea lor exactă.

1.3.3 Analiză comparativă a implementărilor schemei "Pay-as-you-throw"

Schemele Pay-As-You-Throw (PPCA) sunt utilizate de autoritățile locale din Belgia, Țările de Jos și Luxemburg într-un efort de a crește reciclarea și de a reduce deșeurile reziduale colectate din cadrul gospodăriilor. Există metode diferite de aplicare a schemelor PPCA (Card și Schweitzer, 2016), partea din taxă aferentă alegerii rezidenților fiind legată de:

- Mărimea recipientului ales de gospodărie;
- Frecvența de colectare a unui container dat;
- Aplicarea unei taxe pe sac utilizat;
- Greutatea deșeurilor stabilite pentru colectare;
- O combinație a celor de mai sus.

În multe cazuri, elementul variabil se aplică doar colectării deșeurilor reziduale. În altele, colectarea deșeurilor biologice poate fi de asemenea taxată, dar pot fi percepute taxe și pentru colectarea de materiale reciclabile, de obicei mult mai reduse decât pentru deșeurile reziduale. Faptul că există o gamă largă de sisteme, conduce la o orientare către anumite opțiuni care să fie utilizate mai mult în anumite zone decât în altele (de exemplu, colectările bazate pe sac ar putea fi o alegere mai probabilă în zonele urbane, în timp ce colectarea deșeurilor biologice va fi taxată cu precădere în zonele rurale).

Implementarea schemelor PPCA la nivel local sau regional, uneori ca răspuns la îndrumări de la un nivel administrativ superior, a dat naștere unui mozaic de modele diferite de sisteme în diversele jurisdicții (Linderhof et al., 2001; Hogg et al., 2002; Hill et al., 2002). În Olanda, un sistem PPCA a fost introdus pentru prima dată în Oostzaan în 1992; în Luxemburg, în Koerich și Kopstal a fost pilotată o schemă PPCA din 1994 până în 1997; în Belgia au fost testate scheme-pilot în Flandra, la începutul anilor 1990, înainte de adoptarea pe scară largă în 1995. Aceste scheme folosesc mecanisme și taxe diferite, dar toate sistemele urmăresc să descurajeze utilizarea containerelor pentru deșeurile reziduale. De exemplu, sistemul regional Ghent PPCA din Flandra se bazează în zonele urbane și suburbane pe prețul diferențiat al sacilor de deșeuri reziduale, deșeuri reciclabile și deșeuri biologice. În zonele rurale, se aplică un sistem de taxare al rezidenților pentru colectarea deșeurilor, cu taxe mai mari pentru deșeurile reziduale decât pentru deșeurile biologice (OCDE, 2006). Sistemul pilot din Koerich și Kopstal, Luxemburg, a avut o abordare ușor diferită, taxele variind în funcție de greutatea deșeurilor colectate și de volumul

containerului de deșuri reziduale folosit, în timp ce materialele reciclabile au fost colectate gratuit (Hogg, 2002), scheme similare fiind implementate și ulterior în Luxemburg (OCDE, 2010).

Schemele PPCA sunt fundamentate pe nevoia sau dorința de a reduce generarea de deșuri reziduale, dar și pentru a imprima un ritm ascendent sortării deșeurilor la nivel de gospodărie. De exemplu, în Flandra, schemele PPCA sunt parțial reglementate de guvernul regional, care stabilește tarifele minime și maxime pe care autoritățile locale le pot taxa pentru colectarea deșeurilor reziduale. Măsurile PPCA au fost introduse pentru a combate problema tot mai mare a gestionării deșeurilor în regiunea dens populată Flandra, acestea fiind însoțite și de o decizie ministerială de prevenire a înființării de noi gropi de gunoi din 1993 (EEA, 2009) care a ajutat la obținerea sprijinului public.

În Flandra, sistemul PPCA are o conexiune puternică și cu alte instrumente fiscale, cum ar fi taxele de depozitare a deșeurilor și taxele de incinerare, care împreună formează un pachet de instrumente concepute pentru a promova o mai bună gestionare a deșeurilor. În Valonia, mai multe municipalități au introdus schemele PPCA pentru a se asigura că nu li se cere să plătească o taxă pentru exces de deșuri reziduale, care urma să se aplice acelor municipalități în care deșeurile reziduale per locuitor ar fi depășit o anumită cotă.

Schemele din toate națiunile Benelux sunt legate de asemenea de sistemele de responsabilitate extinsă a producătorului (în special de Fost Plus în Belgia, Valorlux în Luxemburg și Nedvang în Țările de Jos), care colectează o parte din deșeurile menajere pentru reciclare.

În unele zone, în special în care coordonarea regională a fost mai slabă decât în Flandra, au existat mai multe bariere în calea implementării schemei PPCA. S-a putut remarca o creștere a eliminării ilegale a deșeurilor, deși alte studii au constatat că acest efect este supraestimat (Hogg et al., 2006). Alte bariere au inclus evitarea taxelor de către persoanele care călătoresc în zone care nu implementează o schemă PPCA pentru eliminarea deșeurilor, deși, din nou, amploarea acestui comportament este mică în comparație cu impactul general pozitiv al PPCA (Linderhof et al., 2001). De asemenea, poate exista un dezacord cu privire la reglementarea PPCA între autoritățile naționale și regionale. De exemplu, în Luxemburg, au existat dezacorduri cu privire la introducerea legislației care transpune Directiva-cadru revizuită a UE privind deșeurile (2008/98/CE), care urmărea să introducă mai multe reglementări privind taxele asociate colectării deșeurilor. Syvicol (sindicat ce reprezintă orașele și comunele din Luxemburg) nu a fost de acord cu intenția Guvernului de a introduce noi reglementări care impun tarife diferențiate pentru gestionarea deșeurilor (Europaforum, 2011). Syvicol a fost îngrijorat de faptul că nu au fost luate costurile în considerare în mod corespunzător la elaborarea legislației și, de asemenea, s-a opus la

impunerea unui model de taxare de la guvernul central. Cu toate acestea, legislația de transpunere a fost adoptată, introducând reguli mai stricte pe baza schemei PPCA în Luxemburg.

Schemele PPCA sunt considerate mecanisme de recuperare (parțială) a costurilor care stimulează plătitorul de taxe să adopte un comportament mai ecologic. Veniturile obținute prin taxe variabile sunt de obicei mai mici decât costurile de gestionare a deșeurilor municipale, astfel încât veniturile sunt de obicei suplimentate cu alte taxe. De exemplu, fondurile strânse de PPCA în Flandra echivalează cu doar aproximativ 50% din fondurile necesare pentru gestionarea corespunzătoare a deșeurilor. Deoarece furnizorii de servicii trebuie să recupereze costurile furnizării serviciilor, poate fi un risc să se bazeze exclusiv pe veniturile din taxe variabile pentru a genera nivelul dorit de venituri. Este tipic ca veniturile din taxe variabile să acopere aproximativ 30-50% din costuri pentru a nu expune furnizorul de servicii la problema instabilității veniturilor.

În țările din Benelux, rezidenții plătesc taxe direct, fie cumpărând saci la un preț stabilit, fie plătind colectarea deșeurilor din containere în funcție de greutate, frecvență sau dimensiune direct autorității locale (Hogg et al., 2011). Vehiculele și containerele pentru deșeuri pot fi proiectate în asemenea mod încât vehiculele să nu golească containerele gospodăriilor care nu au facturile achitate. Acest lucru reduce problemele legate de eventuale datorii neplătite și oferă siguranța că nu există rău-platnici în schemă.

Conform studiului publicat de Foundation Operate (2019), în Austria, taxa pentru colectarea deșeurilor este determinată în funcție de volumul containerului destinat colectării deșeurilor urbane reziduale și de frecvența de golire stabilită (ce poate fi săptămânală sau lunară). Acest lucru s-a putut realiza prin extinderea sistematică a colectării separate a deșeurilor alimentare și prin adoptarea compostării deșeurilor menajere în mai multe zone rurale. În multe zone urbane din Austria (de exemplu, Innsbruck) se aplică o taxă volumetrică atât pentru colectarea deșeurilor urbane reziduale, cât și pentru colectarea separată a deșeurilor biologice. În municipiile cu construcții verticale, sistemul PPCA se aplică la nivelul clădirii individuale și costul pentru apartamentele individuale este adesea calculat pe metru pătrat.

În Finlanda, taxele sunt alcătuite dintr-o parte fixă care depinde de tipul de utilizator și o parte variabilă, care acoperă deșeurile ce urmează a fi eliminate și accesul la punctele de colectare centralizate. Taxa pe deșeurile municipale se bazează, în principal, pe cantitatea de deșeuri urbane colectată și pe frecvența colectării acestuia.

În Germania, legislația federală nu impune aplicarea sistemului PPCA. Cu toate acestea, acest lucru este cerut indirect de legislațiile mai multor regiuni, conform cărora taxele aferente deșeurilor ar

trebuie să promoveze prevenirea generării de deșeuri și valorificarea acestora (de exemplu, regulamentele privind deșeurile din Berlin și Saxonia). Aplicarea efectivă a sistemului PPCA include scheme volumetrice, scheme bazate pe frecvența colectării (în principal a deșeurilor urbane reziduale și uneori a deșeurilor alimentare), precum și scheme în care se înregistrează greutatea deșeurilor livrate. În majoritatea cazurilor, la colectare se aplică o schemă bazată pe măsurare volumetrică sau pe frecvența de colectare a deșeurilor urbane reziduale. Prin urmare, mărimea taxei anuale aplicate utilizatorilor individuali este determinată în general de volumul containerului instalat la sediul utilizatorului și de frecvența de golire.

În Irlanda, legislația națională prevede aplicarea sistemului PPCA încă din anul 2005. Taxarea se realizează în funcție de volumele de deșeuri menajere generate de utilizatori, în funcție de frecvența de golire care este înregistrată prin etichete instalate pe containere și în funcție de greutatea deșeurilor. Privatizarea furnizorilor de servicii de colectare a deșeurilor urbane a condus la dezvoltarea unor relații contractuale specifice și directe cu utilizatorii casnici și, în consecință, la aplicarea extinsă a sistemului PPCA.

Sistemul PPCA în Italia a fost introdus ca parte a Normei Quadro sui rifiuti în 1997, dar este aplicată doar de către municipalități individuale, consorții de municipalități și de câteva orașe mai mari ca Bolzano, Parma, Mantua, etc. La nivelul anului 2019, se estima că aproape 3 milioane de italieni sunt deserviți de sistemul PPCA. Aplicațiile PPCA din această țară recurg aproape întotdeauna la măsurarea volumetrică a deșeurilor reziduale urbane (care, în scopul cuantificării, sunt convertite în kg). În plus, în ultimii zece ani, au fost implementate pe scară largă dispozitive pentru a detecta golirea containerelor și pentru a identifica containerele prin utilizarea etichetelor RFID în combinație cu poziția GPS a locației de golire a containerului. Sistemul PPCA se aplică în principal circuitelor de colectare „din ușă în ușă”, cu realizarea unei colectări separate sistematice a deșeurilor alimentare. În acele cazuri sporadice, în care taxa este bazată pe greutatea deșeurilor, colectarea se realizează în puncte centralizate, care sunt uneori echipate cu dispozitive de reducere volumetrică.

În urma analizei întreprinse, devine evident faptul că implementarea unui sistem PPCA este un demers unic, specific fiecărei națiuni, care ține cont de particularitățile naționale ale procesului de colectare, dar și de alți factori: economic, politic, social. Gradul de conștientizare a populației asupra particularităților colectării, modificarea percepției negative privind reciclarea (care este văzută, adeseori, ca un efort suplimentar în etapa de livrare a deșeurilor) constituie elemente ce pot asigura succesul unei scheme PPCA.

1.3.4 Concluzii

Creșterea populației și a volumului global de deșuri determină omenirea să exploreze noi tehnologii pentru o gestionare eficientă și adecvată a deșeurilor. În plus, ca tendință recentă, orașele inteligente accelerează inovațiile în managementul deșeurilor. Sistemele inteligente de gestionare a deșeurilor remodelează orașele, redefinesc sustenabilitatea, susținând principiile unei economii circulare. În urma analizei întreprinse, am constatat că eficacitatea și durabilitatea sistemelor tradiționale de colectare a deșeurilor sunt afectate de diverși factori, mai ales în mediul urban, de o perioadă foarte lungă de timp. Provocările care apar pentru sistemele tradiționale de colectare a deșeurilor provin din ineficiență, lipsă de responsabilitate și lipsa de monitorizare în timp real. Concluziile la care am ajuns în urma analizei situației actuale a deșeurilor și a sistemelor tradiționale de colectare a deșeurilor impun ca necesitate adoptarea unui sistem modern bazat pe tehnologie care să asigure monitorizarea în timp real a fluxurilor de deșuri menajere, această necesitate decurgând din problemele sesizate la nivelul sistemelor tradiționale.

Unul dintre cele mai semnificative dezavantaje ale sistemelor tradiționale de colectare a deșeurilor este dependența de procese manuale și metode învechite. Rutele de colectare sunt adesea planificate pe baza datelor istorice, fără a lua în considerare factorii în timp real, cum ar fi schimbările în densitatea populației, evenimente speciale intervenite sau închiderea drumurilor. Ca rezultat, camioanele de colectare a deșeurilor pot urma rute suboptimale, ceea ce duce la pierderi de timp, combustibil și emisii crescute.

În plus, evidența manuală a activităților de colectare a deșeurilor este predispusă la erori, ceea ce face dificilă urmărirea cu acuratețe a performanței sistemului.

Responsabilitatea este o altă problemă majoră cu privire la sistemele tradiționale de colectare a deșeurilor. Cele mai multe procese fiind efectuate manual, devine dificil să se identifice responsabilitatea pentru ridicările ratate de deșuri, serviciile întârziate sau manipularea necorespunzătoare a deșeurilor. Această lipsă de responsabilitate nu numai că afectează eficiența colectării deșeurilor, ci contribuie și la nemulțumirea beneficiarilor serviciilor de colectare a deșeurilor.

Separarea adecvată a deșeurilor este un aspect crucial al gestionării deșeurilor, deoarece permite reciclarea și eliminarea adecvată a diferitelor tipuri de deșuri. Sistemele tradiționale au probleme serioase privind separarea deșeurilor. La români, există o percepție negativă privind reciclarea și un grad scăzut de conștientizare privind necesitatea efectuării unei sortări adecvate a deșeurilor. Fără monitorizare în timp real, este posibil ca personalul de colectare a deșeurilor să nu cunoască respectarea



regulilor de separare adecvată de către rezidenți. Ca urmare, materialele reciclabile pot ajunge la gropile de gunoi, împiedicând eforturile de reciclare și conducând la rispa de resurse valoroase.

În plus, lipsa monitorizării în timp real în sistemele tradiționale de colectare a deșeurilor limitează sever capacitatea de a colecta date valoroase pentru planificare și optimizare. Tiparele de generare a deșeurilor se pot schimba în timp, iar autoritățile de gestionare a deșeurilor au nevoie de date exacte pentru a-și adapta strategiile de colectare în consecință. Fără acces la date în timp real, identificarea tendințelor, optimizarea rutelor și alocarea eficientă a resurselor devine dificilă. Integrarea tehnologiei și adoptarea de soluții inteligente de gestionare a deșeurilor pot revoluționa procesele de colectare a deșeurilor. Acest lucru le va face mai durabile, mai rentabile și mai receptive atât la nevoile comunității, cât și ale mediului.

Acestea sunt doar o parte din motivele pentru care construirea unei platforme de monitorizare a fluxurilor menajere și similare devine o necesitate pentru managementul adecvat al deșeurilor. Numeroase studii din literatura de specialitate subliniază că este imperios necesar să existe o soluție de monitorizare în timp real a întregului proces, din etapa de colectare a deșeurilor și până la eliminarea acestora.

2 Tehnologii suport

2.1 Sisteme de gestiune a bazelor de date

Baza de date constituie o componentă indispensabilă soluțiilor informatice de gestiune, prin urmare caracteristicile mediului de implementare, respectiv ale sistemului de gestiune a bazelor de date (SGBD) selectat pentru implementare, sunt determinate pentru maniera în care aplicațiile de acest tip reușesc să satisfacă un set specific de cerințe funcționale, dar și să răspundă unor cerințe nefuncționale care derivă din scenariile de utilizare a aplicației, ca și din necesitatea de a recurge la anumite tehnologii disponibile pentru implementarea acestora (spre exemplu, serviciile web).

Din perspectiva procesului de dezvoltare a unei soluții informatice, selectarea un anumit SGBD are loc în contextul adaptării soluției logice la particularitățile unui mediu software concret, prin tranziția de la o structură de date independentă de platformă (modelul logic al datelor) la o soluție de proiectare detaliată a structurii datelor, destinată implementării (modelul fizic al datelor). În funcție de structura bazei de date și evident, de particularitățile SGBD utilizat, o astfel de tranziție reclamă tratamente specifice, care vizează cu precădere următoarele aspecte:

- Definitivarea câmpurilor de date – În esență, reclamă stabilirea tipurilor de date pentru fiecare câmp din structura logică a bazei de date, precum și adăugarea de câmpuri de natură tehnică, pentru asigurarea condițiilor de funcționare curentă a bazei de date (spre exemplu, câmpuri de tip Timestamp pentru detectarea înregistrărilor care au fost modificate);
- Reprezentarea restricțiilor de integritate – Vizează definirea cheilor specifice modelului relațional și a condițiilor pe care trebuie să le respecte valorile memorate;
- Optimizarea fizică – Permite selectarea celor mai adecvate soluții de memorare și regăsire a datelor, considerând atât posibilitățile SGBD folosit, dar și caracteristicile sistemului informatic, în termenii volumului datelor și tranzacțiilor pe care trebuie să le gestioneze. Printre intervențiile uzuale la acest nivel pot fi menționate indexarea și partiționarea, respectiv secționarea orizontală sau verticală a tabelor; această operație se bazează pe analiza tranzacțiilor și urmărește creșterea performanței interogărilor, prin regruparea fizică a înregistrărilor sau câmpurilor prelucrate frecvent împreună.

Secțiunile următoare oferă o trecere în revistă a principalelor caracteristici a două SGBD reprezentative pentru administrarea și exploatarea bazelor de date relaționale, Microsoft SQL Server și MySQL, precum și a atuurilor specifice fiecărui SGBD, care au stat la baza deciziei echipei de implementare de a recurge la ambele produse software, în funcție de structura de date vizată la nivelul arhitecturii generale a sistemului informatic destinat analizei datelor privind colectările de deșeuri, în cadrul programului PPCA:

- Bazele de date aferente sistemului operațional – reprezintă structuri multiple, cu arhitectură identică, gestionate distinct la nivelul fiecărei unități administrativ-teritoriale (UAT);
- Structura de tip Data Lake, care permite colectarea datelor operaționale furnizate de fiecare UAT;
- Depozitul de date (data warehouse), care vizează cerințe specifice de analiză și raportare a datelor privind colectarea deșeurilor și care este populat periodic, pe baza datelor disponibile în structura de tip Data Lake.

Volumul datelor gestionate, volumul și natura tranzacțiilor, precum și scenariile de utilizare ale fiecăreia dintre structurile de date menționate, au reprezentat principalele elemente care au condiționat alegerea unui anumit SGBD, conducând în final la următoarele opțiuni:

- MySQL, pentru implementarea bazelor de date care oferă suport tranzacțiilor curente la nivelul fiecărei UAT care gestionează colectarea deșeurilor;
- Microsoft SQL Server, pentru implementarea structurilor de date care facilitează integrarea și analiza datelor de la nivel operațional (Data Lake, respectiv data warehouse).

În continuare este furnizată o prezentare succintă a celor două produse de tip SGBD, dar și o comparație relevantă pentru argumentarea opțiunii de a le folosi pe ambele, ca tehnologii complementare, pentru implementarea structurilor de date din cadrul soluției informatice care vizează colectarea deșeurilor prin programul PPCA.

2.1.1 Sistemul de gestiune a bazelor de date MySQL

În condițiile în care gestionarea și manipularea datelor bazată pe SQL beneficiază de suportul mai multor sisteme de gestiune a bazelor de date relaționale, principalul avantaj al MySQL este reprezentat de caracterul open-source. În plus, fiabilitatea, ușurința de utilizare, funcțiile puternice, securitatea și scalabilitatea sunt atribute recunoscute ale MySQL, care determină folosirea acestui SGBD pe scară largă în cadrul aplicațiilor web, împreună cu PHP și serverele Apache.

MySQL AB, compania care a dezvoltat MySQL, a fost achiziționată de Oracle Corporation în anul 2008. De atunci, Oracle a continuat să dezvolte și să susțină MySQL ca proiect open-source, acest SGBD fiind acum parte din familia de produse Oracle, alături de alte softuri pentru baze de date, precum Oracle Database și Oracle NoSQL Database. MySQL este considerat un instrument puternic și versatil, oferind suport pentru diferite sarcini de exploatare și administrare de baze de date, în funcție de aplicațiile-client și scenariile de utilizare. Aceste atuuri sunt confirmate de faptul că MySQL se regăsește în soluțiile tehnice ale multor site-uri web și aplicații extrem de populare, precum Facebook, Twitter, Netflix, Uber, Shopify și Booking.com. De asemenea, MySQL este utilizat pentru implementarea mai multor sisteme de management al conținutului (CMS), cum sunt WordPress, Drupal și Joomla. MySQL poate fi instalat pe diverse sisteme de operare (Linux, Windows, macOS) și poate fi configurat pentru a se potrivi mediului tehnic specific organizației care îl exploatează. Totodată, MySQL este perceput drept un SGBD flexibil și ușor adaptabil la nevoile dezvoltatorilor, oferind suport pentru diferite limbaje de programare (PHP, Java, Python etc.). În plus, caracterul deschis facilitează îmbunătățirile și contribuțiile aduse de comunitatea open-source la dezvoltarea MySQL. Atutul adaptabilității se poate constata și în planul costurilor, MySQL fiind disponibil atât în mod gratuit, cât și în ediții comerciale, ceea ce-l face adecvat pentru diferite cerințe și bugete.

Drept caracteristici esențiale pentru gestiunea datelor MySQL se pot menționa: asigurarea integrității datelor, scalabilitatea, disponibilitatea ridicată a datelor și opțiunile de securitate. De asemenea, MySQL permite definirea unor obiecte necesare oricărei soluții non-triviale de baze de date, precum procedurile stocate, declanșatorii sau vizualizările. Enumerarea de mai jos punctează cele mai importante atribute ale MySQL:

- Performanța – În condițiile în care performanța este determinantă pentru calitatea soluțiilor web, MySQL este conceput pentru a gestiona în mod eficient volume mari de

date și un număr ridicat de conexiuni simultane; performanța bazelor de date este consecința aplicării unor tehnici specifice, pentru care MySQL oferă suport dedicat:

- Indexarea – MySQL acceptă o varietate de tehnici de indexare, inclusiv indecși B-tree și hash;
 - Memorarea în cache – MySQL utilizează diverse tehnici de stocare în cache, care vizează, după caz, interogările, tabelele sau cheile;
 - Procedurile stocate – permit reducerea volumului de date transferate între bazele de date și aplicațiile-client, cu avantaje evidente în planul performanței.
- Fiabilitatea – MySQL este considerat un SGBD robust și stabil, dispunând de componente care vizează următoarele aspecte:
 - Replicarea datelor, pentru creșterea disponibilității acestora și diminuarea riscului pierderii;
 - Backup și recuperare, pentru crearea copiilor de siguranță și restaurarea bazelor de date în cazul producerii unor incidente;
 - Suport pentru tranzații, pentru gruparea logică a operațiilor de prelucrare a datelor, astfel încât fie toate se vor executa cu succes, fie toate vor fi anulate, în caz de eșec.
 - Scalabilitatea – MySQL este un sistem scalabil, permițând adaptarea la noi cerințe și scenarii de exploatare, ce includ creșterea volumului de date și a numărului de conexiuni simultane. Mecanismele care asigură scalabilitatea bazelor de date MySQL includ:
 - Partiționarea – dintr-un tabel unic rezultă subseturi de înregistrări ce pot fi gestionate de servere multiple;
 - Sharding – facilitează împărțirea datelor pe mai multe servere, pe baza unor chei specifice;

- Suport pentru clustering – permite crearea unui cluster de servere care conlucrează pentru realizarea prelucrărilor, cu beneficii importante în planul scalabilității și disponibilității datelor.
- Disponibilitatea datelor – Disponibilitatea ridicată a datelor MySQL este consecința implementării suportului dedicat pentru:
 - Replicare, prin crearea de copii ale bazelor de date și reducerea riscului de pierdere a datelor;
 - Suport cluster, pentru crearea clusterelor de servere;
 - Echilibrarea încărcării (load balancing), prin distribuirea sarcinilor de prelucrare pe mai multe servere, pentru îmbunătățirea performanței și a disponibilității datelor.
- Suportul pentru cloud computing – În condițiile în care această tehnologie oferă mai multe avantaje față de implementările tradiționale, în planul scalabilității și flexibilității, dar și al costurilor și rentabilității, MySQL poate fi utilizat în diverse medii cloud, inclusiv publice, cum ar fi Amazon Web Services (AWS), Google Cloud Platform (GCP) și Microsoft Azure, precum și în cloud-uri private sau de tip hibrid. Furnizorii de cloud pot oferi un mediu MySQL complet gestionat, incluzând backup-uri automate, actualizări de software și setări de securitate. La acestea se adaugă funcții precum scalarea automată, disponibilitatea ridicată și recuperarea în caz de dezastru. Utilizarea unui serviciu MySQL bazat pe cloud permite economii care vizează toate tipurile de resurse, eliminând necesitatea gestionării mediului MySQL. De asemenea, prezintă avantaje în planul securității, furnizorii de cloud asigurând protecția infrastructurii și datelor clienților. Selectarea furnizorului de cloud și a serviciului MySQL se realizează prin considerarea volumului și caracteristicilor tranzacțiilor, respectiv în baza cerințelor de performanță, disponibilitate, scalabilitate, cost și conformitate, specifice fiecărui beneficiar; în plus, este necesar ca aplicațiile care vor utiliza astfel de servicii să fie proiectate pentru a funcționa în mod eficient într-un mediu de tip cloud.

2.1.2 Sistemul de gestiune a bazelor de date Microsoft SQL Server

SQL Server a fost lansat în anii 1980 și, deși a suferit numeroase actualizări de-a lungul timpului, principiile de bază care guvernează acest SGBD relațional au rămas aceleași. Microsoft și Sybase au lansat versiunea 1.0 în 1989; deși acest parteneriat s-a încheiat la începutul anilor 1990, Microsoft a lansat ulterior mai multe versiuni ale SQL Server (2000, 2005, 2008, 2012, 2014, 2016, 2017, 2019, 2022).

Chiar dacă este compatibil cu standardul ANSI SQL (Structured Query Language), SQL Server implementează propriul dialect SQL, T-SQL (Transact-SQL), care adaugă limbajului diferite extensii (pentru declararea de variabile, gestionarea excepțiilor, definirea procedurilor etc). SQL Server Management Studio este aplicația folosită în mod uzual pentru administrarea SQL Server, prin intermediul unei interfețe grafice dedicate.

SQL Server este disponibil în mai multe ediții, diferite sub aspectul funcționalităților oferite și al costurilor pentru beneficiari, asigurând suport pentru o largă varietate de aplicații, de la cele mai simple (pentru care se poate recurge la ediția Standard) până la cele mai sofisticate, în care performanța, securitatea sau disponibilitatea datelor reprezintă aspecte critice (impunându-se instalarea ediției Enterprise). De asemenea, edițiile SQL Server vizează diferite contexte de utilizare, ca server care deservește aplicațiile aflate în exploatarea unei anumite organizații (edițiile Enterprise, Standard, Web) sau ca server ce oferă suport pentru învățare sau pentru procesele de dezvoltare și testare de aplicații (edițiile Developer și Express). O comparație detaliată a funcționalităților și opțiunilor specifice fiecărei ediții (Enterprise, Standard, Web, Developer, Express) este disponibilă în documentația tehnică a SQL Server, pe site-ul Microsoft.

SQL Server vizează cu precădere dezvoltatorii ce folosesc tehnologia Microsoft .NET pentru implementarea a diferite tipuri de aplicații (desktop, web, mobile) care necesită utilizarea unei baze de date și accesul securizat la aceasta. Dintre caracteristicile esențiale ale SQL Server în planul stocării și prelucrării datelor, pot fi menționate: posibilitatea definirii tipurilor de date complexe, controlul tranzacțiilor, indexarea avansată, asigurarea unei disponibilități înalte a datelor și diferite opțiuni de securitate. Funcționalitățile SQL Server constituie apanajul unui set de componente dedicate pentru diverse tipuri de servicii furnizate clienților conectați la server, cele mai importante dintre acestea fiind menționate mai jos:

- **Database Engine** (motorul bazelor de date) este o componentă esențială, care permite crearea și gestionarea bazelor de date, execuția interogărilor și securizarea datelor. Aceste funcții sunt susținute de mai multe servicii auxiliare:
 - Service Broker – Oferă suport pentru comunicarea asincronă între server și clienți. Aceștia din urmă transmit cereri către server și continuă derularea a diferite prelucrări, fără a aștepta răspunsul de la server; într-o astfel de abordare, brokerul are rolul de a asigura procesarea cererilor atunci când serverul devine disponibil.
 - Servicii de replicare – Permit copierea și distribuirea datelor și obiectelor create la nivelul bazelor de date de la un server la altul, cu beneficii în planul performanței (se facilitează accesarea rapidă a datelor, în funcție de localizarea serverelor) și al securității accesului la date.
 - Căutare pe text (full-text search) – Oferă suport pentru căutarea rapidă și inteligentă în baze de date de mari dimensiuni, prin identificarea înregistrărilor care conțin anumite cuvinte și expresii.
 - Servicii de notificare – Permit generarea și trimiterea de mesaje către utilizatori sau administratori, cu privire la evenimente predefinite.
 - **SQL Server Agent** – Permite programarea execuției anumitor prelucrări sau sarcini administrative (spre exemplu, realizarea copiilor de siguranță ale bazelor de date), acestea fiind declanșate în mod automat, conform programării.
- **Integration Services** (servicii de integrare) – Vizează colectarea datelor din surse multiple și eterogene, în vederea integrării acestora într-o structură de date unică, după aplicarea prealabilă a tuturor transformărilor necesare (conversii de date, derivări și centralizări etc.) Scenariul tipic de utilizare al SQL Server Integration Services este reprezentat de popularea depozitelor de date, pe baza proceselor ETL (Extraction-Transformation-Loading / Extragere-Transformare-Încărcare).
- **Analysis Services** (servicii de analiză) – Realizează procesarea datelor din depozitele de date stocate pe server, într-o abordare multidimensională; sunt astfel posibile centralizări

și analize în timp real ale datelor numerice (măsurii ale activității), prin raportarea la reperele descriptive specifice dimensiunilor (criterii de analiză ale măsurilor).

- **Reporting Services** (servicii de raportare) – Oferă un set de instrumente pentru definirea și gestionarea rapoartelor în diferite formate, pe baza datelor disponibile pe server.

Amplul set al funcționalităților specifice SQL Server este prezentat pe larg în documentația oficială Microsoft, dar și în multe alte resurse furnizate de comunitatea software. Întrucât o prezentare extinsă excede cadrul documentului de față, mai jos sunt enumerate doar câteva caracteristici care au contribuit la utilizarea pe scară largă a acestui SGBD pentru implementarea soluțiilor și arhitecturilor software complexe:

- Scalabilitate – Datele stocate în tabele de mari dimensiuni pot fi partiționate în mai multe grupuri de fișiere, serverul având capacitatea accesării simultane a acestor grupuri.
- Integritate cu mediul .NET, prin componenta CLR (Common Language Runtime) – Codul implementat în orice limbaj pentru care .NET Framework oferă suport poate fi salvat sub forma procedurilor stocate, executându-se pe server, similar codului T-SQL.
- Orientarea spre servicii – SQL Server Service Broker facilitează implementarea aplicațiilor complexe, care implică arhitecturi distribuite și interacțiuni asincrone.
- Nivel înalt de securizare a serverului și bazelor de date, ca rezultat al unor controale și mecanisme dedicate, care acționează pe mai multe niveluri (utilizatori individuali, roluri, procese).
- Management bazat pe politici – Vizează definirea unui set de politici pentru configurarea și gestionarea instanțelor SQL Server sau a obiectelor de la nivelul unui server (baze de date, tabele, indecși etc).
- Mecanisme specifice pentru gestionarea volumului de lucru al serverului, prin alocarea și managementul resurselor sistemului (SQL Server Resource Governor).

Integrarea SQL Server - Microsoft Azure a condus la apariția unei familii de produse (Azure SQL) care permit ca serviciile de gestionare a datelor, specifice SQL Server, să devină disponibile în cloud. Abordarea PaaS (Platform as a Service) oferă accesul la un mediu software în care majoritatea sarcinilor

de administrare (upgrade, backup, monitorizare etc.) sunt gestionate în mod automat, fără intervenții explicite din partea utilizatorilor.

Cea mai recentă versiune SQL Server (2022) aduce o serie de perfecționări mecanismelor care vizează securitatea, scalabilitatea și disponibilitatea datelor, menite să determine creșterea siguranței în exploatarea bazelor de date. Dintre atributele acestei versiuni se pot menționa:

- Îmbunătățirea performanței cu ajutorul Intelligent Query Processing, fără a se impune ajustarea codului T-SQL;
- Parameter Sensitivity Plan (PSP) permite stocarea în cache a mai multor planuri de execuție pentru aceeași interogare parametrizată;
- Feedbackul privind estimarea cardinalității (Cardinality Estimation - CE) oferă optimizerului de interogări informații detaliate despre metricile de execuție a interogărilor, facilitând actualizarea corespunzătoare a planurilor de execuție;
- Posibilitatea de a accesa și utiliza alte medii de stocare prin intermediul PolyBase REST API, care oferă suport pentru noi scenarii de integrare cu sistemele externe;
- Noi extensii aduse limbajului T-SQL.

2.1.3 Asemănări și diferențe între MySQL și SQL Server

Asemănări între MySQL și SQL Server

Întrucât ambele SGBD vizează bazele de date relaționale, oferind un important set de funcționalități pentru managementul adecvat și exploatarea eficientă a unor astfel de structuri de date, atât MySQL, cât și SQL Server, reprezintă opțiuni extrem de populare pentru alegerea unui SGBD. În pofida originii separate și a asocierii cu producători software diferiți, MySQL și SQL Server prezintă o serie de similitudini prin prisma cărora ambele SGBD pot fi considerate soluții viabile pentru implementarea bazelor de date ce deservește diverse tipuri de aplicații. Aceste asemănări sunt trecute în revistă în paragrafele următoare:

- Utilizarea limbajului SQL – Ambele SGBD se bazează pe SQL (Structured Query Language), un limbaj standardizat de interogare a bazelor de date relaționale și de gestionare a obiectelor specifice unei astfel de structuri de date (tabele, indecși etc.).

- Managementul tranzacțiilor – Ambele SGBD garantează integritatea datelor și derularea în siguranță a prelucrărilor complexe, prin raportarea la reperele ACID (Atomicity, Consistency, Isolation, Durability).
- Tipuri de date și indecsi – Cele două SGBD dispun de tipuri de date asemănătoare pentru memorarea valorilor text, numerice sau calendaristice; de asemenea, oferă suport pentru definirea indecșilor de diferite tipuri (primari, unici, full-text).
- Suportul pentru proceduri stocate, declanșatori, vederi – Pot fi definite în cazul ambelor SGBD, permițând modularizarea și reutilizarea codului, dar și îmbunătățirea performanței, prin reducerea volumului datelor transferate între bazele de date și aplicațiile-client.
- Securizarea accesului la date – SQL Server și MySQL furnizează diferite mecanisme de protecție a datelor, precum autentificarea utilizatorilor, controlul accesului și drepturilor pe baza rolurilor, diferite opțiuni de criptare a datelor memorate pe disc sau transferate între server și clienți.
- Scalabilitatea și disponibilitatea datelor - Ambele SGBD oferă suport pentru scalare orizontală și verticală, pentru a face față creșterii potențiale a volumului datelor sau numărului de utilizatori și de tranzacții. Totodată, pentru a se asigura un nivel ridicat de disponibilitate a datelor, există suport dedicat pentru replicare și clustering.

Diferențe între MySQL și SQL Server

În pofida asemănărilor menționate mai sus, MySQL și SQLServer se deosebesc în mai multe privințe, care pot determina opțiunea pentru un anumit SGBD, în funcție de un context particular de implementare și exploatare. Principalele diferențe sunt punctate mai jos:

- Compatibilitatea cu diferite platforme – MySQL poate rula pe diferite sisteme de operare, opțiunea uzuală fiind totuși reprezentată de Linux. SQL Server a fost inițial destinat platformei Windows. Începând cu anul 2017, SQL Server poate rula și pe Linux; pentru a rula pe macOS necesită software adițional, precum Docker. În mod evident, Windows rămâne platforma cea mai adecvată pentru SQL Server.

- Limbajul SQL – Deși SQL reprezintă fundamentul ambelor SGBD, acestea folosesc dialecte distincte ale standardului ANSI SQL. SQL Server utilizează T-SQL (Transact SQL), o extensie a SQL, care oferă suport pentru programarea procedurală (incluzând recursivitatea la variabile locale, la funcții pentru diferite tipuri de prelucrări etc.) Drept urmare, utilizarea efectivă a limbajului SQL impune cunoașterea și aplicarea regulilor de sintaxă specifice unui anumit mediu software.
- Compatibilitatea cu diferite limbaje de programare – Ambele SGBD pot fi utilizate pentru sisteme de baze de date accesabile prin aplicații implementate în majoritatea limbajelor uzuale (C#, C++, Java, Visual Basic, PHP, Python etc.), în condițiile în care acestea permit conexiuni bazate pe ODBC. Pe de altă parte, în cazul soluțiilor implementate în mediul Microsoft .NET, SQL Server constituie opțiunea cea mai adecvată, întrucât componenta ADO.NET vizează cu predilecție SQL Server, respectiv facilitarea accesului la bazele de date și serverele de acest tip. În plus, integrarea SQL Server - CLR (Common Language Runtime) permite codului .NET să se execute pe server sub forma procedurilor stocate.
- Performanța – Constituie un atu important pentru ambele SGBD, care permit exploatarea simultană a mai multor baze de date disponibile pe același server. Indecșii pentru acces rapid reprezintă un mecanism uzual pentru îmbunătățirea performanței, atât în cazul bazelor de date SQL Server, cât și al bazelor de date MySQL. Pe de altă parte, analizele asupra timpului de răspuns evidențiază un ușor avantaj în favoarea SQL Server, atât în mediul Windows, cât și Linux.
- Securitatea – MySQL și SQL Server sunt compatibile EC2, conformându-se standardelor de securitate pentru cloud computing. Pe de altă parte, există diferențe semnificative în privința accesului la bazele de date. În MySQL, fișierele aferente bazelor de date sunt disponibile pentru acces și modificare și altor procese, în timp ce SQL Server nu permite altor procese sau utilizatorilor accesul direct la fișierele gestionate de server; astfel de operații sunt posibile exclusiv prin intermediul unor funcții dedicate, furnizate de SQL Server. Drept consecință, SQL Server este considerat o opțiune mai bună ca MySQL în situațiile în care accesul la date reprezintă un aspect critic, iar cerințele de securitate sunt vizate în mod particular. În plus, SQL Server dispune de mecanisme de securitate suplimentare (precum cele de monitorizare și audit sau pentru definirea de politici care

vizează anumite riscuri de securitate și generarea alertelor aferente), care permit detectarea unor breșe de securitate sau a unor setări necorespunzătoare și corectarea acestora.

- Maniera de execuție a interogărilor – În MySQL execuția unei interogări nu poate fi oprită în mod individual, întregul proces trebuie stopat; pe de altă parte, SQL Server permite oprirea execuției interogărilor. Deși această diferență nu influențează activitatea de administrare a bazelor de date, ea poate reprezenta un argument în favoarea SQL Server pentru echipa de implementare a aplicațiilor ce accesează și prelucrează aceste date. În același registru care vizează interogările, pot fi menționate diferențele în privința manierei în care se realizează selecția înregistrărilor din tabele. Dacă sunt implicate surse multiple, în MySQL se recurge la interogări distincte pentru fiecare bază de date. Prin comparație, această operație este considerabil simplificată în SQL Server, care permite utilizarea de baze de date multiple pentru aceeași interogare.
- Realizarea copiilor de siguranță și restaurarea bazelor de date – Pe durata generării fișierelor de backup, bazele de date MySQL sunt disponibile exclusiv în acest scop, devenind inutilizabile altor procese. Acest comportament este cu precădere problematic în cazul bazelor de date de mari dimensiuni, astfel încât trebuie modificat prin intervenții explicite (utilitarul mysqldump); prin comparație, SQL Server permite accesul la date și realizarea de prelucrări în paralel cu derularea procesului de backup.
- Licențe și documentație – MySQL este disponibil în sistem dual, ca produs software open-source sub licența GPLv2, sau în baza unei licențe care presupune costuri. SQL Server este un produs Microsoft, a cărui utilizare pentru exploatare curentă, ca server de producție, necesită deținerea licențelor corespunzătoare. Ambele SGBD beneficiază de o amplă documentație tehnică dezvoltată de companiile care le promovează, dar și de suport extins din partea comunității software, care include numeroase forumuri și resurse disponibile online. Dezvoltatorii și utilizatorii MySQL și SQL Server își aduc propria contribuție la popularizarea acestor sisteme pentru baze de date, prin furnizarea de conținut relevant pentru înțelegerea și utilizarea lor, ca și prin oferirea de suport și soluții concrete pentru diferite probleme de implementare.

- Ediții disponibile și costuri asociate – În cazul ambelor SGBD, edițiile diferă prin setul funcționalităților oferite, ceea ce determină și un anumit nivel al costurilor pentru beneficiari. MySQL este disponibil în edițiile Community, Standard, Cluster CG și Enterprise. Varianta Community este gratuită și open-source; restul edițiilor sunt disponibile contra cost, cel mai amplu set de funcționalități (precum opțiunile avansate pentru securizarea bazelor de date sau pentru backup) fiind furnizat de ediția Enterprise. MySQL oferă și serviciul MySQL Cloud, care dispune de diferite opțiuni tarifare, în funcție de dimensiunea și complexitatea bazelor de date. Și SQL Server este disponibil în mai multe ediții: Express (gratuită, cu un set limitat de funcționalități), Developer (gratuită, expune setul complet de funcționalități SQL Server, dar nu poate fi folosită pentru dezvoltarea de software comercial), Standard și Enterprise (ambele presupun costuri cu licențele, diferă în privința setului de funcții și opțiuni oferite). SQL Server este disponibil și ca serviciu de tip cloud (Azure SQL), costurile pentru beneficiari depinzând de dimensiunea bazei de date și regimul de utilizare. În condițiile în care MySQL este disponibil fără costuri asociate, ca software open-source, iar ediția MySQL Enterprise are un cost mai mic decât ediția echivalentă a SQL Server, se poate aprecia că MySQL reprezintă o opțiune mai avantajoasă din perspectiva costurilor. Pe de altă parte, acest aspect trebuie corelat cu cele de natură tehnică, analizându-se plusurile și minusurile fiecărui SGBD, prin raportarea la cerințele concrete ale unui anumit beneficiar.

2.1.4 Considerații privind utilizarea MySQL și Microsoft SQL Server pentru implementarea platformei de colectare a datelor privind PPCA

În condițiile în care atât MySQL, cât și SQL Server, sunt utilizate pe scară largă pentru diferite tipuri de aplicații, opțiunea pentru un anumit SGBD trebuie să aibă în vedere, dincolo de elementele generale expuse în secțiunile anterioare, o serie de aspecte particulare, de factură IT și organizațională, care definesc un anumit context operațional și regimul de exploatare al bazelor de date ce vor fi gestionate cu ajutorul acestor sisteme. În cele ce urmează, vor fi enunțate principalele motivații care au stat la baza deciziei echipei de dezvoltare de a recurge la ambele SGBD pentru implementarea soluției informatice care vizează colectarea deșeurilor prin programul PPCA.

MySQL: SGBD aferent bazelor de date utilizate de sistemul operațional

Aceste baze de date sunt gestionate distinct de fiecare unitate administrativ-teritorială (UAT) responsabilă pentru colectarea deșeurilor, fiind abordate drept surse care alimentează cele două structuri dedicate pentru integrarea și analiza datelor, care sunt disponibile la nivelul sistemului (Data Lake, respectiv data warehouse). Prin comparație cu acestea din urmă, bazele de date utilizate de sistemul operațional fac obiectul unor cerințe mai relaxate în planul performanței și securității, prin prisma volumului și naturii datelor și tranzacțiilor pe care le gestionează și care vizează o singură UAT. Drept urmare, MySQL a fost considerat o opțiune validă pentru implementarea acestor baze de date, mai ales că, din perspectiva costurilor, constituie o soluție mai accesibilă decât SQL Server. Pe de altă parte, un nivel mai scăzut al costurilor nu reprezintă, în mod necesar, un minus în planul funcționalităților, MySQL fiind un sistem versatil, ce poate fi utilizat de aplicații dezvoltate în diferite limbaje.

De asemenea, infrastructura IT necesară aplicațiilor care oferă suport procesului de colectare a deșeurilor constituie un alt argument pentru utilizarea MySQL pentru implementarea bazelor de date operaționale. Cu alte cuvinte, când Linux este platforma vizată și soluția de implementare se raportează la stiva LAMP (Linux, Apache, MySQL, PHP), MySQL reprezintă cea mai adecvată opțiune.

Microsoft SQL Server: SGBD utilizat pentru Data Lake și Data Warehouse

Structura de tip Data Lake reprezintă un punct central al arhitecturii sistemului, având rolul de a colecta datele din bazele de date ale sistemelor operaționale de la nivelul UAT; ulterior, se realizează prelucrările care permit integrarea datelor din Data Lake la nivelul depozitului de date folosit pentru reportare și analiză. Acest regim de utilizare a soluției informatice, în ansamblul său, relevă diferența

semnificativă între volumul datelor gestionate distinct, de fiecare UAT, pe de o parte, și volumul datelor disponibile la nivelul Data Lake, respectiv al depozitului de date, care trebuie să stocheze datele transmise de toate entitățile de tip UAT. În mod evident, volumul datelor din astfel de structuri, unice la nivelul sistemului, este net superior celui gestionat de bazele de date operaționale care le alimentează periodic. În condițiile în care un volum important de date se corelează cu potențiale probleme de performanță, iar analizele comparative MySQL - SQL Server relevă, în mod uzual, avantajul soluției Microsoft în planul performanței, SQL Server a fost selectat ca SGBD destinat structurilor de tip Data Lake și data warehouse.

Din perspectiva interoperabilității, se impune observația că transferul datelor între sistemele operaționale și Data Lake nu ridică problema incompatibilității datelor, întrucât se bazează pe servicii web și pe JSON. Sistemele-sursă aferente UAT transmit datele MySQL sub forma unui șir de caractere care se conformează sintaxei JSON, acest șir fiind prelucrat ulterior la nivelul Data Lake pentru extragerea datelor elementare ce trebuie mapate câmpurilor-destinație. Aceste prelucrări sunt în mod considerabil facilitate de suportul dedicat pentru JSON disponibil în SQL Server, un alt atu al acestui SGBD. Absența problemelor de interoperabilitate este și mai evidentă în cazul transferului datelor între Data Lake și depozitul de date, în cadrul procesului ETL (Extraction - Transformation - Loading) pentru popularea depozitului, întrucât ambele structuri de date sunt implementate și gestionate în mediul SQL Server.

Dincolo de avantajele pe care le oferă în planul securității și performanței, enunțate în secțiunea dedicată prezentării generale a SQL Server, acest SGBD prezintă două attribute care au influențat în mod decisiv alegerea sa pentru implementarea Data Lake și data warehouse:

- SQL Server dispune de un set robust de instrumente și funcții pentru administrarea serverului și bazelor de date;
- SQL Server permite integrarea nativă cu mediul Microsoft .Net și cu alte aplicații și instrumente software furnizate de Microsoft, care pot fi utilizate la dezvoltarea unor soluții complexe de analiză a datelor. În acest mod, pot fi evitate problemele de integrare și interoperabilitate ce trebuie gestionate în mod explicit atunci când se recurge la tehnologii eterogene, care provin de la producători software diferiți.

Din ansamblul instrumentelor de administrare și monitorizare disponibile în SQL Server, absolut necesare pentru gestionarea unor aspecte critice pentru exploatarea oricărui server de date (precum securitatea), în contextul sistemului destinat analizei colectării deșeurilor, un rol aparte revine serviciului

SQL Server Agent, care a permis decuplarea prelucrărilor derulate la nivelul sistemelor operaționale ale UAT de cele realizate la nivel central, în cadrul Data Lake, respectiv validarea datelor transmise de fiecare UAT și memorarea lor în tabelele din Data Lake. Într-o astfel de abordare, transferul propriu-zis este limitat la importul șirului JSON care conține datele, fără a include salvarea acestora în tabelele-destinație. Memorarea datelor în tabele reclamă validarea lor prealabilă, iar această prelucrare poate varia considerabil ca durată, în funcție de volumul de date transmis, făcând nefezabilă returnarea imediată a rezultatului validării către sistemul care transmite datele. Drept urmare, transferul datelor în format JSON se finalizează cu memorarea acestora în formă brută (șir de caractere) într-un tabel dedicat din structura data-lake, urmând ca validarea să se realizeze ulterior, la intervale prestabilite. În planul implementării, aceste prelucrări asincrone sunt apanajul serviciului SQL Server Agent, care a permis executarea automată, la anumite intervale, a procedurii ce identifică toate șirurile JSON neprocesate (cele care au fost primite după ultima execuție a job-ului) pentru a le prelucra în mod corespunzător (validare, urmată de salvarea datelor în tabele, în cazul șirurilor valide, sau de jurnalizarea erorilor, în cazul șirurilor invalide).

După cum s-a menționat anterior, un alt argument în favoarea utilizării SQL Server pentru implementarea Data Lake și data warehouse îl reprezintă suportul disponibil în mod implicit pentru utilizarea tehnologiilor Microsoft ce permit valorificarea potențialului informațional al acestor structuri de date, prin intermediul unor aplicații sau instrumente software dedicate, destinate dezvoltării soluțiilor complexe pentru analiza datelor. Dincolo de beneficiul integrării native a componentelor unei astfel de soluții, aplicațiile Microsoft care vizează raportarea și analiza datelor de afaceri prezintă o serie de atuuri importante, confirmate de diferite studii care compară produsele dedicate domeniului Business Intelligence. În cazul sistemului de față, situația colectării deșeurilor în timp real poate fi analizată prin intermediul tablourilor de bord realizate și gestionate în mediul Microsoft Power BI; avantajele utilizării Power BI sunt expuse într-o secțiune distinctă din cadrul acestui document. De asemenea, au fost avute în vedere potențiale dezvoltări ulterioare ale sistemului, prin folosirea serviciilor SQL Server din sfera Business Intelligence pentru transferul și integrarea datelor din surse multiple (SQL Server Integration Services - SSIS), raportare (SQL Server Reporting Services - SSRS) sau analiza datelor în timp real (SQL Server Analysis Services - SSAS).

2.2 Soluții tehnice pentru schimbul electronic de date

SOA (Service oriented architecture/Arhitectura orientată pe servicii) reprezintă o arhitectură care se focalizează pe servicii discrete în locul unei abordări monolitice. Un serviciu SOA este o unitate de

software autonomă concepută pentru a îndeplini o anumită sarcină. Aplicațiile folosesc SOA și standarde de interfață simplă pentru a accesa servicii pentru a forma aplicații noi. SOA simplifică sistemele software complexe în servicii reutilizabile care pot fi accesate de alte aplicații și utilizatori denumiți consumatori de servicii. Aceste servicii pot fi utilizate ca elemente de bază ale noilor aplicații. Fiecare serviciu SOA are o sarcină specifică și o interfață care include parametrii de intrare și ieșire ai serviciului, precum și protocolul de comunicație necesar pentru a-l accesa.

Arhitectura bazată pe servicii constă de obicei în patru componente:

1. Serviciul. Acesta este fundamentul SOA. Serviciile pot fi private și disponibile numai pentru utilizatorii autorizați sau open source și disponibile public. Fiecare serviciu conține o implementare a serviciului, care este codul responsabil pentru efectuarea serviciului; un contract de servicii, care descrie parametrii unui serviciu și costul și calitatea acestuia; și o interfață de serviciu, care este nivelul unui serviciu care definește modul de comunicare cu acesta și gestionează comunicarea cu alte servicii și sisteme.
2. Furnizorul de servicii: Componenta care creează sau furnizează serviciul.
3. Consumatorul de servicii: Componenta care generează cereri către furnizorul de servicii. Contractul de servicii descrie regulile de interacțiune dintre un furnizor de servicii și solicitantul de servicii.
4. Registrul de servicii: Cunoscut și ca depozit de servicii, un registru de servicii este un director de servicii disponibile. Acesta are sarcina de a stoca descrierile serviciilor și alte informații relevante despre modul de utilizare a serviciilor unui furnizor de servicii.

Obiectivele principale ale SOA

- Serviciul. Acest obiectiv are ca scop structurarea procedurilor sau componentelor software ca servicii. Serviciile sunt concepute pentru a fi ușor cuplate cu aplicații, așa că sunt utilizate doar atunci când este necesar. De asemenea, sunt proiectate astfel încât dezvoltatorii de software să le poată utiliza cu ușurință pentru a crea aplicații într-un mod consistent.
- Publicarea. SOA își propune să ofere un mecanism de publicare a serviciilor disponibile care să includă funcționalitatea și cerințele de intrare/ieșire ale acestora. Serviciile sunt publicate într-un mod care permite integrarea facilă a acestora în aplicații.
- Securitatea. Acest obiectiv se bazează pe securitatea componentelor individuale din arhitectură, procedurile de identitate și autentificare legate de acele componente și securitatea conexiunilor dintre componentele arhitecturii.

Avantajele SOA

- **Fiabilitate:** testarea și depanarea sunt mai ușor de realizat pentru că se realizează pe servicii mici și independente în loc de bucăți masive de cod.
- **Independent de locație:** serviciile sunt localizate prin intermediul registrului de servicii care poate fi accesat prin URL, prin urmare se poate modifica locația în timp fără întreruperea funcționării aplicației
- **Scalabilitatea:** rularea serviciilor pe diverse platforme, limbaje de programare, servere favorizează scalabilitatea
- **Independența de platformă:** Arhitectura Orientată pe Servicii permite dezvoltarea de aplicații complexe prin integrarea diferitelor servicii din surse diferite care o fac independentă de platformă.
- **Cuplarea slabă:** conceptul de cuplare slabă din SOA este inspirat de paradigma proiectării orientate obiect, care reduce cuplarea dintre clase pentru a favoriza un mediu în care clasele pot fi schimbate fără a rupe relația existentă. SOA încurajează foarte mult dezvoltarea serviciilor independente pentru a spori eficiența aplicației software.
- **Reutilizarea:** serviciile definite în cadrul unei arhitecturi SOA pot fi folosite în mod repetat în diferite aplicații și procese, ceea ce duce la economii semnificative de timp și resurse. Aceasta reduce dezvoltarea și întreținerea duplicată a codului și a funcționalității, creând o mai mare eficiență în procesele de dezvoltare software.
- **Agilitatea:** una dintre principalele beneficii ale arhitecturii SOA constă în flexibilitatea și adaptabilitatea sa. Prin dezvoltarea unei aplicații sau a unui sistem sub forma unor servicii independente, organizațiile pot răspunde mai ușor la schimbările cerințelor de afaceri. Această flexibilitate este esențială într-un mediu în care cerințele și condițiile de piață pot evolua rapid.
- **Mentenanța facilă:** întrucât arhitectura orientată spre servicii este o unitate sau entitate independentă, întreținerea sau actualizările aplicației au devenit mult mai ușoare fără afectarea altor servicii.

JSON (JavaScript Object Notation) este un format de reprezentare și interschimb de date între aplicații informatice. Este un format text, inteligibil pentru oameni, utilizat pentru reprezentarea obiectelor și a altor structuri de date și este folosit în special pentru a transmite date structurate prin rețea, procesul purtând numele de serializare.

În comparație cu XML, JSON este mai concis și mai compact, făcându-l mai rapid de analizat și generat. Acest lucru reduce dimensiunea și lățimea de bandă a transferurilor de date, precum și îmbunătățirea performanței și eficienței procesării datelor. În plus, JSON este mai ușor de citit și de scris decât XML, ceea ce îmbunătățește lizibilitatea și mentenabilitatea codului. De asemenea, acceptă tipuri de date native, cum ar fi numere, booleeni și valori nule, simplificând reprezentarea și validarea datelor. Principalele caracteristici și avantaje ale JSON sunt prezentate în continuare.

- JSON (JavaScript Object Notation) este un format de date text ușor de citit și de generat, utilizat pe scară largă pentru schimbul electronic de date între aplicații și sisteme diferite. JSON este adesea formatul preferat pentru scenariile de procesare a datelor datorită numeroaselor sale avantaje:
- Ușurință în citire și scriere: JSON utilizează un format de text simplu, ușor de înțeles atât pentru oameni, cât și pentru mașini. Acest lucru face ca datele să fie ușor de citit și de generat de către dezvoltatori.
- Interoperabilitate: JSON este un format independent de limbaj, ceea ce înseamnă că poate fi utilizat într-o varietate de limbaje de programare, inclusiv JavaScript, Python, Java, C#, și multe altele. Acest lucru facilitează schimbul de date între diferite sisteme și platforme.
- Eficiență în dimensiunea fișierelor: JSON are o structură de date simplă și compactă, ceea ce face ca fișierele JSON să fie relativ mici în comparație cu alte formate de date, cum ar fi XML. Acest lucru poate reduce timpul de transmitere a datelor și necesită mai puțin spațiu de stocare.
- Suport nativ în majoritatea limbajelor de programare: Multe limbaje de programare oferă suport nativ pentru serializarea și deserializarea datelor JSON, ceea ce simplifică manipularea datelor JSON în aplicații.

- Ușor de folosit în aplicații web: JSON este adesea utilizat în aplicații web pentru a transmite date între clientul web și serverul web. Deoarece JavaScript (un limbaj comun în dezvoltarea web) poate manipula ușor obiecte JSON, acesta se potrivește bine cu cerințele dezvoltării web moderne.
- Suport pentru structuri de date complexe: JSON permite reprezentarea datelor structurate în mod clar și eficient. Poate gestiona obiecte, liste, tipuri de date primitive și valori nule, permițând astfel reprezentarea unei varietăți de informații.
- Ușor de citit și de depanat: Structura JSON este ușor de citit și de înțeles, ceea ce face de depanarea mai eficientă atunci când apar probleme în schimbul de date sau în procesarea acestora.

2.3 Arhitecturi și tehnologii pentru gestiunea și analiza datelor

În efortul de adaptare dinamică la evoluția galopantă a tehnologiilor, integrarea de noi arhitecturi de date, care să poată susține cereri de informații din ce în ce mai frecvente și complexe, reprezintă o constantă peremptorie a demersurilor de modernizare a organizațiilor și societății în general.

Pentru o documentare riguroasă în vederea proiectării și dezvoltării unui sistem de monitorizare a fluxurilor de deșeurii menajere și similare în scopul îmbunătățirii mecanismelor de gestionare a instrumentului economic „Plătește Pentru Cât Arunci”, am realizat căutări de informații în bazele de date științifice recunoscute, interesându-ne totodată îndeaproape de bunele practici adoptate atât de companiile producătoare de software, cât și de cele care au implementat sisteme ce oferă facilități similare celor vizate de echipa de proiect.

Rezultate acestei perioade de cercetare științifică sunt prezentate în secțiunile care urmează.

2.3.1 Conceptul de Big Data

Big Data este un termen folosit pe scară largă pentru a face referire la seturi voluminoase de date structurate, nestructurate și semi-structurate, care sunt dificil de gestionat folosind instrumente tradiționale de procesare a datelor. Aceste date nu sunt generate doar prin modalitățile tradiționale, ci și prin intermediul senzorilor de diverse tipuri, încorporați într-o varietate de medii.

Deși nu se cunoaște cu exactitate cine a utilizat pentru prima dată termenul de „Big Data”, majoritatea studiilor îl creditează pe informaticianul John Mashey de la Silicon Graphics, acest termen

fiind folosit încă de la începutul anilor '90. Conceptul de „Big Data”, apare în literatura de specialitate în 1997, într-un articol publicat de cercetătorii NASA, Michael Cox și David Ellsworth (Cox&Ellsworth,1997). O definiție a conceptului a fost introdusă în 2005 de Roger Magoulas (Halevi&Moed, 2012), acesta asociindu-l cu o cantitate mare de date pe care tehnicile tradiționale de gestionare a datelor nu o pot gestiona și prelucra din cauza complexității și dimensiunii acesteia. Big Data a fost prezent și în cercetare, începând cu anii 1970, dar a devenit subiect de interes în publicațiile științifice abia din anul 2008. Big Data pune un accent deosebit pe problema volumului de informații care depășește capacitatea tehnologiilor tradiționale de gestionare a datelor și creează nevoia de noi instrumente și tehnologii pentru a gestiona acest volum extrem de mare de date. În general, datele utilizate de tehnologiile Big Data provin din baze de date, sunt furnizate de senzori și alte dispozitive audio/video, rețele, fișiere jurnal, aplicații web tranzacționale și rețelele sociale, fiind generate în cea mai mare parte în timp real.

În 2001, Doug Laney, analist la Gartner, scrie un raport în care analizează trei dintre caracteristicile Big Data, care au dat naștere așa-numitului model 3V: volumul, viteza și varietatea. Acest document publicat de compania META Group, care a achiziționase firma de consultanță Gartner, constituie punctul de plecare pentru modelele teoretice care au apărut ani mai târziu și care încorporează caracteristici suplimentare legate de conceptul de Big Data. Celor trei V-uri cantitative (volum, varietate și viteză), li s-au adăugat ulterior două V-uri calitative (veracitate și valoare), constituind o extensie a modelului inițial sub numele de modelul 5V. Modelul evoluează continuu, adăugându-se noi caracteristici, abordările actuale incluzând peste 10 caracteristici.

Principalele caracteristici fac referire la următoarele aspecte:

- Volumul se referă la dimensiunea seturilor de date care invadează organizațiile. An de an, volumele de date gestionate de acestea continuă să crească și este de așteptat ca această tendință să persiste. Potrivit Fortune Business Insights, mărimea pieței globale de analiză a Big Data este estimată să atingă 549,73 miliarde USD în 2028. Astăzi, datele sunt generate la nivel de petabytes și exabytes și provin din surse precum sistemele de procesare a tranzacțiilor, e-mailuri, rețelele sociale, baze de date cu clienți, dispozitive de monitorizare și aplicații mobile. Pentru a gestiona acest volum imens de date, managerii folosesc Data Lake, Data Warehouse sau sisteme de gestiune a datelor, stochează date în cloud sau utilizează furnizori de servicii precum Google

Cloud. Problema majoră o constituie volumul prea mare de date care face dificilă stocarea și care poate avea un impact major asupra complexității proceselor de analiză a datelor;

- Varietatea se referă la diversitatea tipurilor de date, respectiv la gama largă de formate de date disponibile: Parquet, JavaScript Object Notation (JSON), Avro, etc. O organizație poate obține date din diverse surse de date, interne sau externe, care pot varia ca valoare. Provocarea în cazul varietății constă în standardizarea și distribuirea tuturor datelor colectate. Datele colectate pot fi nestructurate, semistructurate sau structurate. Datele nestructurate sunt date neorganizate, disponibile în fișiere sau formate diferite. De obicei, datele nestructurate nu sunt compatibile cu o bază de date relațională principală, deoarece nu se potrivesc cu modelele de date convenționale. Datele semistructurate sunt combinate din mai multe surse eterogene, nu au o structură implicită asociată datelor sau au o structură parțial specificată (Reveiu, 2003). Acest lucru face ca datele semistructurate să fie mai ușor de procesat decât datele nestructurate. Datele structurate sunt date ce pot fi organizate într-un depozit permițând procesarea și analiza eficientă a datelor. În prezent, există o tipologie diversificată de informații pentru analiză furnizate de noile canale și tehnologii apărute ce generează date semistructurate sau nestructurate. Varietatea vizează surse de date precum bazele de date, documente, XML, e-mailuri, blog-uri, mesaje instant, click stream-uri, fișiere log, imagini statice, audio, video, date despre cursul acțiunilor, tranzacții financiare etc. (Alexandru și Coardoș, 2017).
- Viteza se referă la fluxul constant de date dintr-o varietate de surse și, frecvent, dispozitive IoT. Evident, este vizată atât rapiditatea cu care sunt generate datele, cât și rapiditatea cu care datele trebuie să fie prelucrate pentru a satisface cererea. În acest caz sunt implicate fluxuri de date, procese de creare de înregistrări structurate, dar este necesară și disponibilitatea pentru acces și livrare. Viteza de generare, prelucrare și analiză a datelor crește continuu datorită caracteristicii de realizare în timp real a proceselor de generare, cererilor care rezultă din combinarea fluxurilor de date cu procesele de afaceri, specificului proceselor de luare a deciziilor. (Alexandru și Coardoș, 2017).
- Veracitatea este una dintre caracteristicile Big Data ce se manifestă în sens opus comportamentului celorlalte caracteristici. Pe măsură ce volumul, viteza și varietatea cresc, veracitatea (încrederea în date) scade. Pentru a discuta despre această caracteristică a modelului 5 V al Big Data, trebuie avute în vedere două aspecte: fiabilitatea statistică, adică consistența sau gradul de certitudine al datelor precum și fiabilitatea datelor care depinde de o serie de factori

precum originea datelor, metodele de colectare și prelucrare, infrastructura și facilitățile informatice utilizate. Prin urmare, veracitatea Big Data garantează că datele sunt fiabile, autentice și protejate împotriva accesului și modificărilor neautorizate. Se asigură astfel, o interpretare corectă a informațiilor, pe măsură ce varietatea și numărul surselor de date existente crește.

- Valoarea este cea mai importantă caracteristică a Big Data. Este o modalitate de a măsura utilitatea unui set de date pentru o organizație și de a arăta cât de bine se potrivesc datele cu obiectivele acelei organizații, dacă pot ajuta ca situația acesteia să se îmbunătățească financiar sau strategic, fiind o măsură a capacității sale de a genera venituri și de a rezolva problemele de afaceri printr-o percepție corectă a faptelor. Valoarea poate fi măsurată prin intermediul următoarelor variabile: timp, legalitate, context, calitate, cost de achiziție.

Alte caracteristici Big Data relevate de literatura de specialitate sunt:

- Vizualizarea - are în vedere instrumentele prin care poate fi realizată reprezentarea vizuală a datelor, în încercarea de a reprezenta datele și a pune lucrurile în perspectivă. Instrumentele actuale de vizualizare a datelor se confruntă cu provocări tehnice, rezultat al limitărilor tehnologiei în memorie, a scalabilității reduse, a funcționalității și timpului de răspuns redus. Nu pot fi utilizate graficele tradiționale atunci când se încearcă parcurgerea unui volum mare de date, fiind nevoie de modalități diferite de reprezentare a datelor, cum ar fi: gruparea datelor (data clustering), tree maps, coordonate paralele, diagrama de rețea circulară etc.
- Valabilitatea- se referă la exactitatea și corectitudinea informațiilor stocate.
- Volatilitatea se referă la rata modificărilor și durata de viață a datelor. Organizațiile trebuie să înțeleagă cât timp este valabil un anumit tip de date. De exemplu, datele referitoare la sentimente, stări de spirit se modifică frecvent pe rețelele sociale și sunt foarte volatile. Un exemplu de date cu volatilitate scăzută sunt tendințele meteo, care sunt mai ușor de prezis.
- Vulnerabilitatea are în vedere faptul că un volum uriaș de date ridică și multe probleme de securitate, existând un trend ascendent în ceea ce privește atacurile cibernetice și încălcările de securitate IT.
- Variabilitatea vizează faptul că unele fluxuri de date înregistrează perioade de vârf și sezonabilitate, periodicitate. Gestionarea unei cantități mari de informații nestructurate este dificilă și necesită tehnici puternice de procesare.

În condițiile în care discutăm despre varietate, volum și viteză tot mai mari, sistemele tradiționale utilizate în depozitarea datelor nu mai pot susține nevoile de date ale organizației moderne. În prezent au loc transformări importante în managementul datelor în ceea ce privește modul în care datele sunt stocate, procesate, gestionate și furnizate factorilor de decizie. Fiecare organizație trebuie să colecteze un volum mare de date pentru a susține procesul decizional și pentru a extrage corelații prin analiza datelor ca bază pentru decizii.

2.3.2 Baze de date, depozite de date, Data Lake: posibile alegeri pentru elaborarea arhitecturii sistemului dezvoltat

Studiile privind bunele practici în domeniu semnaleză utilizarea frecventă a depozitelor de date – Data Warehouse (DW), Data Lake (DL) și Data Mart (DM), ca soluții de stocare în cloud (AWS, 2023). Potrivit definițiilor Amazon, un Data Warehouse stochează datele preprocesate pentru analiză și business intelligence într-un format structurat. Un Data Mart reprezintă un depozit de date creat special pentru a deservi nevoile unei anumite unități structurale organizaționale, cum ar fi departamentul de finanțe, marketing sau resurse umane. Data Lake-ul este prezentat drept soluție arhitecturală recomandată atunci când este nevoie de un depozit central pentru date brute, ori nestructurate, ce pot fi mai întâi stocate și ulterior procesate (AWS, 2023).

Pentru gestionarea unor volume mari de date, provenite din diverse surse, organizațiile folosesc atât Data Lake-uri, cât și depozitele de date (Oracle, 2023a). În funcție de necesitățile informaționale, se selectează una sau cealaltă dintre variante, ori o combinație a acestora. Data Lake-urile stochează o mulțime de date disparate, nefiltrate, pentru a fi utilizate ulterior într-un anumit scop. Datele din sistemele operaționale, aplicațiile mobile, rețelele sociale, dispozitivele IoT și multe altele sunt capturate ca date brute într-un Data Lake, utilizatorii fiind cei care stabilesc structura și formatul diferitelor seturi de date pe care sunt interesați să dezvolte analize. Un Data Lake poate reprezenta o soluție potrivită pentru situații ce necesită stocare la costuri reduse de date neformatate, nestructurate, din mai multe surse, rezervate unor analize ce se vor derula în scopuri prestabilite, în viitor (Oracle, 2023a).

Depozitele de date sunt alegeri arhitecturale de date recomandate în situația nevoii de analize avansate ale datelor cu caracter istoric, generate la nivel intern; acestea sunt colectate, apoi contextualizate și transformate, cu scopul de a crea perspective bazate pe manipularea unor cantități mari de date din diverse surse.

Potrivit studiilor Oracle, dacă, inițial, depozitele de date erau construite, în cadrul sistemelor de asistare a deciziilor, pentru analiză și raportare pe baza fluxurilor de date operaționale, astăzi volumele mari de date au impus integrarea lor cu tehnologii de inteligență artificială și algoritmi de învățare automată. Devenind astfel mai eficiente, operațiile de colectare, curățare și integrare datelor au fost ameliorate constant, iar platformele actuale de Business Intelligence se circumscriu unor infrastructuri de analiză largi, care găzduiesc o mare varietate de aplicații, ce vizează, spre exemplu, analiza operațională și managementul performanței.

Plusvaloarea incrementală a organizațiilor se datorează următoarelor facilități asigurate de o arhitectură de tip depozit de date (Oracle, 2023a):

- raportare tranzacțională: furnizează "instantanee" ale performanței afacerii;
- instrumentar dedicat aprofundării și consolidării analizelor: tehnici de segmentare și extragere de informații de tip "Slice and dice", interogări ad-hoc, instrumente BI de vizualizare;
- valorificarea datelor în scopul elaborării de predicții de performanță conținând vizualizări și informații specifice prospective;
- efectuarea de analiză tactică (spațială, statistică) și simulări, pe bază de scenarii „ce ar fi dacă” (What if);
- stocare de date timp de multe luni sau ani.

2.3.3 Arhitecturi bazate pe Data Lake

Solicitările lumii manageriale privind fundamentarea deciziilor pe analize complexe ale datelor organizaționale au condus la dezvoltări importante ale științei datelor, ale algoritmilor de învățare automată și analiză avansată, ale instrumentelor de generare a tablourilor de bord în timp real. În prezent, companii precum Google, Amazon și Facebook își revendică afaceri tradiționale prin valorificarea datelor, iar organizațiile furnizoare de servicii financiare și asigurări au fost întotdeauna bazate pe analiza datelor. Mai nou, Internetul obiectelor (IoT) schimbă procesele de producție, transport, agricultură și serviciile medicale, iar inteligența artificială și învățarea automată pătrund în toate aspectele vieții sociale. Pentru a sprijini aceste eforturi și pentru a aborda aceste provocări, are loc o revoluție în gestionarea datelor în ceea ce privește modul în care datele sunt stocate, procesate, gestionate și furnizate factorilor de decizie. Big Data, concept definit prima dată de Doug Laney de la Gartner, este asociat celor trei V-uri deja celebre: volum, varietate și viteză, completate ulterior de un al patrulea, veridicitatea. Tehnologia permite

scalabilitate și eficiență a costurilor, de ordine de mărime mai mari decât cea ce este posibil cu infrastructura tradițională de gestionare a datelor.

Un depozit de date (DW, Data Warehouse) tipic include următoarele elemente (Oracle, 2023a):

- bază de date relațională pentru stocarea și gestionarea datelor;
- o soluție de extracție, încărcare și transformare (ELT, Extract, Transform, Load) menită să pregătească datele pentru analiză;
- facilități de analiză statistică, de raportare și capabilități de extragere a datelor;
- instrumente de analiză propriu-zisă și de vizualizare și prezentare a rezultatelor analizelor;
- alte aplicații analitice avansate ce aplică știința datelor, a algoritmilor de inteligență artificială (AI) și diverse caracteristici grafice și spațiale.

Organizațiile pot opta pentru soluții care combină procesarea tranzacțiilor, analiza în timp real rulate prin interogarea Data Warehouse și Data Lake și învățarea automată într-un singur serviciu de analiză a datelor, fără complexitatea, latența, costul și riscul extragerii, transformării și încărcării (ETL) de duplicate (Oracle, 2023a).

Pentru analizarea unor volume de date aflate în creștere, Amazon avansează următoarea configurare de arhitectură de sistem (Figura 7):

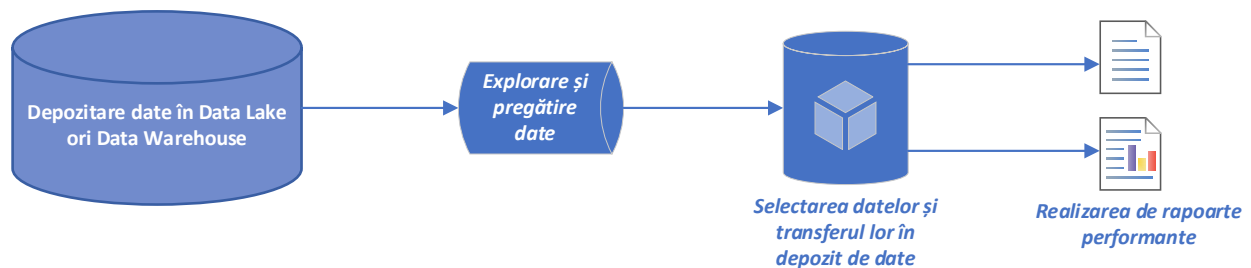


Figura 7. Soluție arhitecturală pentru un sistem eficient de analiză și raportare (sursă: adaptare după AWS, 2023)

Facilitățile pe care specialiștii din domeniu le au în vedere sunt sintetizate în tabloul de mai jos (Tabel 3).

Tabel 3. Caracteristici de lucru ale Data Warehouse și Data Lake (sursă: adaptat după AWS, 2023, MongoDB, 2023 și Coursera, 2023)

PARAMETRI	Data Lake	Data Warehouse
-----------	-----------	----------------

Date	Toate datele, inclusiv structurate, semi-structurate și nestructurate	Date relaționale din sisteme tranzacționale, baze de date operaționale și aplicații de linie de afaceri; date stocate în funcție de metrice și atribute
Schemă	Elaborată în momentul efectuării analizei (schema-la-citire); după stocarea datelor, pentru a oferi agilitate și captare ușoară a datelor	Adeea proiectat înainte de implementarea depozitului de date, dar poate fi scris și în momentul analizei (schema-la-scriere sau schema-la-citire); înainte de stocarea datelor, pentru a oferi securitate și performanță ridicată
Proces	Extract Load Transform (ELT)	Extract Transform Load (ETL)
Preț/Performanță	Rezultatele interogărilor devin mai rapide utilizând stocarea la costuri reduse și decuplarea procesului și stocării	Cele mai rapide rezultate de interogare folosind stocarea locală
Calitatea datelor	Orice date care pot sau nu fi curățate (adică date brute)	Date foarte îngrijite care servesc drept versiunea centrală a realității
Prospețimea datelor	Este posibil să nu fie actualizate (pe baza frecvenței proceselor ELT)	Este posibil să nu fie actualizate (pe baza frecvenței proceselor ETL; caracter "istoric")
Utilizatori	Analiști de afaceri (folosind date selectate), oameni de știință ai datelor, dezvoltatori de date, ingineri de date și arhitecți de date, care au nevoie de analize și instrumente aprofundate (cum ar fi modelarea predictivă) pentru a le înțelege	Analiști de afaceri, oameni de știință de date și dezvoltatori de date; Profesioniști care au nevoie de DW pentru operațiuni
Accesibilitate	Accesibil și ușor de actualizat	Complicată efectuarea de modificări
Analytics	Învățare automată, analiză exploratorie, descoperire de date, analiză de date în flux continuu (streaming), analiză operațională, date mari și profilare	Raportare în loturi, BI și vizualizări
Istoric	Relativ nou pentru piața de tehnologii big data	Conceptul există de mai multe decade

Data Lake - este o nouă abordare îndrăzneță care valorifică puterea tehnologiei Big Data și o îmbină cu agilitatea conferită de capacitatea de autoservire. În literatura se consemnează că majoritatea întreprinderilor mari de astăzi fie au implementat, fie sunt în proces de implementare a Data Lake. Ideea de autoservire - Self-Service - preia abordările anterioare, creatorul conceptului de Data Lake, James Dixon, CTO al Pentaho, aducând în prim-plan principalele două caracteristici ale acestora: (1) datele sunt colectate și stocate în forma și formatul inițial (date naturale sau brute); (2) datele sunt valorificate de către o comunitate mare de diverși utilizatori (Gorelik, 2023).

Diversitatea surselor și a formatelor de informații care pot fi identificate în peisajul digital de astăzi este o componentă a unei noi normalități în ceea ce privește modul în care datele trebuie să fie colectate, transferate, stocate, procesate și analizate.

Facilitățile sistemelor de gestiune a bazelor de date – Databases -, împreună cu cele specifice depozitele de date – Data Warehouses - ale unei organizații pot fi integrate în platforme de tip Data Lake. Cu arhitecturi Data Lake, datele pot fi colectate în formate „brute” și exploatate apoi prin programe robuste de analiză avansată (Hagstroem, 2017).

Construcții arhitecturale valoroase recente, Data Lake-urile au apărut ca platforme scalabile, bazate pe cloud, cu capabilități atât de stocare și procesare rentabile ale volumelor uriașe de date diverse, cât și de analiză și asistare a deciziei (Qlik, 2021). Premise favorabile acestor evoluții au fost considerate creșterea puterii de procesare a computerelor, îmbunătățirea tehnologiilor de stocare în cloud și a conectivității la diferite rețele. Informații detaliate despre profilurile personale ale clienților, date privind tranzacțiile de vânzare, caracteristicile mărfurilor, hărțile proceselor, dintr-o gamă largă de surse (dispozitive IoT - Internet-of-Things, site-uri de social media, sisteme online de vânzări, de colaborare internă și învățare organizațională) au amplificat temerea liderilor din mediul economic de a deveni copleșiți de volumul și varietatea de date pe care ar trebui să le gestioneze, de viteza cu care informațiile circulă prin rețele interne și externe și de costul administrării informațiilor critice de afaceri ori sociale. Politicile manageriale s-au mulat pe obiective de extindere a infrastructurii de gestionare rapidă a datelor masive, structurate și nestructurate, apărând astfel *Data Lake* (Hagstroem, 2017). Principalul avantaj este reprezentat de capacitatea de păstrare a datelor în formatele lor native, acestea suferind prelucrări și reconfigurări numai când și cum este necesar. În literatură sunt semnalate, de asemenea, economii de costuri, seturile de date nenesitând indexări și pregătiri prealabile colectării lor, iar *echipamentul hardware necesar fiind accesibil și ușor de procurat*. O abordare agilă a dezvoltării Data Lake poate ajuta

organizațiile să lanseze rapid programe de analiză și să alimenteze o cultură prietenoasă cu datele pe termen lung (Hagstroem, 2017).

Specialiștii de la Qlik¹ aduc în atenție principalii factori care influențează rentabilitatea investiției într-o arhitectura Data Lake. Primul este viteza de restituire a informațiilor necesare susținerii dezvoltării activității în condițiile intensificării concurenței, apariției multiplelor inovații tehnologice, schimbărilor pieței și a complexității sporite a cerințelor de afaceri. Responsabilii de procese trebuie să poată descoperi informațiile utile *în timp real* și să poată acționa pe baza acestora, fără a fi asumate riscuri de afectare a acurateței ori calității acestora. Apoi, este menționată capacitatea unei arhitecturi Data Lake de a procesa rapid date din surse și în formate foarte diverse – fluxuri continue (streaming), date tranzacționale, informații furnizate de senzori, imagini, text, fișiere video și audio - și de a valorifica imediat rezultatele. În al treilea rând, se consideră că volumele mari de date apărute în flux continuu necesită prelucrări bazate pe algoritmi de inteligență artificială, esențiali pentru furnizarea unor informații de calitate în timp real. Urmează scalabilitatea, flexibilitatea și agilitatea infrastructurilor de date și a celor de analiză, obligate să se conformeze cerințelor în continuă schimbare. Considerentele de ordin economic precum costuri mai mici, productivitate mai mare și timp de lansare pe piață mai scurt completează lista factorilor de influențare a rentabilității unei astfel de investiții.

2.3.4 Etape în dezvoltarea Data Lake

În viziunea Mckinsey, organizațiile care dezvoltă Data Lakes parcurg următoarele patru etape (Hagstroem, 2017):

1. *Faza de inițiere (landing) și a datelor brute.* Aici, lacul de date este un mediu de „captură pură” cu costuri reduse, scalabil, construit separat și neinvaziv față de sistemele IT existente, care permite stocarea datelor brute pe termen nelimitat înainte de a fi pregătite pentru valorificare. Sunt relevate drept elemente importante de guvernare a datelor etichetarea și clasificarea riguroasă a acestora.

2. *Nivelul de studiu al datelor (Data Science environment).* Lacul de date poate deveni acum operațional ca platformă pentru experimentare, prin extinderea de la funcționalități de colectare de date la cele de analize de date, de construire de prototipuri pentru programe de analiză. Cu ajutorul unor

¹ Conform informațiilor de pe site-ul oficial, compania Qlik oferă soluții de analiză și integrare a datelor în timp real, bazate pe Qlik Cloud, menite să reducă decalajele dintre date, informații și acțiuni și deservește peste 38.000 de clienți activi în peste 100 de țări.

instrumente open-source și comerciale implementate alături de lacul de date pot fi generate bazele de testare necesare.

3. *Descărcarea depozitelor de date.* În această etapă, prin procese ETL (Extract, Transform, Load – Extragere, Prelucrare, Încărcare), Data Lake încep să fie alimentate din depozitele de date (DW) existente. Printre avantajele, este menționat, printre altele, faptul că, cu costuri scăzute de stocare tipice unui Data Lake, pe baza datelor „reci” (folosite rar, latente sau inactive) pot fi generate informații utile, fără a afecta limitările de stocare sau dimensiunea depozitelor de date tradiționale, ori ritmul intens de transfer al datelor relaționale în DW-uri. De asemenea, operațiunile de căutare de tip „ac în carul cu fân”, dificil de realizat în baze de date prin interogări care nu sunt acceptate în structurile tradiționale de indexare, pot fi migrate către Data Lake.

4. *Componentă critică a operațiunilor de date.* Este faza în care Data Lake devine componentă a infrastructurii de date, asigurând opțiuni de lucru de tipul „Data-As-A-Service”. Pot fi efectuate analize avansate pentru a implementa programe de învățare automată, se pot construi aplicații precum generatoare de tablouri de bord pentru managementul performanței, ori de combinare a informațiilor colectate în lacul de date cu cele extrase din alte aplicații.

Din perspectiva Qlik (2021), o arhitectură modernă Data Lake înglobează platforme cloud (Amazon, MS Azure, GCP), instrumente de stocare și de calcul decuplate, magazie de obiecte bazată pe cloud + magazie de fișiere (Amazon S3, ADLS, GCS), procesare în timp real bazată pe Spark și tehnologie IoT, fără server și componente scalabile.

2.3.5 Elementele unei abordări agile în construirea unui Data Lake

Pentru asigurarea imperativelor de flexibilitate a platformei Data Lake, literatura consemnează ca necesară colaborarea strânsă între echipele IT și cele ale entității beneficiare pentru elaborarea soluțiilor optime privind tehnologia și proiectarea. Sunt pe lista de lucru întrebările relevante legate de găzduirea în cloud sau on-premise a lacului de date (folosind servere private, publice sau hibride off-site), de modalitatea de populare cu date și de momentele în care aceasta este realizată, de procesele de integrare și de scalabilitatea aplicației.

Se recomandă, de asemenea, surprinderea dificultăților de implementare și problemele de performanță prin colectarea și încorporarea de feedback de la unitățile beneficiare. O echipă de dezvoltare poate fi considerată agilă dacă va fi capabilă să adapteze procesele și protocoalele de



guvernare a datelor dimensiunilor crescânde ale lacului de date, tehnologiilor noi de analiză și stocare, ori cerințelor de afaceri în transformare continuă (Hagstroem, 2017).

Totodată o guvernare agilă a lacului de date creează premisele trasării clare a responsabilităților privind gestionarea surselor de date și a drepturilor de acces la acestea. Devine, în acest context, esențială colectarea și stocarea meta-datelor, utile pentru structurarea și întreținerea unui catalog central de date pentru toate părțile interesate. La elaborarea protocoalelor de gestionare a datelor, bunele practici recomandă corelarea procedurilor de lucru cu reglementările în vigoare vizând asigurarea confidențialității informațiilor cu caracter personal de către deținătorii datelor și comunicarea drepturilor de acces tuturor părților interesate relevante (Hagstroem, 2017).

2.3.6 Considerente privind maturitatea unui Data Lake: Data Swamp, Data Pond și Data Puddles

Cercetările, inclusiv la nivel terminologic, s-au extins de la depozit de date – Data Warehouse, la Data Lake, dar și la ceea ce nu ar trebui să devină un Data Lake un Data Swamp. De asemenea, în literatură pot fi întâlnite și conceptele Data Pond și Data Puddle. Așa cum se poate observa în figura de mai jos (Figura 8), inclusiv noțiunea Data Ocean – este vehiculată în spațiul cercetărilor privind valorificarea valențelor informaționale ale volumelor mari de date. Noțiunile de mai sus au – toate – legătură cu etapele evolutive observabile, parcurse de către un Data Lake până la un nivel de maturitate considerat acceptabil și au menirea de a sprijini specialiștii în sesizarea diferențelor dintre aceste etape.

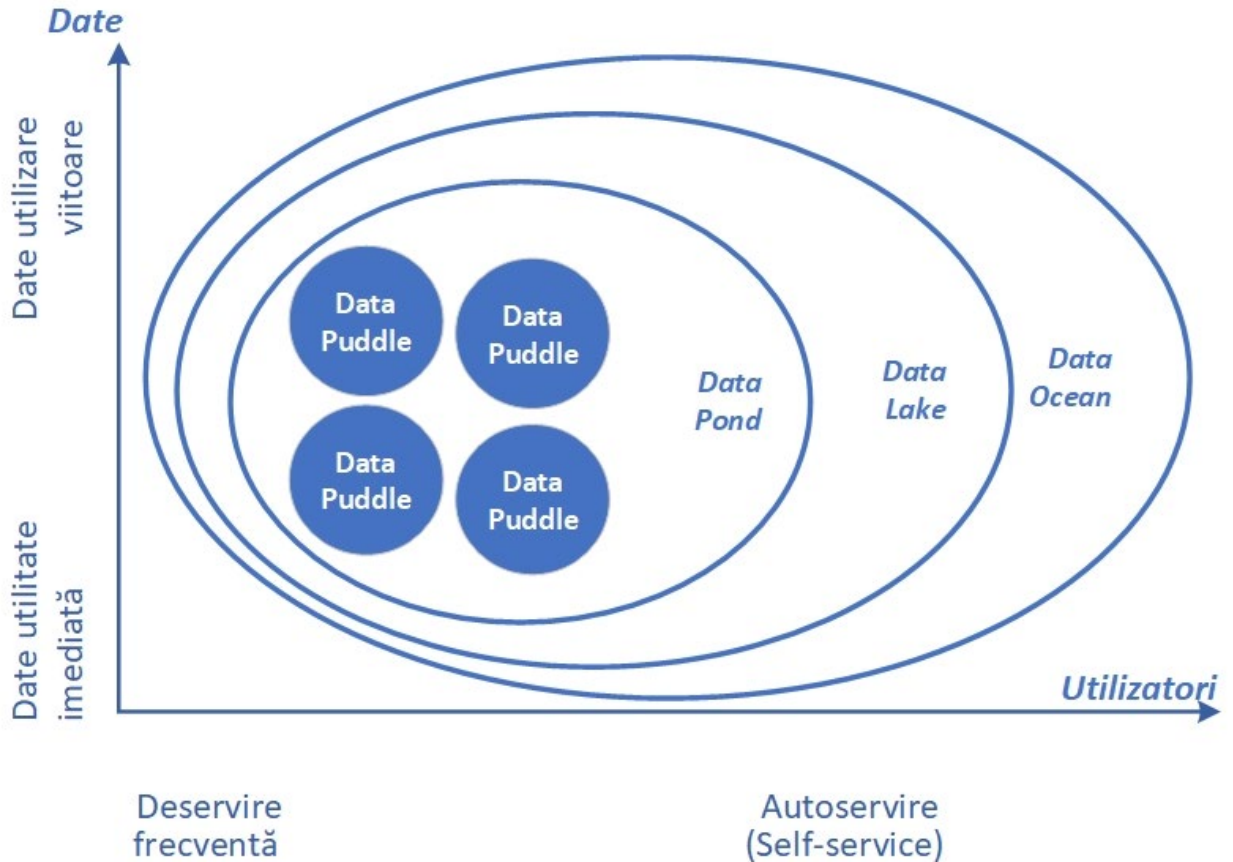


Figura 1 Concepte din literatură: clarificare grafică. Adaptare după Gorelik (2023).

Un Data Puddle este definit drept un Data Mart cu un singur scop sau cu un singur proiect, construit folosind tehnologia Big Data, fiind văzută ca un prim pas în adoptarea acestei tehnologii. Datele dintr-un Data Puddle sunt încărcate în scopul derulării unui singur proiect, al unei singure echipe. De obicei, motivul pentru care tehnologia Big Data este utilizată în locul depozitării tradiționale de date este acela de a reduce costurile și de a oferi performanțe mai bune în analiza și valorificarea datelor.

Un Data Pond este o colecție de Data Puddle. Dacă Data Puddle ar putea fi asimilate unor magazii de date (Data Mart) construite folosind tehnologia Big Data, atunci un Data Pond ar putea fi comparat cu un depozit de date construit folosind tehnologia Big Data, ori cu defecte de proiectare, ori ca o concatenare de Data Puddle-uri, ori reprezentând pur și simplu o descărcare a unui depozit de date existent. Poate apărea organic, pe măsură ce mai multe bălți sunt adăugate platformei de gestiune a volumelor mari de date (Gorelik, 2023). Motivația principală de a crea un Data Pond este fondată pe costurile mai mici ale tehnologiei și o scalabilitate mai bună; totuși, specialiștii avertizează asupra

examinării necesității unui nivel ridicat de participare IT în cazul dezvoltării și implementării unei astfel de soluții. În plus, un Data Pond limitează selectarea datelor doar la cele necesare derulării proiectului cărui i-a fost alocat. Generarea de costuri IT relativ ridicate și disponibilitatea limitată a datelor afectează atingerea obiectivelor de democratizare a utilizării datelor, ori de stimulare a autoservirii și de fundamentare a deciziilor bazate pe date de către responsabilii proceselor de afaceri.

Un Data Lake se diferențiază de un Data Pond prin două elemente esențiale. În primul rând, asigură capabilități de autoservire, grație cărora utilizatorii pot găsi și utiliza seturi de date de interes pentru ei fără a fi nevoiți să solicite pe ajutorul departamentului IT. În al doilea rând, poate găzdui date suplimentare, neutilizate în proiectele aflate în derulare, în ideea de a le conserva până la posibile solicitări viitoare. De la Data Lake se ajunge la Data Ocean, reperat atunci când datele cu autoservire și luarea deciziilor bazate pe date sunt extinse către toate datele unei organizații, oriunde s-ar afla, indiferent dacă au fost încărcate în Data Lake sau nu (Gorelik, 2023).

Cercetările au relevat și apariția conceptului de Data Swamp, considerat o situație nefavorabilă garantării acurateței rezultatelor returnate de procesele de analiză a datelor. Pentru a se asigura că un Data Lake nu va deveni Data Swamp, specialiștii trebuie să se preocupe de elaborarea cu claritate și furnizarea meta-datelor necesare, de conectarea analizelor la Data Lake, de stabilirea modului de măsurare a performanței Data Lake-ului și de partiționarea datelor (Greenfield, 2022).

2.3.7 Data Lakehouse

Un alt concept relativ recent apărut în literatură este cel de Data Lakehouse, definit ca o arhitectură hibridă de gestionare a datelor ce combină beneficiile de flexibilitate și scalabilitate ale unui Data Lake cu structurile de date și caracteristicile de gestionare a datelor ale unui depozit de date (HPE, 2023). Sistemele informatice bazate pe cloud computing permit organizațiilor să îmbine, din ce în ce mai frecvent, tehnologiile Data Lake și Data Warehouse într-o singură arhitectură denumită „Data Lakehouse”. Beneficiile unui Data Lakehouse includ o mai bună integrare, mai puține transferuri de date, o mai bună guvernare a datelor și asistență pentru mai multe cazuri de utilizare (Oracle, 2023b).

2.3.8 Abordări agile și flexibile: deschideri către tendințe privind evoluția arhitecturilor de date

Încă de acum aproape 15 ani, în literatura de specialitate erau consemnate arhitecturi SWARMS - sisteme de management al rețelei de senzori pe suprafață largă -, implementate pentru gestiunea unor incendii în Idaho, USA (Gruenwald, et al., 2007). Se semnala atunci nevoia unor sisteme software

middleware, capabile să gestioneze de la distanță numeroase noduri ale unei rețele de senzori wireless (în engleză, *Wireless Sensor Network, WSN*) instalați în regiuni geografice disparate și se anunțau cercetări în direcția sporirii scalabilității și flexibilității SWARMS, în vederea funcționării la cote bune de performanță atunci când middleware-ul este supus unor încărcări cu date senzoriale de mii de pachete pe secundă. Mai jos este prezentată arhitectura propusă de autori pentru evaluarea performanței acestui tip de sistem (Figura 9).

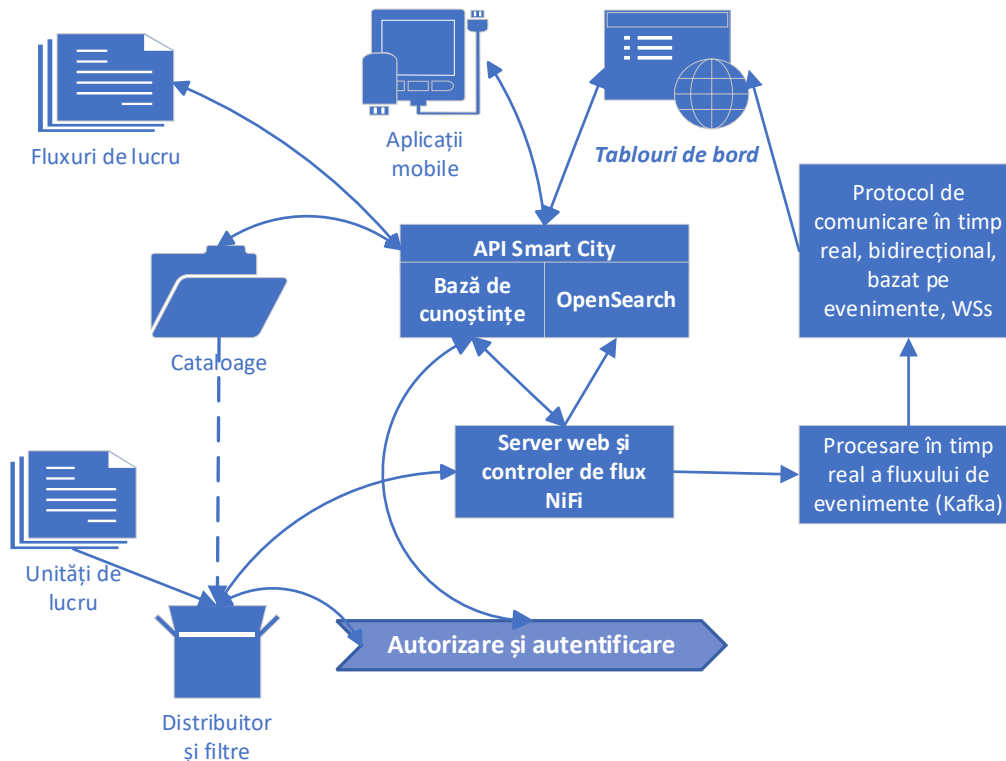


Figura 9. Arhitectura de evaluare a performanței sistemelor de tip SWARMS (sursa: Bellini, et al., 2023)

În prezent, sunt propuse demersuri de cercetare pentru dezvoltarea unor arhitecturi bazate pe broker, centrate pe date, în timp ce majoritatea soluțiilor tehnologice funcționale, centrate pe depozit de date (*Data Warehouse, DW*) și lac de date (*Data Lake, DL*) includ în arhitecturile lor instrumente ETL (Extract/Transform/Load), "data-driven" mai degrabă decât conduse pe evenimente (Bellini, et al., 2023). Cercetătorii atrag atenția asupra trendului de creștere a modelelor de date și a necesității asigurării



interoperabilității back-office a sistemelor, a reducerii timpului de configurare pentru a detecta eficient și a putea procesa structuri necunoscute de date.

Pinon et al. (2023) investighează rolul instrumentelor de tip Self-Service Business Intelligence (SSBI) în sporirea gradului de reactivitate a decidenților cu formare profesională non-IT. Din punct de vedere practic, aceasta înseamnă că managerii din business sunt asistați în formularea interogărilor proprii, SSBI integrând soluții de traducere a jargonului tehnic (utilizat în limbajele formale precum SQL) și de evitare a supraîncărcării datelor. Cercetătorii aplică metodologia Design Science Research pentru a elabora, pe de o parte, o ontologie de domeniu orientată spre vocabularul aferent depozitelor de date și tipologiei afacerii implicate (Figura 9) și pe de altă parte, o componentă de filtrare a datelor apărute la supraîncărcare prin apelarea unui motor hibrid de recomandare ce combină sisteme semantice și reguli de afaceri (Figura 10).

Artefactele propuse utilizează un sistem de recomandări (*Recommender System, RS*) configurat pentru a detecta, dintr-un volum de itemi, pe cei mai susceptibili de a interesa utilizatorul. Dintre cele trei categorii de RS existente în practica de specialitate - *Collaborative Filtering RS*, *Content-Based RS* și *Knowledge-based RS* –, Pinon et al. (2023) au apelat la *Knowledge-Based RS (KBRS)* bazat pe Ontologie, adică un sistem dotat cu o bază de cunoștințe formalizate ontologic. În baza de cunoștințe a unui sistem RS sunt stocate volume mari de informații despre utilizator și itemi, cu scopul de a identifica elementele care se potrivesc cel mai bine cerințelor formulate de specialiștii din business atunci când compun o interogare și a le oferi recomandări relevante.

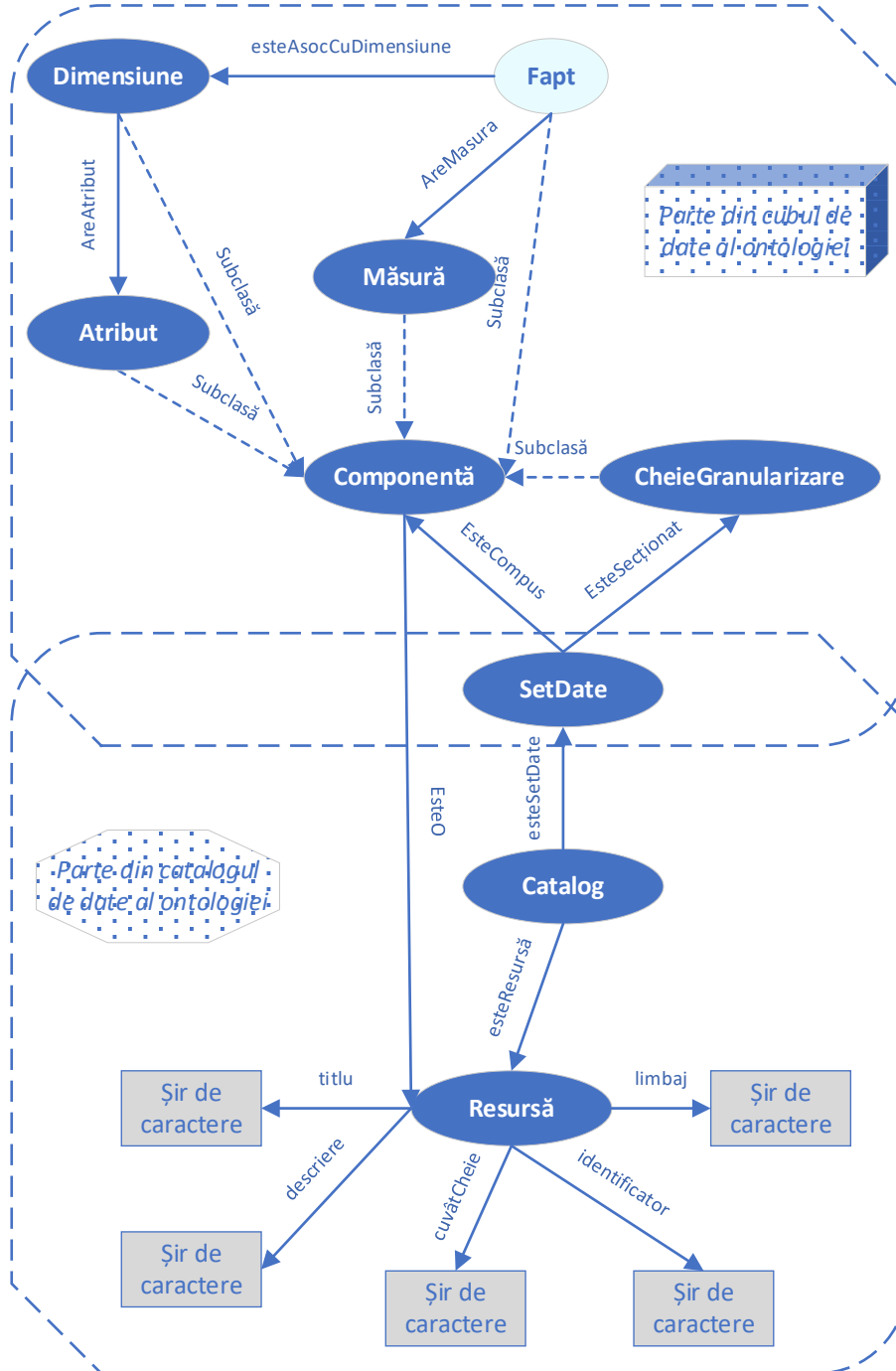


Figura 10. Ontologie dedicată exploatării depozitelor de date (Data Warehouse, DWH) și raportărilor tipice BI (sursă: Pinon et al., 2023)

Deși cadrul general proiectat returnează utilizatorului câmpurile din structurile de date necesare alcătuirii interogării sale, scrierea propriu-zisă a interogării și execuția sa la nivelul aplicației de Business Intelligence pentru a obține datele de analizat rămâne în sarcina acestuia. Așadar, studiul Pinon et al.

(2023) furnizează recomandările câmpurilor de date, valorificând ontologia de afaceri DWH și motorul de recomandări.

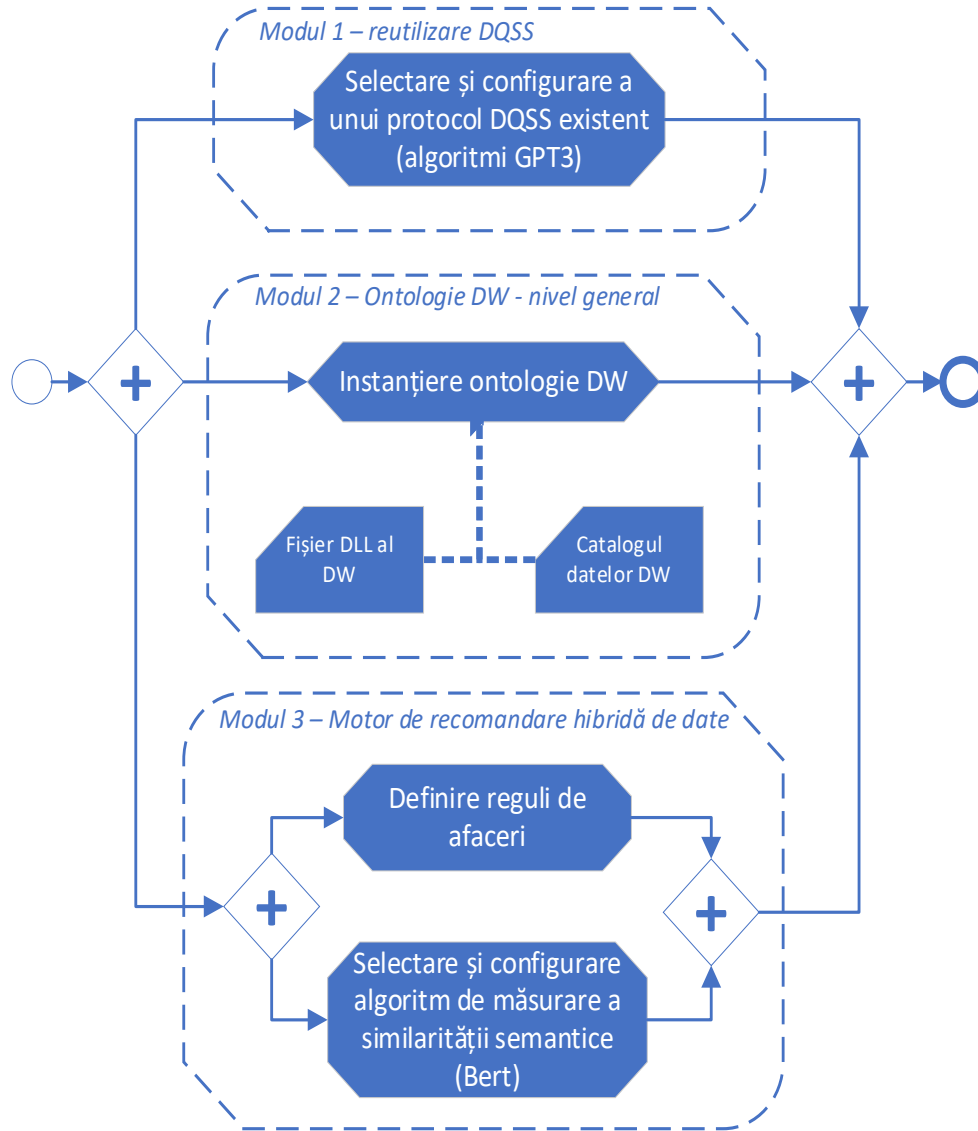


Figura 11. Cadru general de derulare a procesului de utilizare a ontologiei DWH în formularea interogărilor într-o aplicație de Business Intelligence (sursă: Pinon et al., 2023)

2.3.9 Depozite de date

În societatea actuală bazată pe cunoștințe, capacitatea de a valorifica și de a utiliza datele în mod eficient este esențială pentru obținerea unui avantaj competitiv, dar și pentru o dezvoltare sustenabilă a organizațiilor. Acest fapt a determinat apariția conceptului de depozit de date (data warehouse), un depozit centralizat care permite organizațiilor stocarea și analizarea unor volume mari de date.

Încă de la sfârșitul anilor 1980, când organizațiile începeau să se confrunte cu provocările legate de gestionarea unor volumelor din ce în ce mai mari de date, a apărut idee necesității stocării și prelucrării acestora într-un mod eficient. Termenul „depozit de date” a fost introdus de Bill Inmon în 1990, care l-a definit ca „o colecție de date orientată pe subiect, integrată, variabilă în timp și nevolatilă în sprijinul procesului decizional al managementului” [Inmon, 1992]. Definiția lui Inmon a evidențiat caracteristicile fundamentale ale unui depozit de date:

- Orientat către subiect: datele dintr-un depozit sunt organizate în jurul unor subiecte specifice sau zone de interes din cadrul unei organizații, cum ar fi vânzările, comportamentul clienților, gestiunea colectării deșeurilor etc.
- Integrat: Datele provin din surse diverse și sunt integrate într-un formă unitară și consistentă, asigurându-se că pot fi analizate împreună fără conflicte sau inconsecvențe.
- Variabil în timp: Un depozit de date menține date istorice, permițând analiza tendințelor și modelelor în timp, ceea ce este esențial pentru ca procesul adoptării deciziilor să fie eficient.
- Non-volatil: din momentul integrării datelor în depozit, acestea nu mai pot fi modificate, asigurându-se că înregistrările istorice rămân intacte pentru analiză.

Cu toate ca datele stocate într-un depozit de date provin din surse eterogene, acesta este organizat și optimizat pentru interogare și analiză, fiind folosit în generat pentru fundamentarea deciziilor tactice și strategice.

Depozitele de date sunt diferite de bazele de date tradiționale din mai multe perspective. În primul rând, depozitele de date sunt de obicei proiectate pentru sarcini de lucru analitice, în timp ce bazele de date tradiționale sunt concepute pentru sarcini de lucru tranzacționale. În al doilea rând, depozitele de date stochează datele la nivel agregat, ceea ce facilitează interogarea și analiza acestora într-un interval de timp rezonabil. Depozitele de date sunt un instrument esențial pentru organizațiile de toate dimensiunile, fiind utilizate pentru îmbunătățirea procesului decizional, optimizarea operațiunilor curente și identificarea de noi oportunități.

În ultimii ani, depozitarea datelor (data warehousing) a devenit o componentă indispensabilă a afacerilor moderne, apariția big data, facilitată de dezvoltarea utilizării dispozitivelor IoT (Internet of Things) și Industrial IoT, a subliniat necesitatea unor capacități solide de stocare și procesare a datelor.

Progrese tehnologice

- Cloud-Based Data Warehousing: Platforme cloud precum Amazon Web Services (AWS), Google Cloud Platform (GCP), și Microsoft Azure oferă soluții scalabile și avantajoase pentru stocarea și procesarea unor volume mari de date comparativ cu depozitele de date stocate pe serverele locale.
- In-Memory Computing: tehnologii precum SAP S/4 HANA și Oracle Exadata folosesc calculul în memorie, permițând extragerea și analiza mult mai rapidă a datelor.
- Integrarea Data Lakes: Data Lakes, care poate stoca atât date structurate, cât și nestructurate, sunt integrate cu soluții de depozitare a datelor pentru a oferi un mediu analitic mai cuprinzător.

Securitatea și governanța datelor

Pe măsură ce confidențialitatea și conformitatea datelor au devenit din ce în ce mai importante, governanța datelor este un aspect esențial al depozitării datelor. Organizațiile implementează politici și proceduri solide pentru a asigura calitatea, integritatea și securitatea datelor.

Integrarea tehnicilor avansate de analiză

Platformele de depozitare a datelor evoluează pentru a sprijini tehnici avansate de analiză, inclusiv Machine Learning, Deep Learning și inteligența artificială. Acest fapt poate ajuta companiile să reducă costurile și să îmbunătățească eficiența operațiunilor de depozitare a datelor prin automatizarea proceselor asociate cu depozitarea datelor (de exemplu extragerea, transformarea și încărcarea datelor). Totodată, această integrare oferă posibilitatea organizațiilor să obțină perspective și estimări mai profunde din analiza datelor existente. În același timp, integrarea tehnologiilor de tip de business intelligence (BI), cum ar fi vizualizarea datelor și analiza de afaceri vine în sprijinul organizațiilor pentru obținerea de informații suplimentare din datele existente, adoptarea unor decizii mai bune și implicit creșterea performanțelor organizaționale.

Viitorul depozitelor de date este promițător. Depozitele de date devin o componentă din ce în ce mai importantă pe măsură ce organizațiile colectarea și stocarea datelor crește într-un ritm exponențial.

Una dintre tendințele cheie în viitorul depozitelor de date este adoptarea depozitelor de date în timp real. Depozitele de date în timp real pot permite organizațiilor să analizeze datele de îndată ce

acestea sunt generate, fapt care vine în sprijinul adoptării mai rapide a deciziilor și la răspunsuri mai rapide la schimbările din piață.

Augmented analytics combină Machine Learning (ML) și Natural Language Processing (NLP) pentru a automatiza pregătirea datelor, generarea de informații și storytelling. Această tendință va democratiza accesul și analiza datelor, permițând unei game mai largi de utilizatori să obțină informații semnificative.

Arhitectura Data Mesh propune o abordare descentralizată a arhitecturii datelor, în care datele sunt tratate ca un produs. Această abordare pune accent pe proprietatea și scalabilitatea orientate către domeniu, permițând organizațiilor să gestioneze și să utilizeze eficient datele.

Apariția calculatoarelor cuantice promite să revoluționeze capacitățile de procesare a datelor. Calculatoarele cuantice au potențialul de a rezolva probleme analitice complexe la viteze care înainte erau de neimaginat, deschizând noi frontiere în depozitarea și analiza datelor.

Construirea unui depozit de date la nivel național pentru sistemul "Plătești pentru cât arunci" (PPCA) este esențială din mai multe motive:

- Gestionarea centralizată a datelor: un depozit de date la nivel național permite centralizarea tuturor datelor relevante referitoare la generarea, colectarea și eliminarea deșeurilor menajere, fapt care permite o gestionare și accesibilitate mai eficientă a datelor.
- Gestionarea surselor de date eterogene: un depozit de date poate agrega date din diverse surse (localități, companii de colectare a deșeurilor, organisme de reglementare etc.), setul de date obținut poate oferi o viziune holistică a practicilor de gestionare a deșeurilor la scară națională.
- Integrarea surselor de date multiple: sistemul PPCA implică mai multe părți interesate (localități, companii de colectare a deșeurilor, organisme de reglementare etc.), iar într-un depozit de date se pot integra date din aceste surse diverse, permițând o vedere de ansamblu a întregului proces de gestionare a deșeurilor menajere.
- Calitatea și consistența datelor: prin procesul de extragere, transformare și încărcare (Extract, Transform, and Load – ETL), un depozit de date stochează date validate, curățate

și standardizate, acest fapt conduce ca informațiile utilizate pentru adoptarea deciziilor să fie exacte, fiabile și consistente în regiuni sau localități.

- Gestionarea datelor istorice: un depozit de date stochează datele istorice, permițând analiza tendințelor, aspect esențial pentru înțelegerea tiparelor pe termen lung în generarea deșeurilor, ratelor de reciclare și eficacitatea programelor PPCA în timp.
- Analize și instrumente de raportare avansate: un depozit de date acceptă utilizarea instrumentelor avansate de analiză și raportare, fapt care permite generarea de rapoarte și tablouri de bord care ajută la adoptarea deciziilor și formularea viitoarelor politici privind colectarea deșeurilor menajere.
- Modelare predictivă și estimarea colectărilor: datorită stocării datelor istorice, un depozit de date poate permite elaborarea de modele predictivă și de estimare a colectărilor, facilitate care este esențială pentru planificarea viitoarelor strategii de gestionare a deșeurilor menajere, optimizarea alocării resurselor și stabilirea de obiective realiste.
- Monitorizarea și evaluarea performanței: un depozit de date oferă infrastructura necesară pentru monitorizarea indicatorilor cheie de performanță (Key Performance Indicators - KPI) aferente gestionării deșeurilor menajere, oferind suportul pentru evaluarea continuă a eficacității și eficienței programelor PPCA, ajustarea și îmbunătățirea acestora în timp util.
- Conformitatea cu reglementările și auditul: organismele de reglementare pot utiliza depozitul de date pentru a monitoriza conformitatea cu reglementările și standardele naționale/europene de gestionare a deșeurilor menajere, având un instrument deosebit de valoros pentru efectuarea de audituri și pentru asigurarea faptului că toate părțile interesate respectă cerințele legale în vigoare.
- Suport pentru politici bazate pe date: un depozit de date oferă decidenților politici la nivel central datele și informațiile necesare pentru adoptarea deciziilor bazate pe date, oferind sprijin pentru implementarea și ajustarea politicilor de gestionare a deșeurilor menajere, asigurându-se că acestea sunt eficiente și aliniate cu obiectivele naționale și europene.
- Alocarea resurselor: datorită tipologiei datelor stocate privind gestionarea deșeurilor, autoritățile pot alocă resursele într-un mod mai eficient, prin bugetarea pentru

îmbunătățirea infrastructurii, extinderea serviciilor și intervențiilor direcționate în zonele cu provocări specifice de gestionare a deșeurilor menajere etc.

- Evaluare comparativă și partajare a celor mai bune practici: un depozit de date la nivel național permite evaluarea performanței în și între diferite regiuni sau localități, fapt care poate promova schimbul de cunoștințe și adoptarea celor mai bune practici, conducând la obținerea de practici de gestionare a deșeurilor menajere mai eficiente și mai durabile la nivel național.

2.3.10 Extragerea, transformarea și încărcarea datelor

Datele brute sunt adesea dispersate, nestructurate și eterogene, ceea ce face dificilă stocarea, prelucrarea și obținerea de informații semnificative. Aici intervine procesul de extragere, transformare și încărcare a datelor (Extract, Transform, and Load - ETL), care servește este mecanism fundamental pentru integrarea, curățarea și pregătirea datelor pentru analiză.

Procesul de extragere, transformare și încărcare este cheia de bază a strategiilor moderne de gestionare a datelor. Acesta cuprinde trei etape cruciale: extragerea, transformarea și încărcarea datelor, fiecare jucând un rol distinct în călătoria transformării datele brute în informații utile.

Etapa de extracție inițiază procesul ETL prin preluarea datelor din diverse surse eterogene de date. Aceste surse pot varia de la baze de date relaționale, fișiere text și foi de calcul până la servicii web, baze de date online etc. Procesul de extracție variază în funcție de tipul și complexitatea sursei de date. Sursele găzduiesc adesea date în formate eterogene, de la date structurate (cum ar fi tabelele din baze de date relaționale) până la date semistructurate (cum ar fi fișierele JSON sau XML) și date nestructurate (de exemplu, documente text sau imagini).

Unul dintre considerentele principale în timpul fazei de extracție este reprezentat de volumul și viteza de "generare" datelor. Pentru seturi mari de date sau fluxuri de date în timp real, sunt utilizate metode de extracție incrementală, cum ar fi Change Data Capture (CDC). Această abordare identifică și extrage doar înregistrările modificate sau noi de la extragerea anterioară, reducând timpul de procesare și utilizarea resurselor. În plus, există o serie de instrumente specializate care facilitează etapa de extracție. De exemplu, instrumente precum Apache Nifi și Talend simplifică integrarea datelor oferind interfețe grafice pentru configurarea fluxurilor de date. Aceste instrumente acceptă diverse protocoale și API-uri, facilitând conectarea la diferite surse de date.

Etapa de transformare este cea în care datele brute sunt supuse unei serii de operațiuni pentru a le asigura calitatea, consistența și compatibilitatea cu sistemul în care urmează să fie stocate. Acest pas este adesea cel mai complicat și mai consumator de timp în cadrul procesului ETL.

Transformarea datelor cuprinde mai multe operațiuni cheie:

- Curățarea și validarea datelor: Această operațiune implică identificarea și corectarea erorilor, duplicatelor și inconsecvențelor din cadrul datelor. De exemplu, valorile lipsă pot fi completate automat pe baza unor șabloane care se pot modifica dinamic în funcție de datele existente, iar datele eronate pot fi corectate. Verificările de validare confirmă faptul că datele respectă restricțiile definite sau regulile de afaceri.
- Agregarea și normalizarea datelor: în unele cazuri, este necesară agregarea datelor, condensându-le în forme mai ușor de gestionat. În schimb, normalizarea reduce redundanța prin organizarea datelor în tabele structurate, optimizând spațiul utilizat pentru stocare.
- Îmbogățirea și standardizarea datelor: Sursele de date externe pot fi valorificate pentru a spori seturile de date existente, oferind context suplimentar sau îmbogățindu-le cu date suplimentare. Standardizarea implică conversia datelor într-un format uniform, asigurând coerența întregului set de date.

Tehnicile de transformare variază foarte mult în funcție de cerințele specifice ale datelor și de obiectivele procesului ETL. De exemplu, interogările SQL, limbaje precum Python sau R și instrumentele ETL specializate oferă o gamă de opțiuni pentru executarea acestor operațiuni.

Etapa de încărcare încheie procesul ETL prin transferarea datelor transformate într-un sistem destinație (un depozit de date, o bază de date, un instrument de raportare sau orice altă destinație concepută pentru a facilita analiza și raportarea).

Alegerea metodei de încărcare adecvate depinde de factori precum volumul de date, frecvența actualizărilor și capacitățile sistemului destinație. Pentru sistemele orientate pe loturi, "bulk loading" este adesea folosită, datele fiind încărcate în loturi predefinite la intervale de timp prestabilite. Pe de altă parte, sistemele în timp real necesită încărcare continuă, asigurându-se că sistemul țintă reflectă cele mai recente date colectate.

Diverse instrumente și platforme susțin suportă etapa de încărcare, prin oferirea de funcționalități precum validarea datelor, gestionarea erorilor și procesarea paralelă, caracteristici care sporesc eficiența și fiabilitatea procesului de încărcare a datelor.

2.3.11 Tablouri de bord

În lumea actuală bazată pe colectare, stocarea și prelucrarea datelor, în care termeni ca baze de date, depozite de date, big data sunt foarte prezenți, tabloul de bord este instrument esențial pentru organizații. Un tablou de bord reprezintă o interfață vizuală, prezentând date și informațiile esențiale într-un format condensat, ușor de înțeles și care vine în sprijinul fundamentării deciziilor.

Se poate afirma că, conceptul de tablouri de bord își are rădăcinile în cabina de pilotaj a unei aeronave, unde era necesar ca informațiile critice să fie afișate într-un format concis pentru ca piloții să poată adopta deciziile cele mai bune în timp util. Această paradigmă a început să fie aplicată și în domeniul business intelligence, iar în contextul creșterii puterii de procesare a calculatoarelor, a apariției unor interfețe ușor de utilizat, tablourile de bord au evoluat de la simple afișări statice la interfețe dinamice, interactive și user-friendly cu utilizatorul final.

O clasificare a tablourilor de bord le grupează pe acestea în:

- **Tablouri de bord operaționale**
Tablourile de bord operaționale se concentrează pe date în timp real, oferind perspective imediate asupra operațiunilor zilnice ale unei organizații. Acestea permit monitorizarea indicatorilor cheie de performanță (key performance indicator - KPI), urmărirea progresului înregistrat pentru atingerea obiective și răspunsul rapid la problemele emergente.
- **Tablouri de bord strategice**
Tablourile de bord strategice sunt destinate pentru a sprijini planificarea și adoptarea deciziilor pe termen lung. Ele oferă o viziune la nivel înalt asupra performanței organizaționale și ajută la alinierea activităților cu obiective strategice ale organizației.
- **Tablouri de bord analitice**
Tablourile de bord analitice sunt destinate pentru o analiză aprofundată a datelor. Acestea permit explorarea de seturi de date complexe, efectuarea de analize avansate și

descoperirea tendințelor și modelelor "ascunse", care nu sunt evidente sau ușor de observat/identificat.

Tablourile de bord au evoluat de la originile în aviație și au devenit instrumente indispensabile în majoritatea industriilor. Adaptabilitatea lor, funcționalitatea în timp real și capacitatea de a prezenta date complexe în elemente vizuale ușor de înțeles și utilizat le-au transformat în instrumente esențiale pentru factorii de decizie de la toate nivelurile. Pe măsură ce tehnologia continuă să avanseze, ne putem aștepta ca tablourile de bord să joace un rol și mai important în modelarea modului în care organizațiile funcționează și iau decizii. În lumea actuală, cu schimbări extrem de rapide și "condusă" de date, adoptarea acestei paradigme nu este doar o opțiune, ci o necesitate strategică în contextul în care organizațiile doresc să se dezvolte și să rămână competitive în domeniul lor de activitate.

2.3.12 Considerente în alegerea unui instrument pentru vizualizarea datelor analitice la nivel tactic și strategic: Power BI vs. Tableau vs. QlikView

Lumea afacerilor moderne este extrem de dinamică, iar adoptarea unor decizii informate este esențială pentru orice organizație. Un rol important pentru îndeplinirea acestui deziderat este reprezentat de instrumentele de Business Intelligence (BI), care oferă organizațiilor mijloacele de a transforma datele în informații utile. Printre principalii competitori în arena BI se numără Power BI, Tableau și QlikView.

2.3.12.1 Power BI

Power BI, dezvoltat de gigantul tehnologic Microsoft, a apărut ca un competitor formidabil în domeniul Business Intelligence. De la începuturile sale, Power BI a suferit o evoluție remarcabilă, aliniindu-se la cerințele în continuă evoluție ale procesului decizional bazat pe date.

Printre caracteristicile cheie ale Power BI se numără multitudinea de funcții și numărul semnificativ de grafice. Totodată, capabilitățile sporite de pregătire a datelor simplifică procesul de curățare, transformare și integrare a datelor din diverse surse. Unul dintre punctele forte ale Power BI este integrarea sa perfectă în ecosistemul Microsoft. Această relație sinergică cu Excel, SQL Server, SharePoint și alte aplicații Microsoft oferă utilizatorilor un mediu familiar, accelerând curba de învățare. În plus, Power BI beneficiază de o comunitate de utilizatori robustă și activă, asigurând o mulțime de resurse, forumuri și tutoriale pentru utilizatorii care caută îndrumare. În ciuda numeroaselor sale avantaje, Power BI nu este lipsită de limitări. Gestionarea seturilor de date foarte mari poate reprezenta

o provocare, care poate necesita configurații și resurse avansate sau migrarea în cloud-ul oferit de Microsoft.

2.3.12.2 Tableau

Tableau, un produs Tableau Software, s-a impus ca unul dintre primele și cei mai importanți instrumente de tip Business Intelligence. Prin inovațiile pe care le-a adus și abordarea centrată pe utilizator, Tableau a devenit una dintre alegerile preferate pentru organizațiile care doresc o reprezentare dinamică și ușor de utilizat a datelor din punct de vedere vizual.

Caracteristica principală a Tableau este reprezentată de capabilitățile oferite pentru vizualizarea datelor. Produsul oferă o interfață intuitivă care permite utilizatorilor să creeze fără eforturi tablouri de bord și rapoarte interactive. Tableau dispune de o bibliotecă extinsă de opțiuni de vizualizare, asigurându-se că utilizatorii își pot transmite în mod eficient datele și informațiile. Principalul avantaj oferit de Tableau constă în designul său centrat pe utilizator, care facilitează explorarea rapidă a datelor și crearea de tablouri de bord. Produsul oferă posibilitatea utilizatorilor, indiferent de expertiză tehnică, să creeze relativ ușor tablouri de bord. În plus, comunitatea utilizatorilor Tableau încurajează o cultură a partajării cunoștințelor și a asistenței, oferind o sursă vastă de resurse. Tableau are o serie de limitări, în special, costul edițiilor pentru întreprinderi poate fi prohibitiv. Totodată, în timp ce tablourile de bord realizate cu Tableau sunt dinamice și captivante, elementele mai complexe pot necesita o curbă de învățare mai abruptă pentru utilizatorii care nu sunt familiarizați cu funcțiile avansate.

2.3.12.3 Qlikview

QlikView, dezvoltat de Qlik, oferă o abordare distinctă a Business Intelligence, și-a creat o nișă cu modelul său de date asociativ, care revoluționează modul în care utilizatorii interacționează și explorează datele.

Caracteristica remarcabilă a QlikView este modelul său asociativ de date, care permite utilizatorilor să exploreze în mod dinamic datele fără a fi nevoie de ierarhii predefinite sau interogări. Această abordare permite o flexibilitate fără precedent în analiza datelor. În plus, QlikView utilizează procesarea în memorie (in-memory processing), care asigură accesarea, procesarea și explorarea extrem de rapide a datelor. Unul dintre punctele forte ale QlikView constă în capacitățile sale rapide de explorare a datelor. Utilizatorii pot naviga fără efort prin seturi de date vaste, realizând conexiuni intuitive și descoperind informații relevante ușor. Mai mult, QlikView excelează în manipularea seturilor mari de



date, oferind o platformă robustă pentru organizațiile care se ocupă cu volume mari de date. Cu toate acestea, QlikView nu este lipsit de limitări. Curba inițială de învățare poate fi mai abruptă în comparație cu alte instrumentele de tip BI. În plus, costurile legate de licențierea QlikView și infrastructura necesară pot să fie destul de ridicate.

3 Arhitectura informatică a sistemului propus

Sistemul informatic pentru analiza și modelarea datelor privind colectările de deșeuri trebuie să răspundă și să se alinieze proceselor care alcătuiesc arhitectura de afaceri a programului PPCA. Arhitectura sistemului informatic suport pentru acest program este alcătuită din două componente majore: sistemul operațional prin care datele sunt colectate la nivelul fiecărei unități administrativ-teritoriale (UAT) și sistemul de asistare a deciziei, având scopul de a include procese de analiză multidimensională a datelor și de modelare a acestora, prin instrumente și funcționalități specifice. În timp ce prima componentă software (definită ca sistem independent) va fi gestionată la nivel local, cea de-a doua va funcționa la nivel central, având rol de monitorizare a activității operaționale din teritoriu. În acest sens au fost identificate mai multe cerințe de afaceri care au stat la baza definirii proceselor și a funcționalităților IT care le deservesc.

- CA1. Fiecare unitate administrativă teritorială (UAT) va transmite în mod individual situația colectărilor în vederea raportării periodice;
- CA2. Analiza și modelarea datelor primite se va realiza în mod centralizat pentru toate UAT care vor implementa programul PPCA;
- CA3. Fiecare raportare privind colectarea deșeurilor prin programul PPCA se va realiza la cererea UAT-ului respectiv pentru o anumită perioadă de timp, denumită perioada de raportare;
- CA4. Fiecare UAT va putea decide care este perioada de raportare a colectării deșeurilor, dar toată raportările unui UAT trebuie să fie în ordine cronologică. Nu se pot transmite date legate de colectările de deșeuri dintr-o perioadă dacă aceasta nu este continuarea ultimei perioade raportate de UAT-ul respectiv;
- CA5. Datele transmise de un UAT pot fi șterse doar dacă acestea nu au fost încă procesate în sistemul informatic de asistare a deciziei;
- CA6. Datele transferate de către un UAT în baza de date centralizată se vor procesa independent de procedura de transfer a acestora. Procesarea datelor va avea ca rezultat popularea bazei de date centralizate cu înregistrările în oglindă ale bazei de date operaționale pentru o anumită perioadă de timp.
- CA7. Componenta software de integrare a sistemului informatic trebuie să asigure coerența datelor transmise de către fiecare UAT, să nu existe lipsuri în datele transmise, iar acestea să fie validate înainte de înregistrarea lor în sistemul pentru asistarea deciziei;
- CA8. Sistemul informatic operațional și cel de asistare a deciziei vor funcționa independent și decuplat unul de celălalt;
- CA9. Sistemul informatic trebuie să permită în orice moment definirea unui nou UAT care să poată raporta situația colectării deșeurilor;
- CA10. Sistemul informatic trebuie să fie ușor integrabil, prin definirea de puncte de acces la care să se conecteze sistemele operaționale de colectare a deșeurilor.



Cerințe tehnice generale:

Sistemul informatic pentru analiza datelor privind colectarea deșeurilor prin programul PPCA a fost dezvoltat utilizând tehnologii multiple, în funcție de componentele software care se regăsesc în arhitectura sa.

Figura 12 prezintă schema generală de funcționare și amplasare a sistemului informatic pentru analiza și modelarea datelor privind colectarea deșeurilor prin programul PPCA.

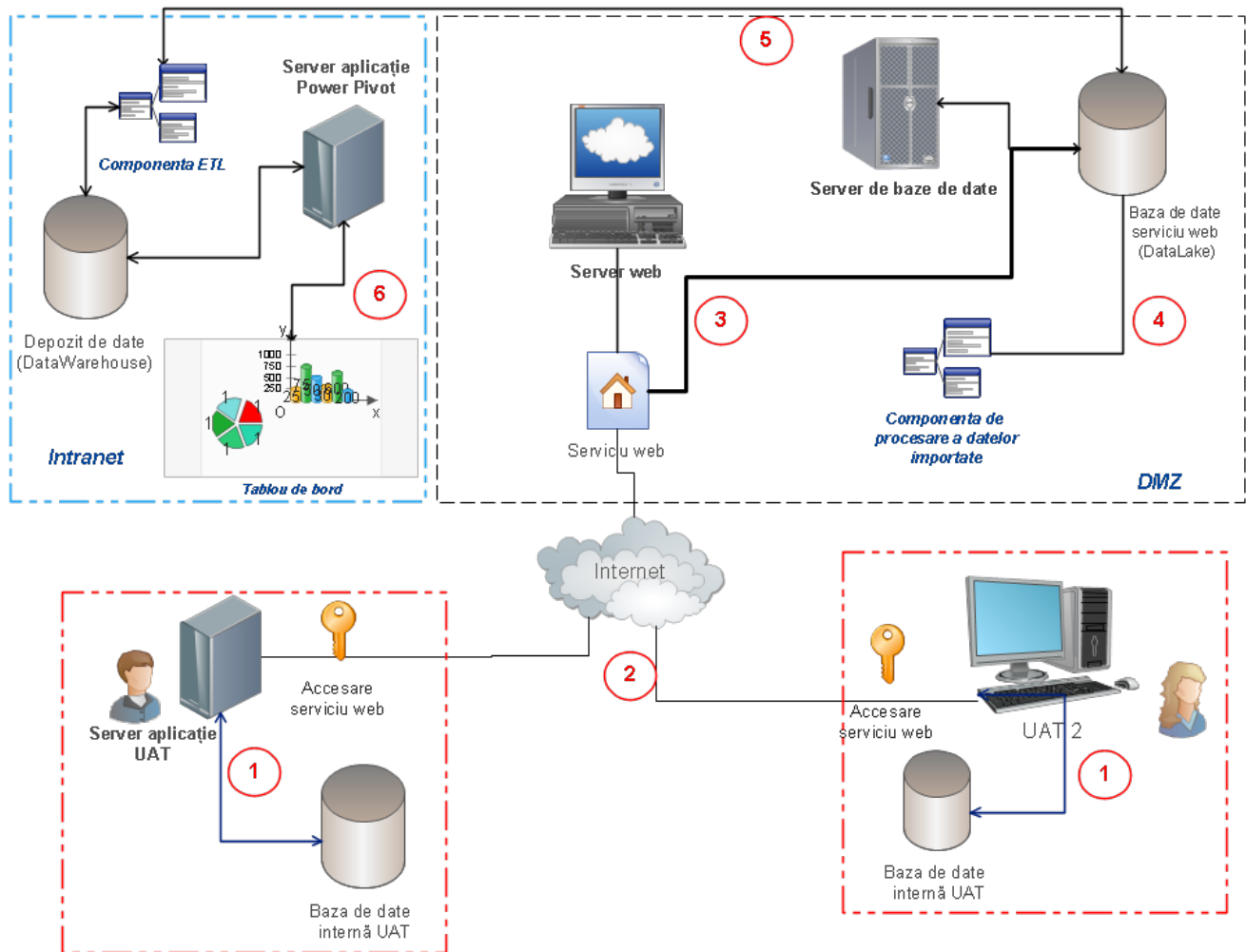


Figura 12. Modelul arhitecturii funcționale a sistemului informatic de analiză a datelor privind colectările de deșeurii prin programul PPCA

1. Extragerea și pregătirea datelor în vederea efectuării transferului acestora în baza de date centralizatoare. Această etapă este realizată de către fiecare UAT printr-o metodă automatizată

de extragere și încărcare a datelor într-un fișier definit într-un format standardizat pentru transferul datelor prin internet.

2. Autentificarea la un serviciu web care asigură comunicarea între cele două componente software și inițierea unui nou transfer de date. Serviciul web este securizat printr-un certificat de tip SSL și permite autentificarea fiecărui UAT pe bază de utilizator și parolă.
3. Fiecare va putea genera un nou transfer de date în format standardizat către baza de date centralizatoare. Într-o primă instanță, acest serviciu web va stoca informația în format standardizat (JavaScriptObjectNotation - JSON) în cadrul bazei de date colectoare, împreună cu informațiile privind importul de date, precum instituția care a generat transferul, data și ora, tipul operației vizate de transfer (adăugare, modificare, ștergere²).
4. Ulterior, toate transferurile reușite vor fi procesate printr-un fir de execuție diferit de cel al serviciului web, spre a nu bloca transferurile de date. Metoda procesării fișierelor primite de la fiecare UAT cu înregistrările ce urmează a fi colectate în baza de date centralizatoare este First In First Out (FIFO). Astfel, la un anumit interval de timp (definit în prealabil de către administratorul bazei de date) sunt citite toate transferurile efectuate de UAT-uri și acestea vor genera înregistrări sau, după caz, vor actualiza sau șterge înregistrări deja existente în cadrul bazei de date centralizatoare de tip DataLake.
5. Componenta ETL a depozitului de date va prelua informațiile din baza de date centralizatoare și le va încărca în structura de date de tip DataWarehouse.
6. Pe baza datelor din depozitul de date, aplicația de tip Tablou de bord permite analiza multidimensională și modelarea datelor în scopul asistării deciziei.

² Detalii privind aceste operațiuni cu datele transferate vor fi oferite mai târziu când se va explica modul de procesare a datelor.

3.1 Descrierea proceselor de afaceri privind componenta de integrare a sistemelor informatice operaționale și de analiză a datelor privind colectările de deșeuri

3.1.1 Etapizarea procesului de afaceri principal privind centralizarea și analiza datelor

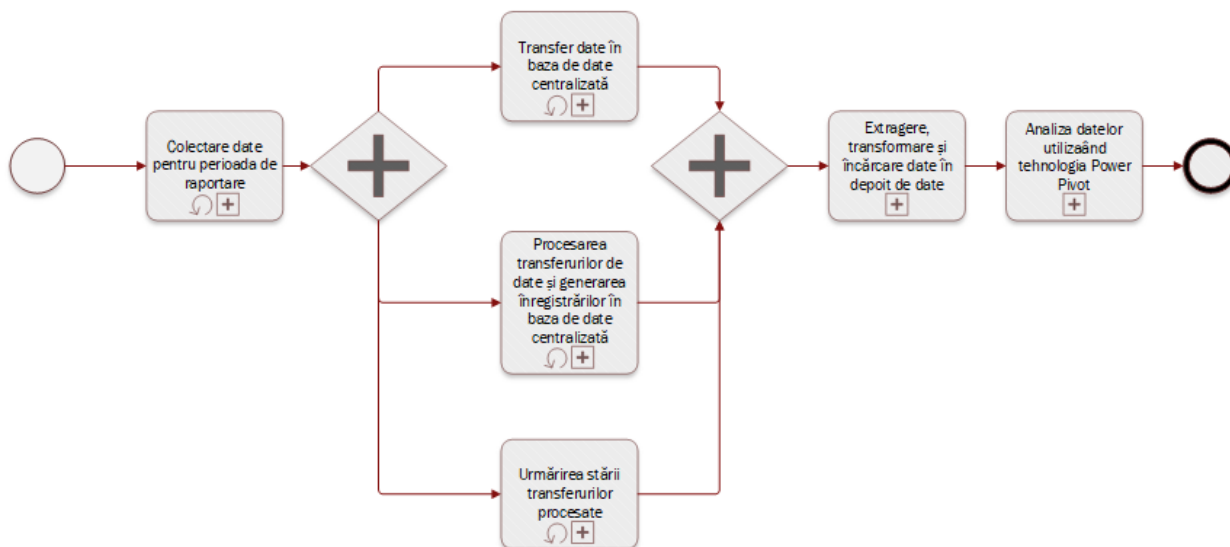


Figura 13. Diagrama BPMN 1. Funcționarea generală a sistemului de analiză și modelare a datelor privind deșeurile colectate prin programul PPCA.

Descriere proces BPMN1:	
Titlu	Funcționarea generală a sistemului de analiză și modelare a datelor privind deșeurile colectate prin programul PPCA
Scop	Detalierea înlănțuirii etapelor privind procesul de colectare, încărcare și analiză multidimensională a datelor privind deșeurile colectate de către UAT
Actori implicați	Sistemul informatic operațional privind evidența colectării deșeurilor Sistemul informatic de asistare a deciziei privind centralizarea și analiza datelor colectate
Pre-condiții de îndeplinire	UAT-ul trebuie să fie înregistrat în prealabil în baza de date centralizatoare
Descriere	Pasul 1: Printr-o procedură automatizată, sistemul informatic operațional de la nivelul fiecărui UAT va colecta datele din cadrul bazei de date pentru o perioadă

	<p>de raportare. Acesta va fi încărcat și stocat printr-o metodă a serviciului web în baza de date centralizatoare a sistemului de asistare a deciziei în format JSON.</p> <p>Pasul 3: Pe un alt fir de execuție sistemul de asistare a deciziei va realiza procesarea datelor din obiectul de tip JSON pentru toate transferurile care încă nu au fost procesate. Practic, toate înregistrările din obiectul JSON vor fi operate în baza de date centralizatoare, marcând importul ca fiind procesat; În paralel cu această etapă, sistemul informatic operațional va putea urmări starea fiecărui transfer realizat anterior pentru a verifica dacă au fost identificate erori de procesare a datelor ce urmează a fi prelucrate de către sistemul informatic de analiză.</p> <p>Pasul 4: Utilizarea tehnicilor de tip ETL din baza de date centralizatoare (de tip DataLake) pentru popularea depozitului de date al sistemului pentru asistarea deciziei;</p> <p>Pasul 5: Analiza multidimensională a datelor și procesarea informațiilor centralizate prin instrumentul Power Pivot;</p>
Frecvența de realizare a fluxului	Diferă pentru fiecare subproces în parte.
Cerințe de afaceri la care răspunde	Toate cerințele de afaceri

3.1.2 Subprocesul de colectare date pentru perioada de raportare

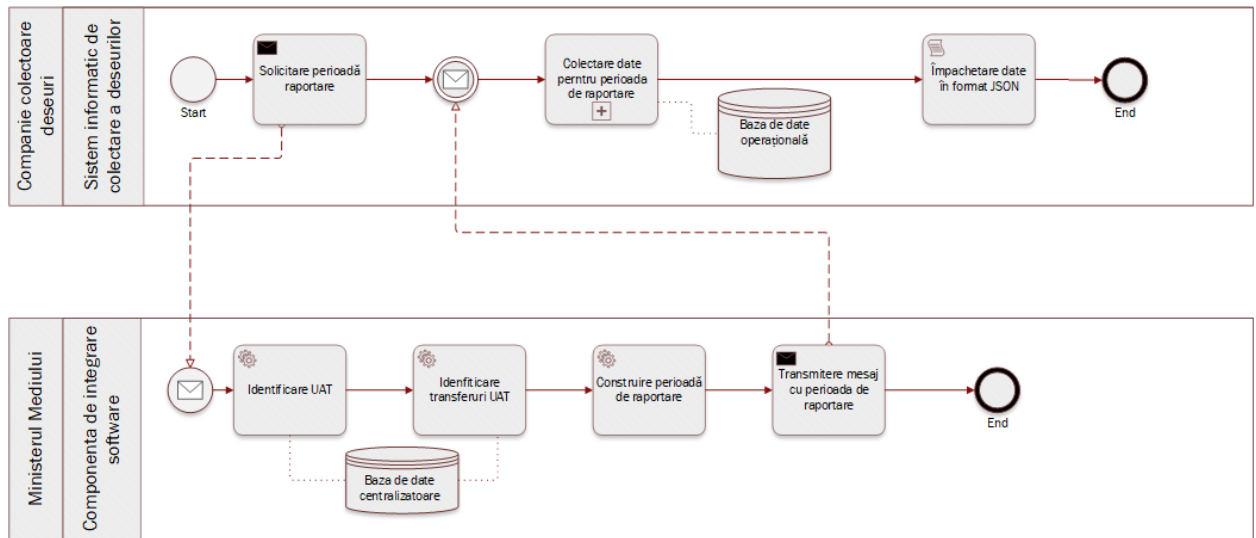


Figura 14. Diagrama BPMN 1.1 privind pregătirea de către un UAT a unui nou transfer de date în baza de date centralizatoare

Descriere proces BPMN 1.1:	
Titlu	Pregătirea datelor privind deșeurile colectate prin programul PPCA în vederea încărcării lor în baza de date centralizatoare a sistemului de asistare a deciziei
Scop	Detalierea înlănțuirii etapelor privind procesul de pregătire a transferului de date către sistemul de asistare a deciziei
Actori implicați	Sistemul informatic operațional privind evidența colectării deșeurilor Sistemul informatic de asistare a deciziei privind centralizarea și analiza datelor colectate
Pre-condiții de îndeplinire	UAT-ul trebuie să fie înregistrat în prealabil în baza de date centralizatoare
Descriere	Sistemul operațional va solicita sistemului de asistare a deciziei perioada pentru care trebuie să facă următoarea raportare. Sistemul informatic de asistare a deciziei va returna perioada pentru care ar trebui să se realizeze următoarea raportare în formatul unui obiect serializabil, pornind de la ultimul import de date de la respectivul UAT. Perioada maximă de raportare este până la ziua anterioară celei la care se solicită o nouă

	<p>raportare. În prealabil este necesară identificarea UAT-ului care a solicitat cererea, precum și datele privind ultimul import realizat de către acesta, dacă este cazul;</p> <p>Colectarea datelor vizate spre a fi transferate către sistemul de tip SIAD va fi realizată prin interogarea tuturor tabelor din baza de date operațională și extragerea înregistrărilor din perioada de raportare.</p> <p>Informațiile alcătuite din înregistrările extrase, precum și tabellele din care acestea fac parte vor fi împachetate în formatul standard JSON de către sistemul informatic operațional.</p>
Frecvența de realizare a fluxului	Frecvența de realizare se va decide în funcție de numărul de tranzacții înregistrate în baza de date operațională de producție. De regulă, o dată la 2-3 zile, dar se poate realiza zilnic.
Cerințe de business la care răspunde	CA1, CA3, CA4

3.1.3 Subprocesul de transfer de date din baza de date operațională în cea centralizatoare

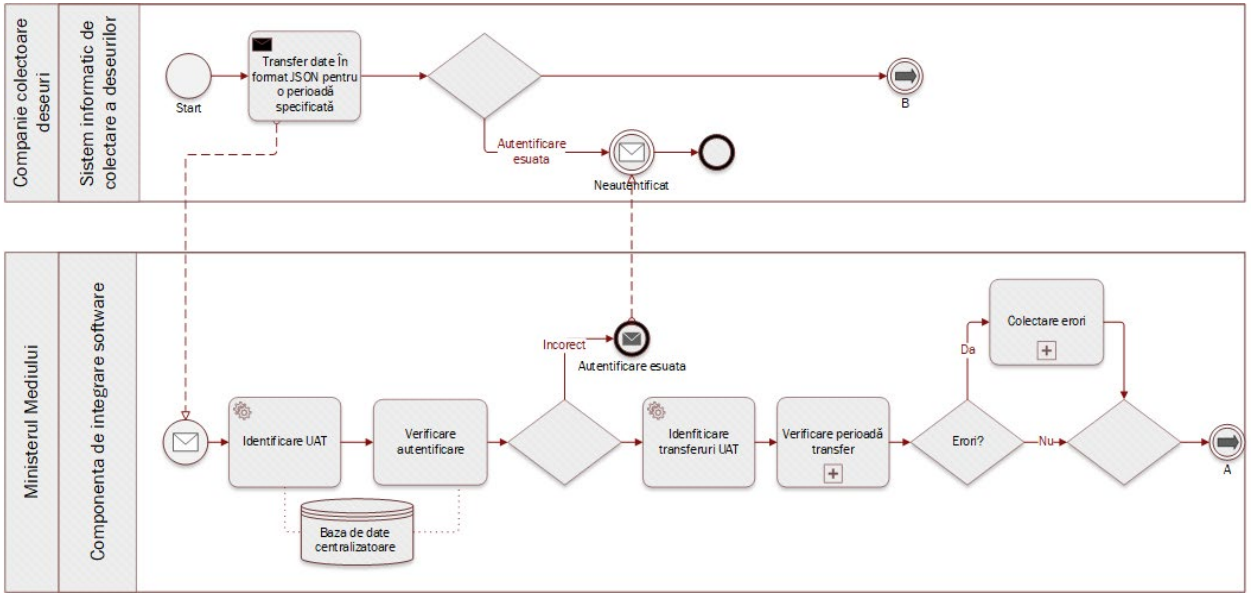


Figura 15. Diagrama BPMN privind subprocesul de transfer date în format JSON în baza de date centralizatoare (partea 1)

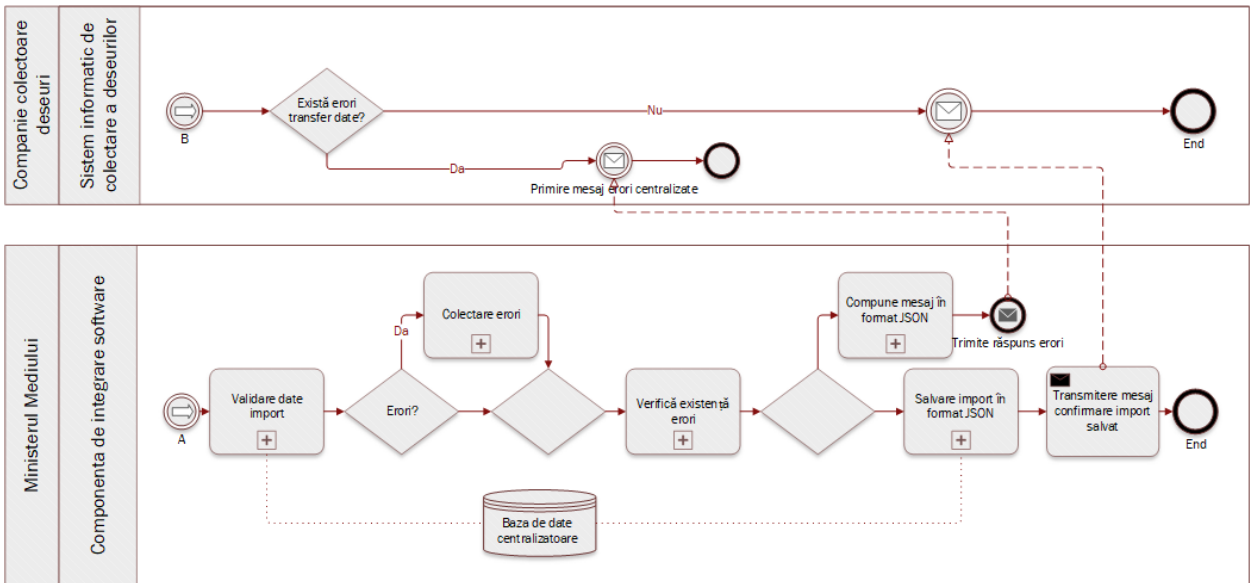


Figura 16 Diagrama BPMN privind subprocesul de transfer date în format JSON în baza de date centralizatoare (partea a doua)

Descriere proces BPMN 1.2:	
Titlu	Transferul datelor din baza de date privind colectările de deșeuri prin PPCA în baza de date centralizatoare a sistemului de asistare a deciziei
Scop	Descrierea etapelor privind procesul de transfer de date către sistemul de asistare a deciziei, în vederea populării ulterioare a depozitului de date
Actori implicați	Sistemul informatic operațional privind evidența colectării deșeurilor Sistemul informatic de asistare a deciziei privind centralizarea și analiza datelor colectate
Pre-condiții de îndeplinire	UAT-ul trebuie să fie înregistrat în prealabil în baza de date centralizatoare
Descriere	<p>Datele preluate din baza de date operațională și pregătite în format JSON potrivit subprocesului descris anterior, se vor transfera în baza de date centralizatoare prin utilizarea unei metode a serviciului web al componentei de integrare a sistemului de asistare a deciziei.</p> <p>În cadrul metodei de transfer de date în format standardizat JSON se vor efectua mai întâi o serie de verificări pentru validarea inițială a importului.</p> <p>Dacă autentificarea prin serviciul web a eșuat, verificările ulterioare nu mai au sens, procesul fiind încheiat prin notificarea sistemului operațional printr-un mesaj.</p> <p>Pentru verificarea validității importului, serviciul web va colecta erorile de validare, dacă acestea există, cu privire la formatul JSON al fișierului primit, existența datelor de import, sau validitatea perioadei de raportare.</p> <p>Dacă există erori de import, serviciul web va compune un mesaj în format JSON pentru a transmite aceste erori cât mai explicit sistemului operațional.</p> <p>În cazul în care nu sunt identificate erori de transfer, serviciul web salvează în baza de date centralizatoare datele de import și fișierul în format JSON, conținând înregistrările care urmează a fi procesate.</p>

	Erorile înainte de înregistrarea transferului în format JSON în baza de date vor fi returnate sistemului operațional ca răspuns la încercarea de import date, fără ca această operațiune să fie finalizată.
Frecvența de realizare a fluxului	Frecvența de realizare se va decide în funcție de numărul de tranzacții înregistrate în baza de date operațională de producție.
Cerințe de business la care răspunde	CA1, CA3, CA4

3.1.4 Subprocesul privind procesarea transferurilor de date și generarea înregistrărilor în baza de date centralizatoare

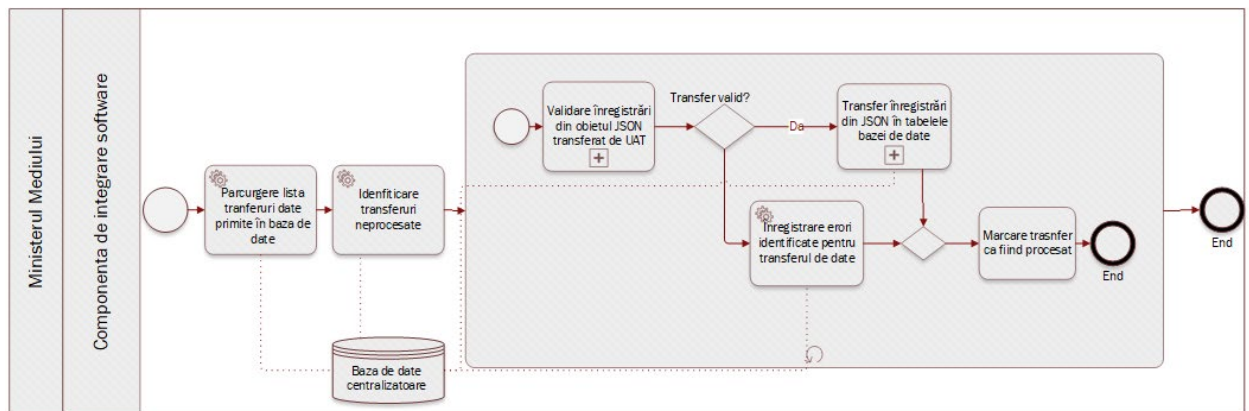


Figura 17. Diagrama BPMN privind procesarea transferurilor de date privind colectarea deșeurilor prin programul PCCA în baza de date centralizatoare a sistemului de asistare a deciziei

Descriere proces BPMN 1.3:	
Titlu	Procesarea transferului de date în baza de date centralizatoare
Scop	Descrierea etapelor privind procedura de procesare a transferului de date din format JSON în înregistrări ale bazei de date centralizatoare a sistemului de asistare a deciziei
Actori implicați	Sistemul informatic operațional privind evidența colectării deșeurilor Componenta de procesare a datelor a sistemul informatic de asistare a deciziei privind centralizarea și analiza datelor colectate

Pre-condiții de îndeplinire	Obiectul în format JSON ce conține datele de transferat trebuie să existe în baza de date centralizatoare.
Descriere	<p>La un interval de timp stabilit de administratorul sistemului de asistare a deciziei, componenta de procesare a datelor va citi din baza de date importurile de date (obiectele de tip JSON) neprocesate.</p> <p>Pentru fiecare transfer neprocesat sistemul de asistare a deciziei, prin componenta de procesare a datelor va efectua verificările necesare pentru asigurarea coerenței și integrității datelor importate.</p> <p>Dacă în urma validării rezultă erori identificate în cadrul înregistrărilor care ar trebui procesate, sistemul informatic va jurnaliza aceste erori în relația cu obiectul JSON procesat.</p> <p>Un obiect JSON care nu deține erori în urma validării va putea fi procesat prin operarea în baza de date centralizată a modificărilor necesare sincronizării acesteia cu baza de date operațională care a transmis datele. În acest sens, operațiile efectuate în tabelele bazei de date centralizatoare pot fi cele de adăugare de noi înregistrări, de modificare a înregistrărilor existente sau de ștergere a unor înregistrări care nu mai există în baza de date operațională.</p> <p>Serviciul de procesare a obiectelor JSON din cadrul componentei de procesare a sistemului de analiză a datelor colectate va efectua prelucrarea tuturor transferurilor.</p> <p>În urma procesării datelor sistemul operațional poate vizualiza prin intermediul unei metode a serviciului web lista transferurilor și statusul acestora.</p>
Frecvența de realizare a fluxului	Frecvența de realizare se va decide de către administratorul sistemului prin precizarea intervalului de verificare a transferurilor neprocesate de către sistemul de analiză a datelor colectate prin programul PPCA.
Cerințe de business la care răspunde	CA6, CA7, CA8

3.1.5 Subprocesul privind urmărirea stării transferurilor efectuate în baza de date centralizatoare

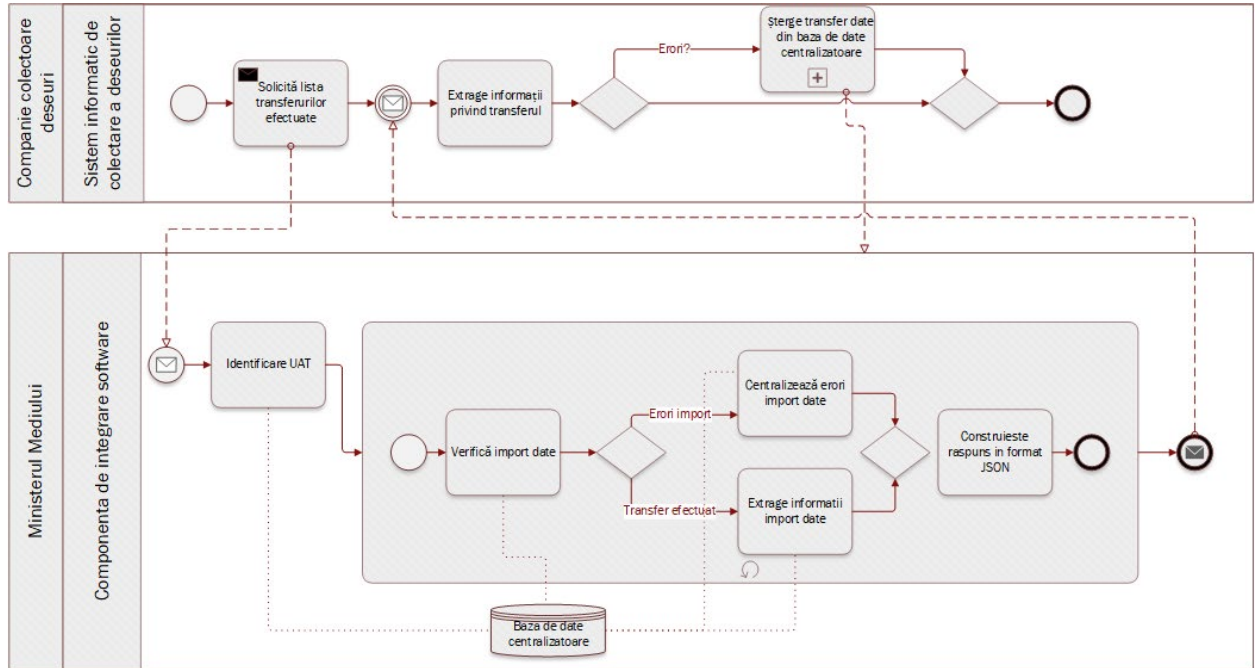


Figura 18. Diagrama BPMN privind urmărirea stării transferurilor efectuate în baza de date centralizatoare

Descriere proces BPMN 1.4:	
Titlu	Urmărirea stării transferurilor de date în baza de date centralizatoare
Scop	Descrierea etapelor privind procedura de urmărire a transferurilor din sistemul informatic operațional de colectare a deșeurilor în baza de date centralizatoare a sistemului de analiză a datelor
Actori implicați	Sistemul informatic operațional privind evidența colectării deșeurilor Componenta de transfer a datelor a sistemul informatic de asistare a deciziei privind centralizarea și analiza datelor colectate
Pre-condiții de îndeplinire	UAT trebuie să dețină un cont activ în baza de date centralizatoare a sistemului de analiză a datelor provenind din colectarea deșeurilor prin programul PPCA

<p>Descriere</p>	<p>La solicitarea sistemului informatic operațional privind colectarea deșeurilor serviciul web pune la dispoziția acestuia o metodă privind centralizarea transferurilor efectuate de respectivul UAT.</p> <p>Pentru fiecare transfer sistemul de asistare a deciziei va centraliza informații privind starea fiecărui import de date salvat în baza de date centralizatoare.</p> <p>Pentru importurile pentru care nu s-a realizat procesarea datelor prin înregistrarea lor în baza de date centralizatoare, precum și pentru cele care au fost deja procesate răspunsul va fi un obiect de tip JSON care va conține informații privind data importului, data procesării, starea procesării. Pentru acele importuri care în urma încercării de procesare a datelor s-au înregistrat erori răspunsul va conține, pe lângă informațiile privind identificarea importului și lista acelor erori, în format JSON. Această etapă va fi una recurentă, pentru fiecare import identificat în baza de date ca fiind încărcat de către UAT-ul care a solicitat lista.</p> <p>Centralizarea importurilor va da naștere listei cu starea acestora și va fi trimisă către sistemul operațional ca răspuns furnizat pentru metoda web apelată prin serviciul web.</p> <p>În urma parcurgerii listei importurilor efectuate, sistemul operațional poate extrage acele importuri pentru care s-au înregistrat erori de procesare și le poate șterge, în vederea transferurilor ulterioare a datelor din perioada respectivă.</p> <p>Ștergerea unui import deja salvat în baza de date operațională se poate realiza prin apelarea unei metode dedicate în acest sens, doar dacă importul nu a fost deja procesat.</p>
<p>Frecvența de realizare a fluxului</p>	<p>Frecvența de realizare se va decide de către administratorul sistemului prin precizarea intervalului de verificare a transferurilor neprocesate de către sistemul de analiză a datelor colectate prin programul PPCA.</p>
<p>Cerințe de business la care răspunde</p>	<p>CA5, CA7, CA8, CA10</p>

3.2 Sistemul operațional de colectare a deșeurilor

De adăugat

3.3 Schimbul electronic de date între componentele sistemului

Pentru dezvoltarea componentei software de integrare a sistemului informatic operațional cu sistemul informatic de asistare a deciziei a fost utilizat mediul de dezvoltare Microsoft Visual Studio în corelație cu gestionarul de baze de date Microsoft SQL Server. La modul general aceasta constă într-o nouă aplicație de tip serviciu web, care va putea fi accesată în mediul internet de către fiecare instanță a sistemului de colectare a deșeurilor care va fi funcțională la nivelul fiecărei UAT. S-au utilizat tehnologii specifice de conectare la bazele de date SQL Server necesare importului de date și transferului informațiilor între baza de date colectoare (de tip DataLake) și bazele de date operaționale.

Popularea tabelor din structura Data Lake

După cum s-a explicat anterior, datele transmise de fiecare unitate administrativ-teritorială (UAT) sunt memorate la nivelul Data Lake doar în ipoteza reprezentării lor conform sintaxei JSON. Pe măsură ce sunt recepționate, aceste șiruri JSON sunt stocate ca atare, în formă brută, în Data Lake (tabelul JSON_import), prelucrarea lor urmând a fi realizată în două etape:

- Validarea, prin raportarea la o structură prestabilită a șirurilor JSON, dar și la un set de reguli care vizează structura și conținutul Data Lake;
- În funcție de tipul operației indicate la importul datelor, se va realiza, după caz, adăugarea, editarea sau ștergerea anumitor înregistrări la nivelul anumitor tabele din Data Lake, toate datele necesare acestor operații putând fi extrase din șirurile JSON validate în prealabil.

Indiferent de natura operației vizate, aceasta este condiționată de rezultatul validării: doar șirurile valide sunt procesate în scopul actualizării conținutului data take, cele invalide generând erori care sunt jurnalizate într-un tabel dedicat (ValidationErrors). Validarea șirurilor JSON importate este o prelucrare absolut necesară, din mai multe considerente:

- Menținerea coerenței datelor, prin conservarea restricțiilor de integritate specifice bazelor de date ce deservește UAT care transmit șirurile JSON. În condițiile în care cheile primare și cheile externe de la nivelul bazelor de date operaționale sunt preluate în Data Lake doar ca nume, nu și ca funcție (tabelele din Data Lake având propriile câmpuri cu

roluri de chei, conform structurii descrise într-o secțiune anterioară), este important ca restricțiile specifice modelului relațional (unicitatea și nenulitatea valorilor cheilor primare, respectiv restricția de integritate referențială, care vizează cheile externe) să se verifice la nivelul subsetului de înregistrări care corespund fiecărei UAT.

- Posibilitatea realizării efective a adăugării, editării sau ștergerii de înregistrări la nivelul Data Lake este condiționată de necesitatea ca datele care fac obiectul acestor operații să se conformeze schemei structurale (metadatelor) aferente fiecărui tabel-destinație. O astfel de verificare este obligatorie întrucât datele importate sunt în format text, iar utilizarea lor presupune, în primă instanță, identificarea tabelelor și câmpurilor vizate din cadrul Data Lake; în plus, în cazul adăugării sau editării de înregistrări, este necesară realizarea tuturor conversiilor care se impun pentru ca datele să poată fi memorate în tabelele-destinație. În mod evident, în ipoteza că datele importate nu sunt compatibile cu câmpurile-destinație (tip de date, acceptarea sau nu a valorilor nule), o astfel de operație ar genera erori, care pot fi însă prevenite prin validarea corespunzătoare a șirurilor JSON salvate în Data Lake.
- Problema menținerii coerenței și acurateței datelor nu se aplică doar la nivel individual, pentru fiecare UAT care transmite date, ci și la nivelul Data Lake, ca structură centrală, care colectează date din surse operaționale multiple, reprezentând el însuși sursa de date a depozitului care oferă suport pentru analiza activității de colectare a deșeurilor în cadrul programului PPCA. Drept urmare, toate erorile de reprezentare a datelor în cadrul Data Lake se vor propaga la nivelul depozitului de date, cu evidente consecințe negative în planul analizei și raportării și implicit, în plan decizional.

Figura 2 redă logica generală a prelucrărilor care vizează datele importate de la UAT și permit actualizarea înregistrărilor disponibile în tabelele din cadrul Data Lake. Enunțurile de mai jos completează reprezentarea grafică printr-o serie de precizări cu privire la maniera specifică în care sunt procesate importurile:

- Prelucrarea datelor importate este declanșată prin intermediul SQL Server Agent; într-o astfel de abordare, procedura stocată care implementează fluxul captat de diagramă este programată să se execute în mod automat, la intervale prestabilite.

- SQL Server Agent este un serviciu Windows, componentă a SQL Server care are rolul de a executa sarcini administrative programate în prealabil. Sarcinile (jobs) pot fi simple scripturi T-SQL, proceduri stocate, pachete SSIS sau baze de date SSAS. Acest mecanism oferă flexibilitatea executării sarcinilor pe baza unor reguli cu o minimă intervenție umană.
- În soluția implementată, agentul rulează la fiecare 30 de minute având drept sarcină prelucrarea datelor încărcate de către aplicația operațională în Data Lake. La fiecare rulare agentul preia datele încărcate în Data Lake din baza de date operațională în format JSON, realizează validarea JSON-ului și a datelor conținute după care fie inserează aceste date în baza de date relațională fie marchează JSON-ul recepționat ca fiind eronat, inserând într-un log mesaje de eroare care să fie utilizate de către aplicația care încarcă datele pentru verificări și corectarea erorilor semnalate. Job-ul (sarcina) marchează JSON-urile deja procesate astfel încât la următoarea rulare să fie prelucrate doar datele încărcate între timp. În situația în care sarcina anterioară nu a fost finalizată din cauza timpului mare de execuție, jobul nu prelucrează prelucrarea șirurilor JSON, urmând ca acestea să fie prelucrate la următoarea rulare a Job-ului.
- Prelucrarea vizează toate importurile care nu au fost încă procesate, indiferent de proveniența datelor (UAT-sursă).
- Importurile sunt prelucrate secvențial, în ordinea în care au fost memorate în Data Lake (tabelul Import), orice eroare înregistrată la procesarea unui import anterior realizat de aceeași UAT (indiferent de cauza sau etapa în care survine eroarea – validare sau operațiile propriu-zise de adăugare / editare / ștergere) conducând la generarea unei erori și în cazul importului ce face obiectul prelucrării curente. Desigur, erorile cauzate de procesarea unui anumit import vor avea consecințe similare asupra următoarelor importuri realizate de aceeași UAT, prelucrarea acestora nemaifiind posibilă. Pentru reluarea prelucrării importurilor de la aceeași sursă, este necesar ca respectiva UAT să realizeze ștergerea logică a tuturor importurilor care au generat erori, importurile ulterioare urmând a fi procesate în mod obișnuit.

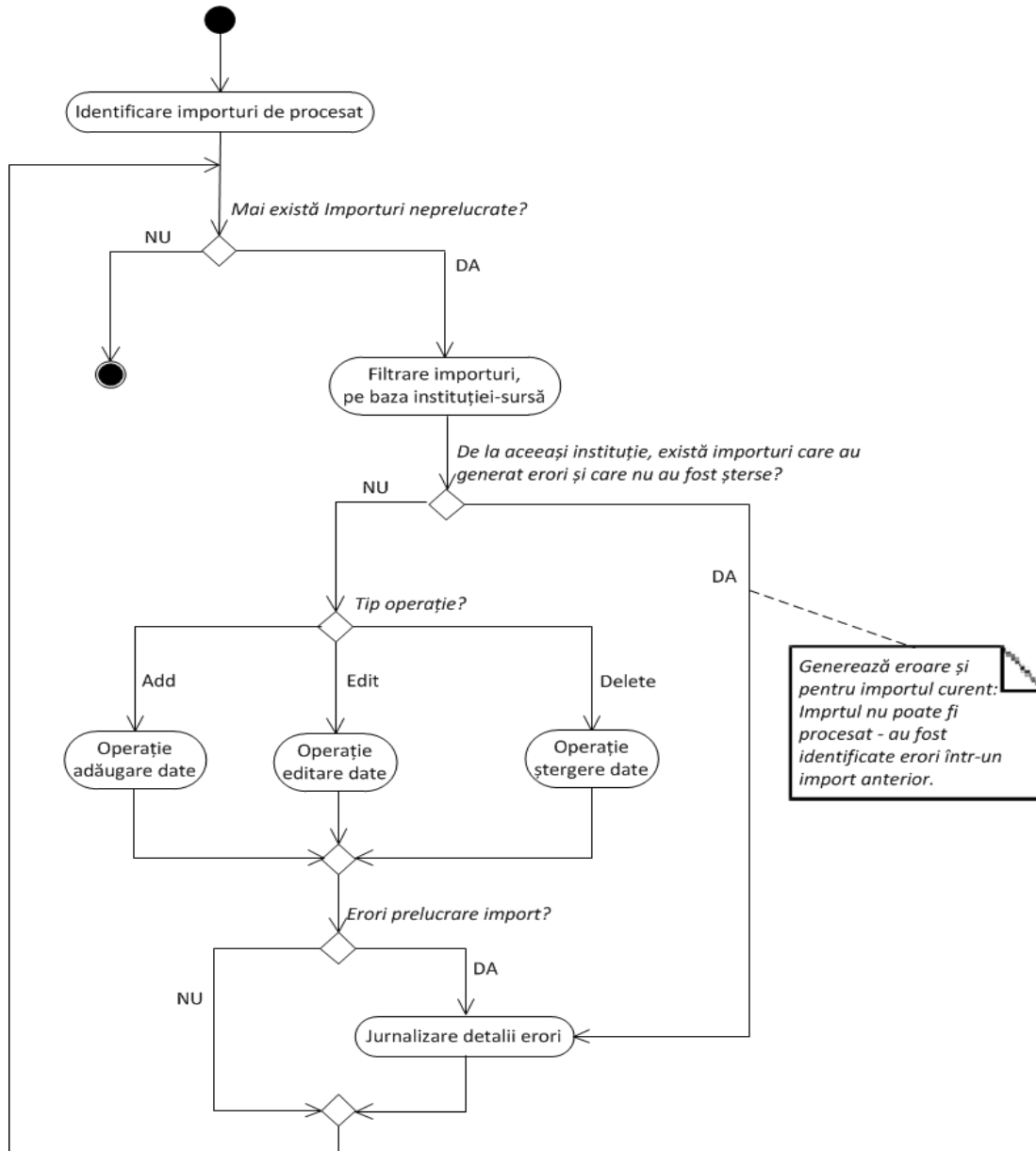


Figura 2 Procesarea importurilor de date: fluxul general al prelucrărilor

- Operațiile de adăugare, editare sau ștergere includ etapa validării șirurilor JSON furnizate de UAT, urmată, în cazul șirurilor JSON valide, de executarea tranzacțiilor specifice fiecărei operații. Fiecare din cele două etape pot genera erori care trebuie jurnalizate în tabelul prevăzut în acest scop în cadrul Data Lake (ValidationErrors). Jurnalizarea presupune înregistrarea unor detaliilor relevante pentru fiecare tip de eroare, aflat în strânsă

legătură cu tipul operației. Prin urmare, tipologia erorilor va fi abordată în contextul prezentării prelucrărilor specifice fiecărui tip de operație.

- Validarea vizează șiruri care conțin datele transmise de UAT pentru a fi memorate în Data Lake și care au fost în prealabil înregistrate în tabelul dedicat (JSON_import). În condițiile în care sunt memorate doar șiruri care se conformează sintaxei JSON, validarea se raportează la un format JSON prestabilit pentru indicarea tabelelor și câmpurilor-destinație, urmărind să verifice dacă datele importate sunt compatibile cu structura tabelelor și restricțiile privind valorile admise de câmpurile din Data Lake.
- În esență, natura prelucrărilor realizate pentru fiecare import depinde de scopul aceluși transfer de date, respectiv de tipul operației vizate (adăugare, modificare sau ștergere). Pe de altă parte, prelucrările specifice operației de adăugare și celei de editare prezintă similitudini care justifică reprezentarea acestora la nivelul aceleiași diagrame care captează maniera de procesare a importurilor; fluxul prelucrărilor specifice operației de ștergere va fi redat printr-o reprezentare separată.

Adăugarea și editarea înregistrărilor în tabelele din cadrul Data Lake

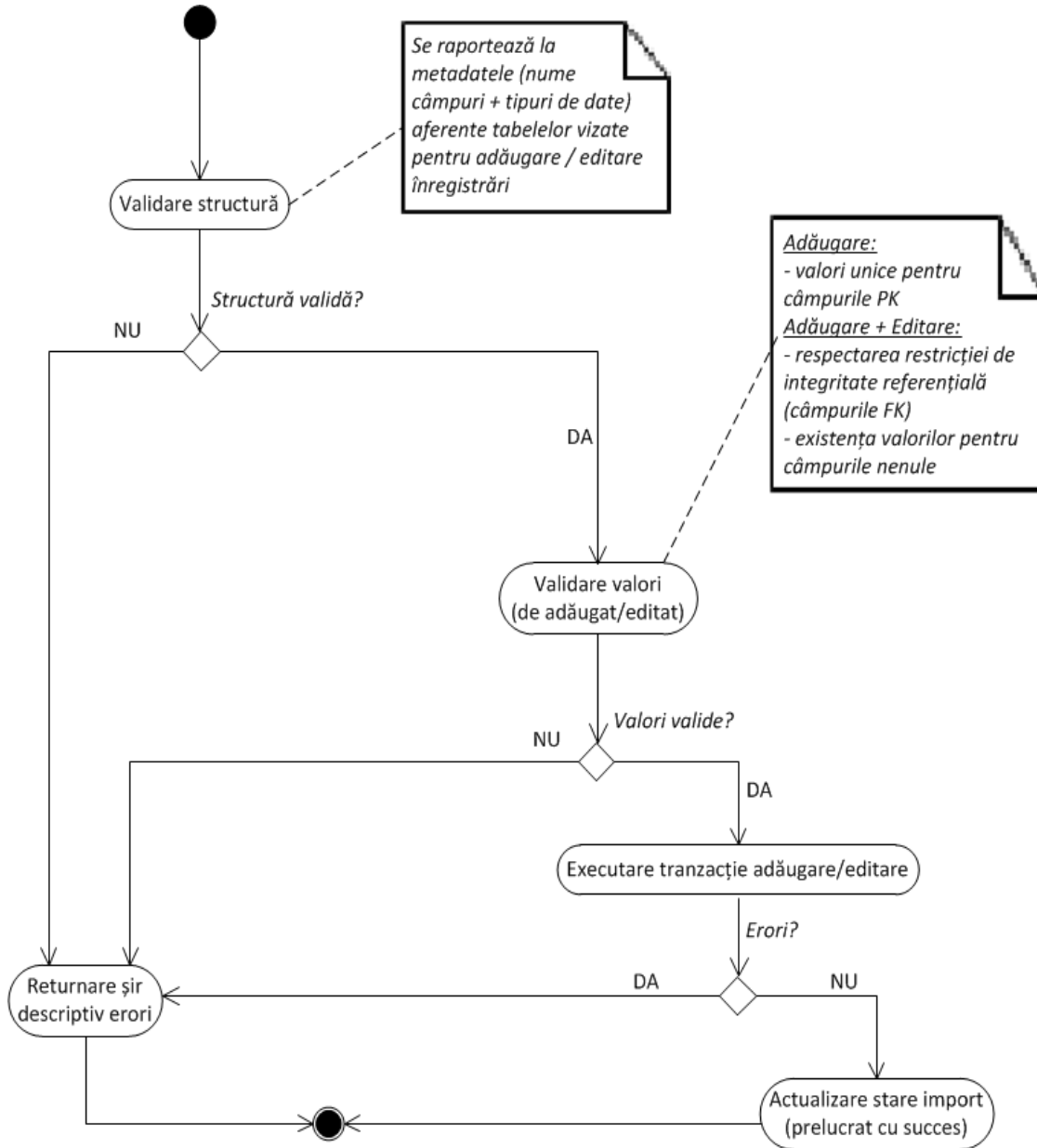


Figura 3 Procesarea importurilor de date: prelucrări specifice operațiilor de adăugare și de editare

Enunțurile de mai jos descriu maniera specifică în care sunt abordate operațiile de adăugare și de editare de date, la nivelul Data Lake:

- Pentru ambele tipuri de operații (adăugare sau editare), validarea șirului JSON aferent unui import se realizează în două faze succesive: validarea structurală, urmată de validarea valorilor care trebuie memorate în Data Lake.
- În ipoteza identificării unor erori de structură în șirul JSON care face obiectul prelucrării, nu se mai realizează validarea valorilor care trebuie memorate în Data Lake; acesta ar fi un demers inutil și chiar imposibil, dacă în șirul JSON nu se pot identifica subșiruri care să corespundă unor înregistrări compatibile cu structura tabelor-destinație.
- Orice tip de eroare (sub aspectul structurii sau valorilor) aferentă șirului JSON procesat cauzează stoparea prelucrării șirului, fără adăugarea sau editarea datelor. Totuși, validarea nu este întreruptă de apariția primei erori, verificarea completă a datelor conducând la identificarea și documentarea tuturor tipurilor de erori, astfel încât să se poată evita repetarea lor pentru importurile viitoare de la aceeași UAT.
- Atât operația de adăugare, cât și cea de editare, vizează șiruri JSON valide și sunt gestionate cu ajutorul tranzacțiilor. În cazul adăugării, șirurile JSON furnizează valorile aferente unor înregistrări noi, care trebuie inserate în unul sau mai multe tabele din Data Lake; în cazul editării, șirurile JSON conțin date menite să înlocuiască înregistrări preexistente din unul sau mai multe tabele din Data Lake. În aceste condiții, abordarea ambelor operații prin intermediul tranzacțiilor garantează că modificările vor fi operate cu succes, la nivelul tuturor tabelor-destinație; în caz contrar, erorile vor fi semnalate în mod corespunzător, distinct de cele care pot rezulta în urma validării.

Validări necesare pentru operațiile de adăugare și de editare: validarea structurii datelor

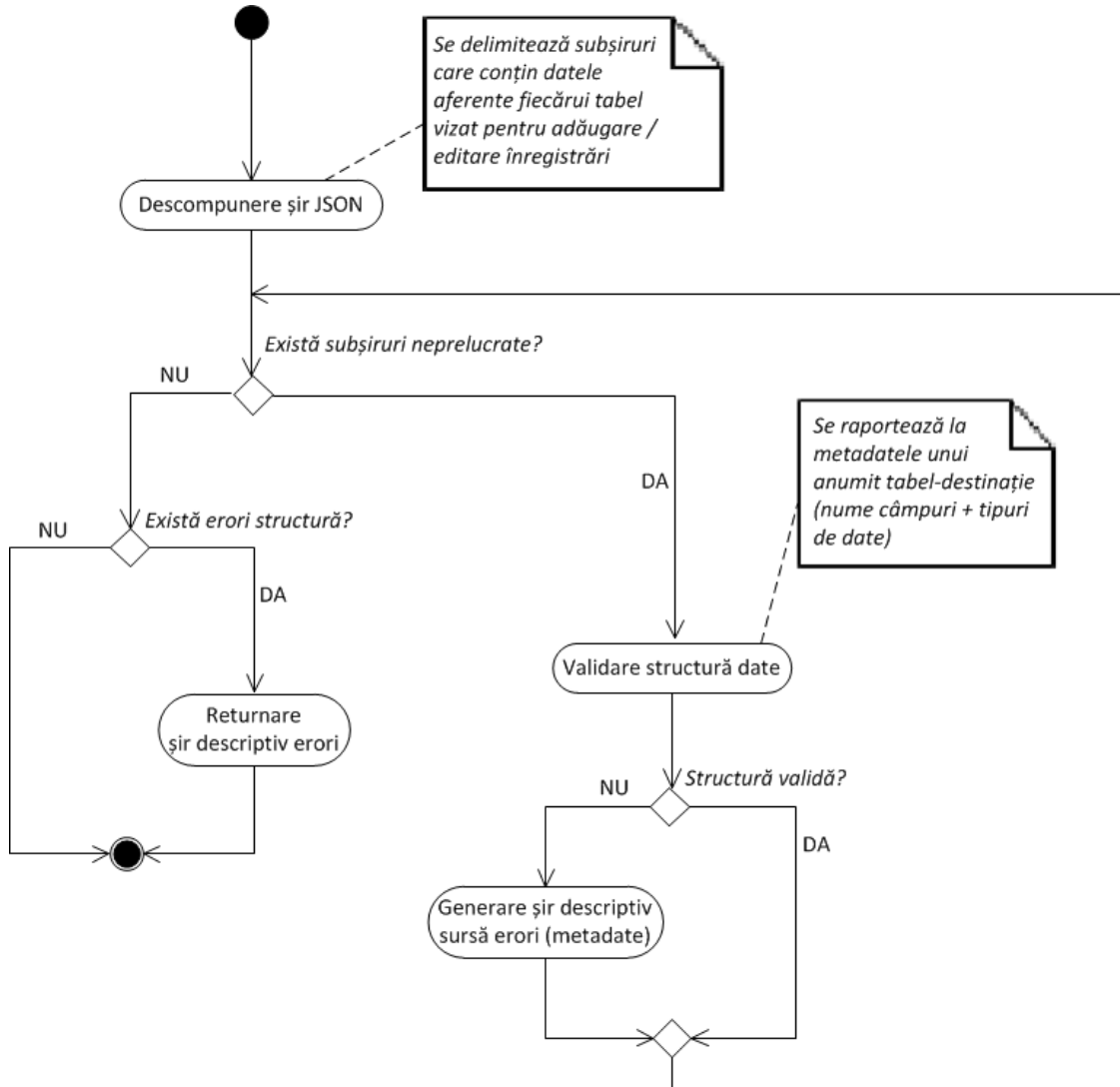


Figura 4 Procesarea importurilor de date: validarea structurii datelor, în cazul operațiilor de adăugare și de editare

Paragrafele următoare sintetizează principalele aspecte care definesc validarea datelor sub aspectul structurii, în cazul operațiilor de adăugare și de editare:

- Validarea structurii șirurilor JSON aferente importurilor se raportează la metadatele aferente tabelelor care reprezintă destinația datelor. Fiecare șir JSON trebuie indice numele unor tabele valide și numele unor câmpuri-destinație de asemenea valide,

precum și valorile care le vizează, cu respectarea formatului prestabilit. În plus, valorile trebuie să fie compatibile cu tipurile de date ale câmpurilor-destinație.

- Erorile de structură nu permit trecerea la următoarea etapă necesară prelucrării șirului JSON (validarea valorilor ce trebuie memorate în Data Lake) și, implicit, nici realizarea operației pentru care s-a realizat de fapt transferul datelor (adăugare sau editare).
- Detaliile erorilor de structură sunt înregistrate în tabelul ValidationErrors din cadrul Data Lake, indicându-se tabele și câmpurile la care se referă datele invalide din șirul JSON prelucrat.
- Validarea structurii datelor importate se aplică distinct pentru seturile de înregistrări care corespund fiecărui tabel-destinație, delimitate prin prelucrarea șirului JSON utilizat drept sursă a datelor. Ca urmare, șirul care indică eventualele erori de structură este generat dinamic, fiind actualizat în urma validării realizate în raport cu fiecare tabel-destinație; se obține astfel o descriere completă a tuturor surselor de eroare identificate în șirul JSON prelucrat.



Validări necesare pentru operațiile de adăugare și de editare: validarea valorilor care trebuie memorate

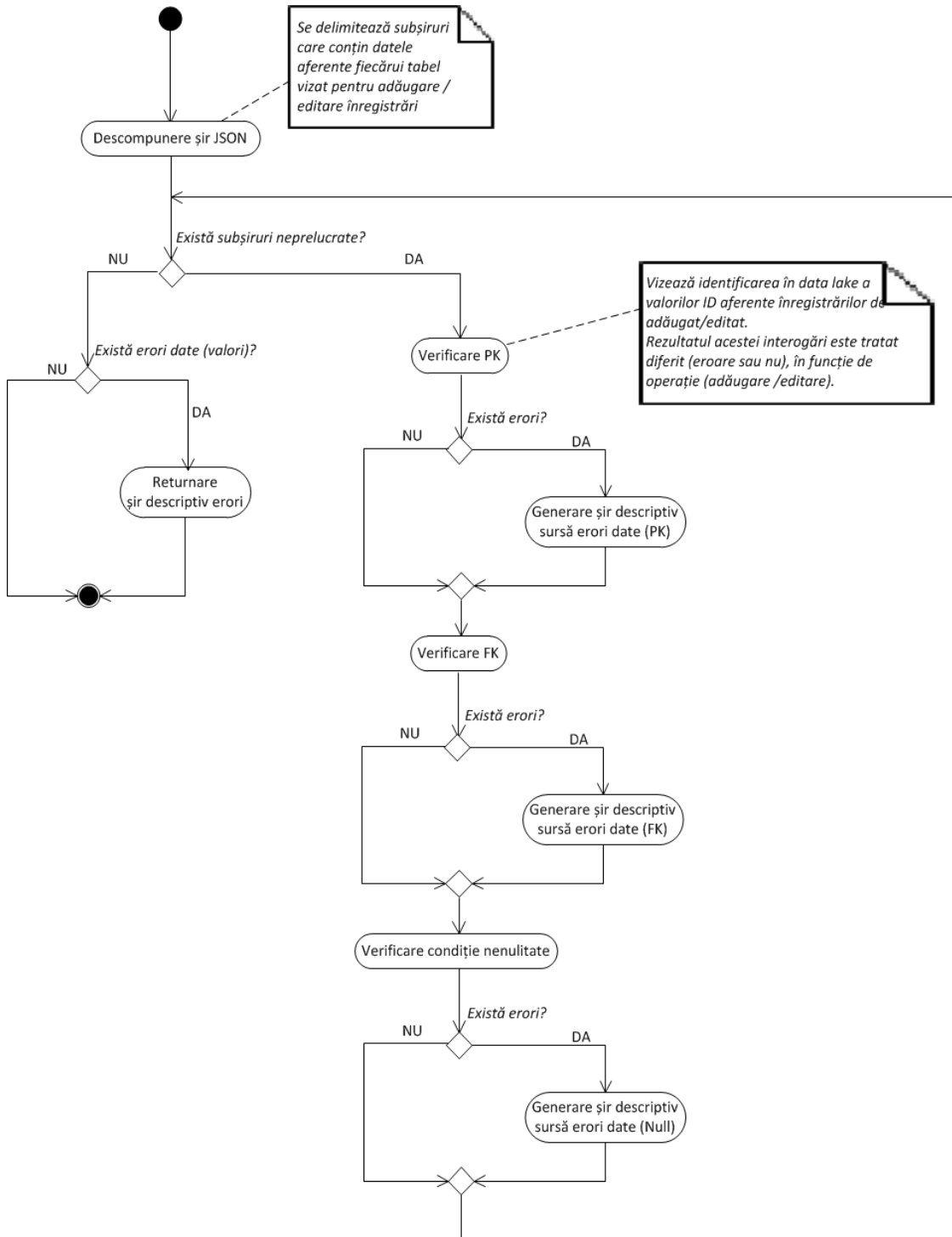


Figura 5 Procesarea importurilor de date: validarea valorilor, în cazul operațiilor de adăugare și de editare

Enunțurile de mai jos punctează elementele definitorii pentru validarea datelor importate în Data Lake, sub aspectul valorilor care fac obiectul adăugării sau editării:

- Validarea valorilor aferente înregistrărilor care trebuie adăugate sau editate la nivelul Data Lake tratează 3 aspecte:
 - Verificarea valorilor aferente cheilor externe – Vizează restricția de integritate referențială la nivelul înregistrărilor care provin de la aceeași UAT, chiar și în condițiile în care această restricție nu are o reprezentare explicită în Data Lake (din cauza recodificării uniforme a tuturor înregistrărilor, indiferent de UAT-sursă, la momentul salvării lor în tabelele din cadrul Data Lake). Vor genera erori toate înregistrările pentru care valorile aferente câmpurilor chei externe (FK) nu se regăsesc printre valorile cheilor primare asociate (PK), această corespondență fiind verificată prin raportarea la cheile și structura specifică bazelor de date operaționale aflate în exploatarea UAT.
 - Verificarea valorilor cu caracter obligatoriu – Vizează câmpurile care nu admit valori nule, identificabile pe baza definiției (metadatelor) tabelor de care aparțin; vor genera erori toate înregistrările care nu furnizează valori pentru aceste câmpuri, indiferent de natura operației care impune validarea (adăugare sau editare).
 - Verificarea valorilor aferente cheilor primare – Vizează restricția de unicitate a valorilor câmpurilor chei primare, pentru aceeași UAT, chiar și în condițiile în care aceste chei nu sunt reprezentate ca atare în Data Lake (care dispune de proprii identificatori la nivelul tabelor populate cu datele operaționale transmise de UAT). Și pentru adăugare, și pentru editare, se urmărește identificarea în Data Lake a valorilor ID aferente datelor importate, însă rezultatul acestei interogări este tratat diferit (eroare sau nu), în funcție de natura operației. În cazul adăugării de înregistrări, se urmărește ca valorile identificatorilor să nu se regăsească în înregistrările deja disponibile în tabelele din Data Lake, care provin de la aceeași UAT. În cazul operației de editare, este important ca valorile cheilor primare să existe deja în setul înregistrărilor importate anterior de la aceeași UAT, tocmai pe baza acestui ID fiind posibilă identificarea înregistrărilor de actualizat (prin înlocuirea valorilor memorate în Data Lake cu cele specificate de șirul JSON care face obiectul prelucrării). Având în vedere diferențele menționate, rezultă că

același șir JSON și același conținut al Data Lake vor conduce sau nu la o eroare de validare, în funcție de natura operației (adăugare sau editare).

- Erorile ce pot surveni la un moment dat nu cauzează întreruperea validării, scopul fiind prelucrarea completă a șirului JSON, pentru identificarea și documentarea tuturor tipurilor de erori, astfel încât să se prevină repetarea lor în cazul unor noi importuri de la aceeași UAT.
- Validarea valorilor destinate inserării sau editării se aplică distinct pentru seturile de înregistrări care corespund fiecărui tabel din Data Lake, delimitate prin prelucrarea șirului JSON utilizat ca sursă a datelor; drept consecință, șirul descriptiv care indică eventualele erori este generat în mod dinamic, fiind actualizat în urma validării realizate în raport cu fiecare tabel-destinație, pe baza criteriilor menționate (unicitatea valorilor cheii primare, integritatea referențială, condiția de nenulitate).

Ștergerea înregistrărilor din tabelele de la nivelul Data Lake

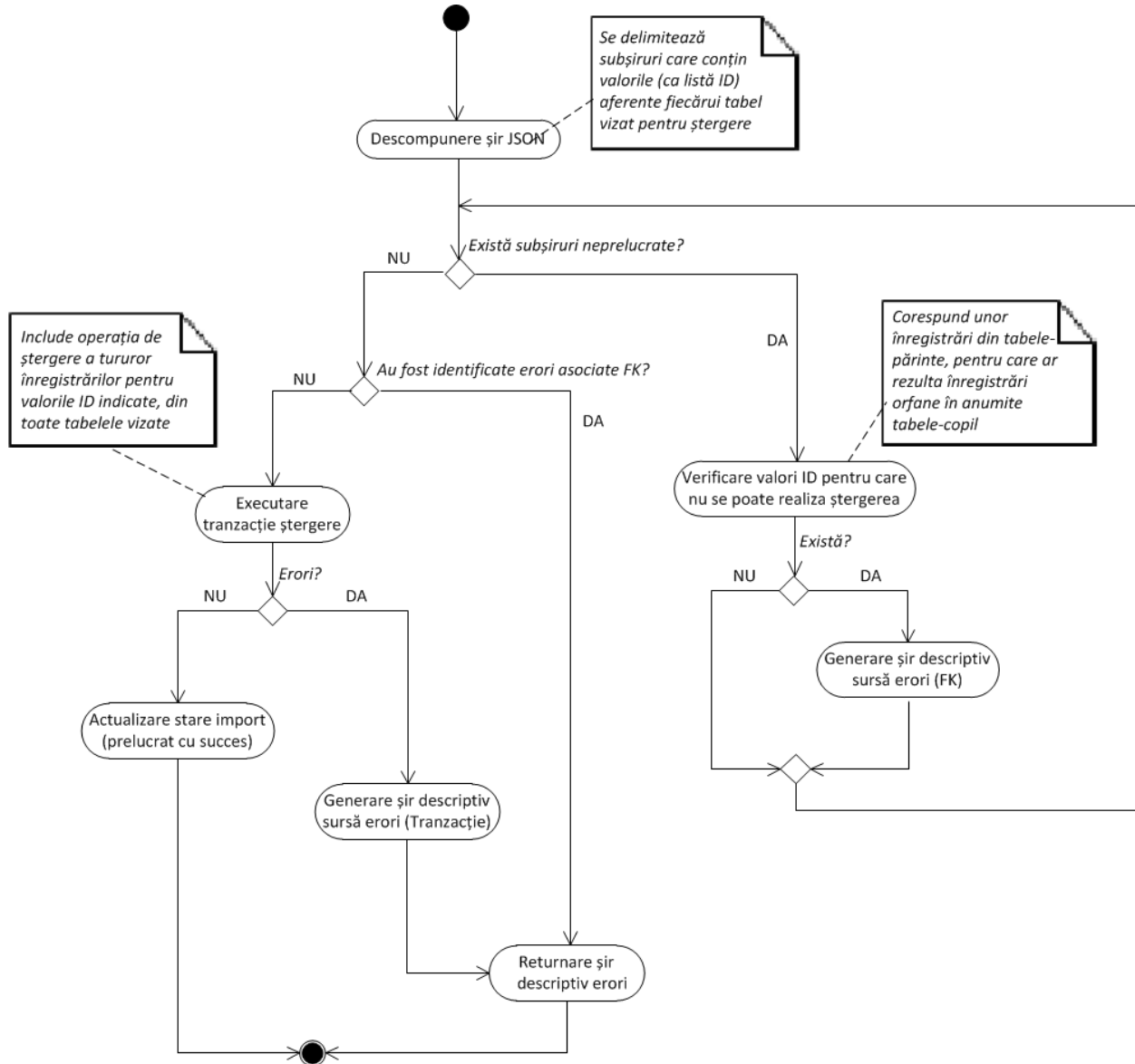


Figura 6 Procesarea importurilor de date: prelucrări specifice operației de ștergere

Paragrafele de mai jos expun maniera în care sunt tratate importurile asociate operațiilor de ștergere a datelor de la nivelul Data Lake:

- Operația de ștergere presupune ca șirul JSON importat să indice valorile cheilor primare (valori de tip ID, specifice UAT-sursă) care corespund înregistrărilor ce trebuie șterse din diferite tabele din cadrul Data Lake și care provin de la aceeași UAT.
- Pentru fiecare dintre aceste tabele, este necesar ca operația de ștergere să nu cauzeze apariția unor înregistrări orfane la nivelul UAT vizate. Drept urmare, anterior ștergerii propriu-zise trebuie să se verifice dacă este complet setul valorilor ID specificate în șirul JSON pentru identificarea înregistrărilor ce trebuie eliminate, cu alte cuvinte, dacă pentru ștergerea unei înregistrări-părinte se urmărește ștergerea tuturor înregistrărilor corelate, care conțin referințe la aceasta (toate înregistrările-copil); cazul contrar corespunde unei erori de ștergere care trebuie descrisă în mod corespunzător (înregistrările care nu pot fi șterse sunt indicate prin intermediul valorilor ID care le-au fost alocate la nivelul UAT-sursă).
- Verificarea setului valorilor ID pentru care nu este posibilă ștergerea se realizează pentru fiecare tabel indicat în șirul JSON prelucrat; drept consecință, șirul descriptiv asociat eventualelor erori este generat dinamic, fiind actualizat în urma verificării dependențelor specifice fiecărui tabel.
- Erorile asociate înregistrărilor ce nu pot fi eliminate din cauza dependențelor mai sus menționate determină stoparea prelucrării șirului JSON, fără a se realiza ștergerea propriu-zisă. Pe de altă parte, pentru evitarea acestor erori în cazul unor importuri viitoare de la aceeași UAT, se urmărește identificarea întregului set de valori ID pentru care nu se poate realiza ștergerea; drept urmare, detectarea unei prime valori de acest tip nu cauzează oprirea verificării pentru restul valorilor ID, această operație vizând setul complet al datelor din șirul JSON.
- În absența erorilor ce vizează înregistrările orfane, operația de ștergere este gestionată cu ajutorul tranzacțiilor. În condițiile în care șirurile JSON indică înregistrări ce trebuie eliminate din unul sau mai multe tabele din cadrul Data Lake, abordarea tranzacțională a operației de ștergere garantează că toate modificările vor fi operate cu succes, la nivelul tuturor tabelelor; în caz contrar, erorile urmează fi semnalate în mod corespunzător, distinct de cele care pot rezulta în urma verificării valorilor ID specificate în șirul JSON.

Concluzii privind prelucrarea datelor transmise de UAT

În condițiile în care operațiile ce se pot realiza asupra înregistrărilor din Data Lake sunt condiționate de validarea prealabilă a șirurilor JSON care au fost transmise de UAT și care conțin datele aferente acestor prelucrări, dar totodată și de executarea cu succes a tranzacțiilor necesare pentru menținerea coerenței la nivelul Data Lake, în Figura 7 și Figura 8 este prezentată o sinteză a tipologiei erorilor care pot surveni cu ocazia procesării șirurilor JSON, dar și a manierei în care sunt jurnalizate în tabelul prevăzut în acest scop în Data Lake. Expunerea detaliată a contextului specific fiecărui tip de operație (adăugare, modificare, ștergere) și a aspectelor care vizează identificarea erorilor este disponibilă în secțiunile anterioare.

Stare șir JSON	Cauze erori	Tip operație și erori asociate	
		Adăugare / Editare	Ștergere
Șir neprelucrat	Importurile anterioare de la aceeași UAT	Erori generate de importurile anterioare, care nu au fost șterse (logic)	
Șir prelucrat	Prelucrările premergătoare operației vizate	- Erori validare structură date	Erori aferente înregistrărilor ce nu pot fi eliminate
		- Erori validare valori	
	- Erori execuție prelucrări		
Prelucrările circumscrise operației vizate	- Erori din tranzacții		
	- Erori execuție (non-tranzacționale)		

Figura 7 Tipologia erorilor asociate importurilor de date

Tip operație		Tipologie erori procesare importuri	Etichetare specifică în data lake	Detalii înregistrate în data lake (exemple ipotetice)
Adăugare Editare	Ștergere			
X	X	Erori generate de importurile anterioare, care nu au fost șterse (logic)	Batch	Importul nu poate fi procesat - au fost identificate erori într-un import anterior.
X		Erori validare structură date	Data	Șablon descriptiv: <i>nume_tabel(...listă câmpuri cu erori...)</i> bin_collections(tag_code, tag_type) cars(type, volme)
X		Erori validare valori	Schema	Șabloane descriptive: <i>nume_tabel(#PK: ...listă valori invalide pentru cheia primară...)</i> <i>nume_tabel(#FK: ...listă câmpuri chei externe cu valori invalide...)</i> <i>nume_tabel(#Null: ...listă câmpuri nenule, pt.care nu există valori...)</i> bin_collections(#PK:1,2,3) bin_collections(#FK:location_id(1) tag_id(1,2) car_id(1,2,3) bin_collections(#Null:tag_code, tag_type, tag_volume)
	X	Erori aferente înregistrărilor care nu pot fi eliminate	Data	Șablon descriptiv: <i>nume_tabel(...listă ID pt.înregistrări ce nu pot fi șterse...)</i> bin_collections(1,2,3) cars(1,2,3)
X	X	Erori din tranzacții	Transaction	Operația nu se poate realiza.
X	X	Erori execuție	Unknown	Importul nu poate fi procesat.

Figura 8 Jurnalizarea erorilor în cadrul Data Lake

Datele transmise de UAT în format JSON sunt memorate în Data Lake și prelucrate în funcție de natura operației vizate (adăugare, modificare, ștergere), rezultatele procesării fiecărui șir fiind disponibile în tabelul Import, sub forma unei înregistrări distincte. Diagrama din Figura 9 oferă o imagine de ansamblu asupra secvenței de stări asociate fiecărui import de date, tranziția între aceste stări fiind determinată de prelucrările pe care le suportă șirul JSON aferent importului, prezentate pe larg în secțiunile anterioare.

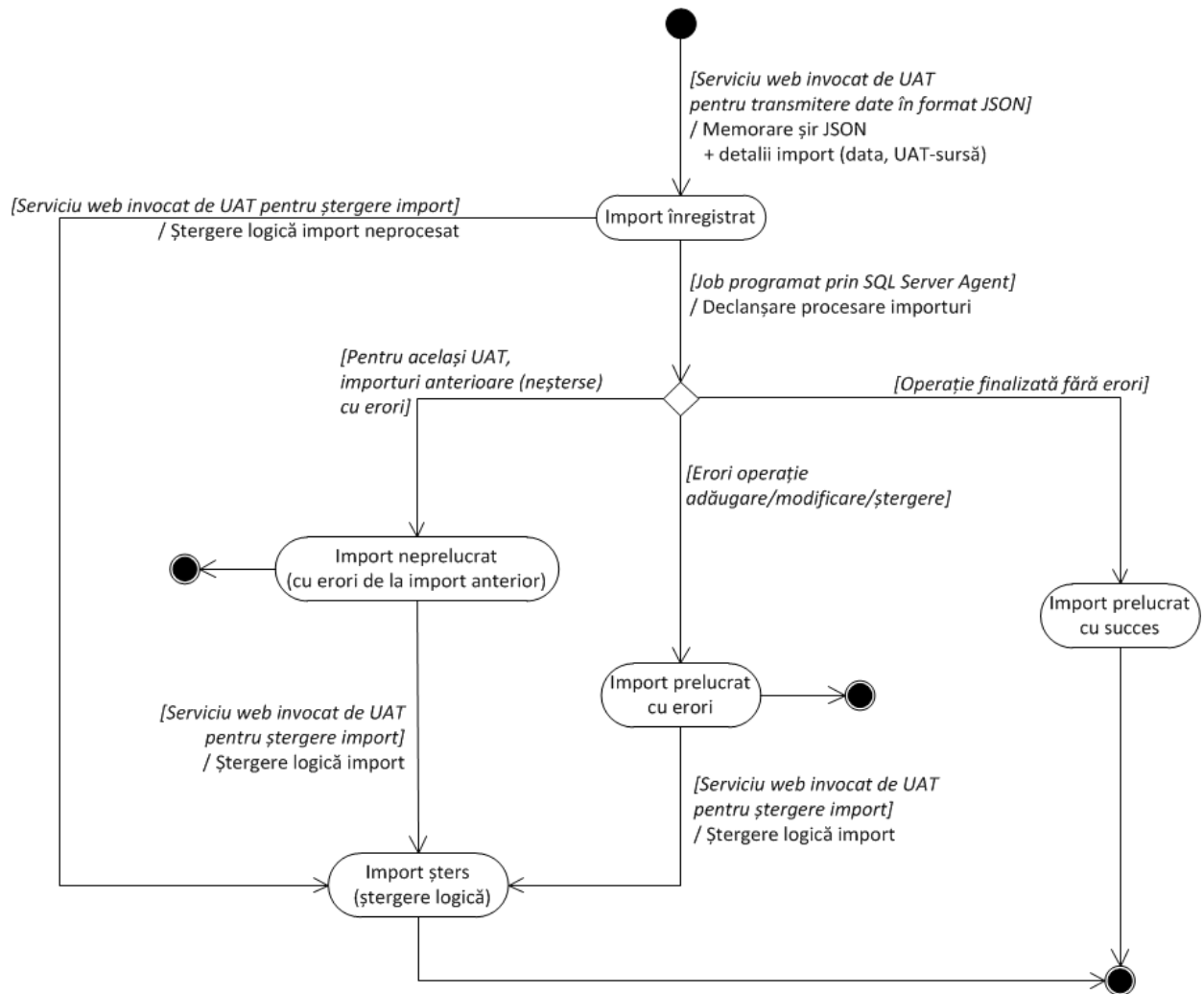


Figura 9 Diagrama stărilor aferente importurilor de date de la UAT

Diagrama include suita stărilor pe care le poate traversa orice import aferent unui șir JSON transmis de UAT. Starea curentă a fiecărui import de la nivelul Data Lake este disponibilă în tabelul Import, care conține detalii actualizabile în urma procesării secvențiale a importurilor, declanșată în mod automat prin intermediul serviciului SQL Server Agent. De asemenea, fiecare UAT poate consulta în orice moment starea curentă a importurilor care îi corespund, pe baza metodelor oferite de serviciul web. Din ansamblul stărilor posibile, doar cele care corespund unui import de date procesate cu succes sau unui import șters (logic) de UAT-sursă reprezintă stări finale, care nu mai suportă prelucrări suplimentare. Pe de altă parte, în cazul unui import cu erori (fie că sunt cauzate de prelucrarea sa directă, fie de faptul că este precedat de un alt import cu erori, care provine de la aceeași UAT și pentru care nu s-a cerut ștergerea logică),

procesarea se poate încheia cu simpla jurnalizare a erorilor, fără ca UAT-sursă să solicite ștergerea importului; desigur, o astfel de situație va influența toate importurile ulterioare de la aceeași UAT, prelucrarea lor nefiind posibilă până nu se operează ștergerea logică a importurilor cu erori.

3.3.1 Modelul de date necesare componentei software de transfer a datelor din sistemul operațional în depozitul de date

Baza de date care va deservi componenta de transfer a datelor va conține o structură relațională care să permită reținerea datelor privind salariile pentru o lună de raportare și pentru fiecare instituție în parte. În acest sens, s-au identificat două categorii de câmpuri necesare pentru diferențierea corectă a informațiilor:

- câmpurile preluate din baza de date operațională ce se vor regăsi în tabele similare ca denumire și structură cu cele din bazele de date operaționale. Acestea vor stoca datele preluate de la fiecare UAT;
- câmpuri necesare pentru evidențierea istoricului transferurilor efectuate de fiecare UAT în baza de date centralizatoare de tip DataLake. Pentru a identifica ulterior în mod distinct proprietarii datelor importate în baza de date centralizatoare, la nivelul fiecărui tabel care stochează date preluate de la UAT-uri, s-a utilizat un câmp care definește cheia externă în raport cu tabelul ce va reține informații generale despre fiecare import de date.

Nume câmp	Tip de dată	Dimensiune	Obligatoriu	Tabel	Descriere
ID	Număr întreg		Da	Import	Reține id-ul importului pentru care definește transferul de date
import_time	Datetime		Da	Import	Reține data și ora la care s-a realizat transferul obiectului de tip JSON
from_date	Date		Da	Import	Data începând de la care se realizează importul înregistrărilor din baza de

Nume câmp	Tip de dată	Dimensiune	Obligatoriu	Tabel	Descriere
					date operațională în cea centralizată
to_date	Datetime		Da	Import	Data până de la care se realizează importul înregistrărilor din baza de date operațională în cea centralizată
operation_type	Nvarchar	10	Nu	Import	Definește tipul operațiunii. Poate fi Add – în cazul în care noi înregistrări sunt raportate, Modify – pentru modificarea unor înregistrări existente, sau Delete – pentru ștergerea unor înregistrări preizate prin ID-urile acestora
ProcessingStartTime	Datetime		Nu	Import	Data și ora la care a început procesarea datelor din obiectul de tip JSON
ProcessingEndTime	Datetime		Nu	Import	Data și ora la care s-a finalizat procesarea datelor din obiectul de tip JSON atașat unui transfer

Nume câmp	Tip de dată	Dimensiune	Obligatoriu	Tabel	Descriere
Deleted	Bit		Nu	Import	Permite marcarea unui import realizat anterior ca fiind șters
DeletionDate	Datetime		Nu	Import	Reține data și ora la care s-a marcat importul ca fiind șters prin metoda de ștergere a serviciului web
InstitutionName	Nvarchar	100	Da	Institution	Reține numele unui UAT care poate transfera date
Username	Nvarchar	100	Nu	Institution	Permite stabilirea unui nume de utilizator pentru autentificarea instituției respective
Password	Nvarchar	100	Nu	Institution	Reține parola pentru autentificarea instituției respective (UAT)
Json_string	Nvarchar	Max	Da	JSON_import	Reține obiectul de tip JSON care deține datele transferate de un UAT în cadrul unui import de date
ShortName	Nvarchar	100	Nu	ValidationErrors	Permite un nume scurt pentru o eroare identificată la validarea datelor unui JSON

Nume câmp	Tip de dată	Dimensiune	Obligatoriu	Tabel	Descriere
Description	Nvarchar	Max	Da	ValidationErrors	Reține descrierea explicită a unei erori identificate la procesarea datelor unui transfer
ErrDateTime	Datetime		Da	ValidationErrors	Permite reținerea datei la care a fost înregistrată eroarea de procesare
ValidationType	Nvarchar	100	Da	ValidationErrors	Permite definirea unui tip de eroare pentru cea înregistrată, în vederea grupării erorilor raportate pe categorii

3.3.2 Etape principale în dezvoltarea sistemului informatic de analiză a datelor colectate prin programul PPCA

3.3.2.1 Etape în realizarea bazei de date centralizatoare necesare sistemului de analiză a datelor

După cum a fost reprezentat în Figura 6, sistemul informatic pentru analiza datelor colectate prin programul PPCA este alcătuit din mai multe componente software și sisteme integrate care comunică prin intermediul serviciilor web. La nivelul datelor s-a considerat necesară constituirea unui nivel intermediar de date între bazele de date operaționale și depozitul de date, cu rolul bine definit de a centraliza datele într-o bază de date de tip DataLake de unde depozitul de date se va alimenta prin procesul de tip ETL.

Întrucât este necesar ca datele transferate de către sistemul operațional al programului PPCA în baza de date centralizată să fie păstrate în aceleași structuri de date pentru a simplifica procesul de sincronizare fiecare din tabelele provenind din baza de date care deservește sistemul operațional (tranzacțional) a fost replicat în cadrul bazei de date de tip DataLake.

Totodată s-a evidențiat necesitatea de a reține metadatele care descriu transferurile efectuate. Aceasta a condus la o structură de tabele suplimentare care completează structura replicată a bazei de date operaționale în baza de date centralizată (o diagramă a modului de asociere între acestea poate fi observată în Figura 13). În cadrul fiecărei tabele replicată din baza de date operațională s-a adăugat câmpul ImportId care desemnează legătura cu tabelul Importuri în care se evidențiază datele privind importurile salvate în baza de date centralizată.

Deși baza de date operațională se supune rigorilor modelului relațional, cu toate constrângerile care țin de integritatea datelor, tabelele din baza de date centralizată și-au pierdut asocierile dintre ele. Pentru aceasta au fost necesare o serie de prelucrări suplimentare privind eliminarea cheilor primare și a cheilor externe din cadrul acestor tabele și definirea de chei primare surogat. Anexa 1 prezintă lista procedurilor stocate care au condus la acest rezultat.

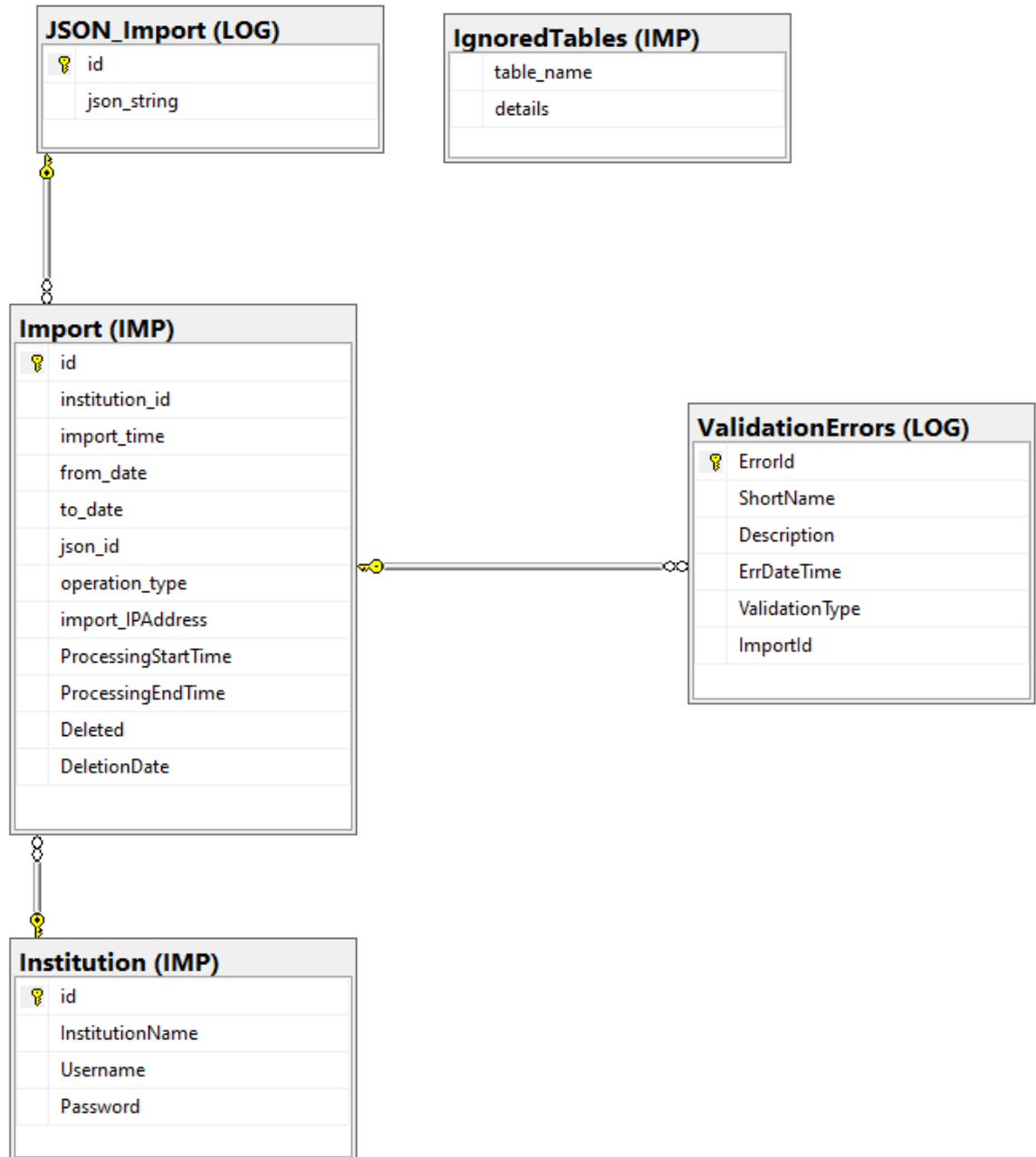


Figura 19. Diagrama bazei de date aferente componentei de transfer de date către baza de date operațională și procesare a acestora

3.3.3 Construirea serviciului web pentru transferul informațiilor în baza de date centralizatoare a sistemului de analiză a datelor prin programul PPCA

Pentru asigurarea comunicării în timp real între sistemele informatice operaționale (câte unul pentru fiecare UAT) și sistemul informatic de analiză a datelor, a fost construit un serviciu web care funcționează ca instrument de transfer de date și integrare între cele două soluții software (Figura 14). Acest serviciu va funcționa pe serverul web al beneficiarului și va fi securizat prin politicile pe care acesta le aplică tuturor site-urilor web publicate. Serviciul web a implementat securitatea transportului prin politici de tip SSL pe bază de certificat. Metodele dezvoltate în cadrul său au urmărit implementarea din punct de vedere informatic a proceselor de afaceri descrise anterior în cadrul prezentului raport.

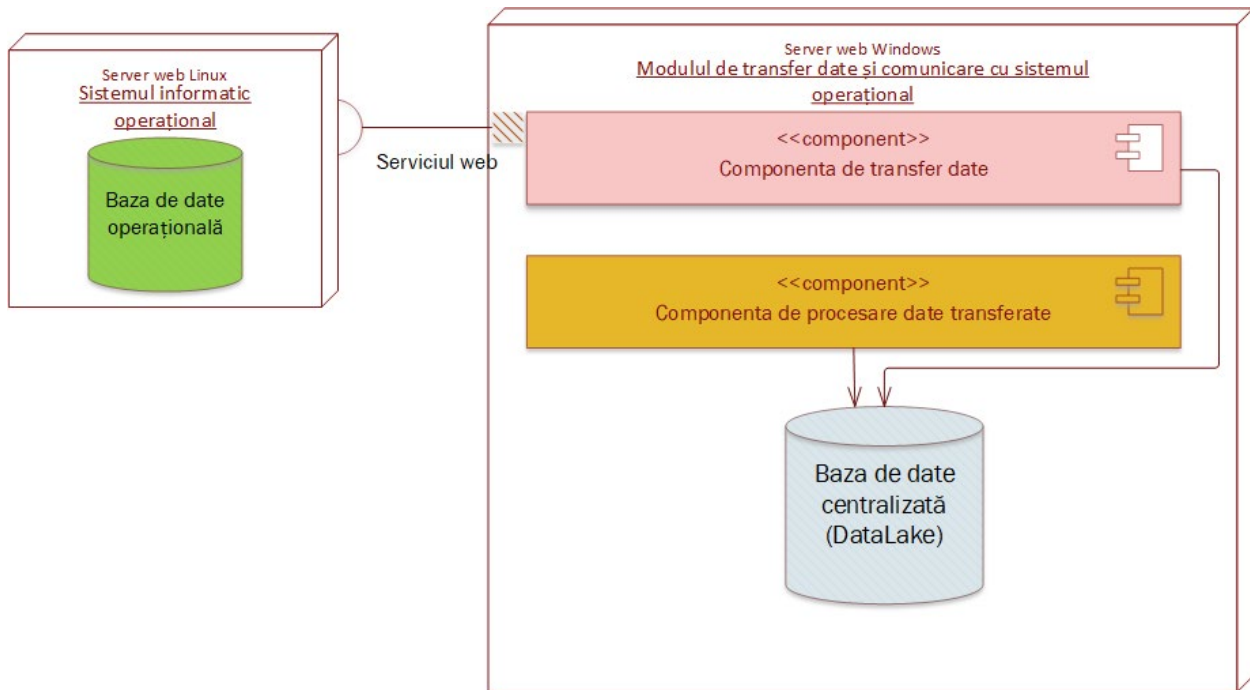


Figura 20. Diagrama de amplasare care prezintă comunicarea între soluțiile software prin intermediul serviciului web dezvoltat

Serviciul web dezvoltat în vederea asigurării integrării sistemului operațional cu sistemul de asistare a deciziei este disponibil, temporar, la adresa:

<https://sipocaweb.ase.ro/SIPOCAWebService/ImportService.svc>

O aplicație de test a fost, de asemenea, dezvoltată în vederea verificării funcționalității serviciului web și a metodelor aferente și este disponibilă online, temporar, la adresa:

<https://sipocaweb.ase.ro/AppTest/>

Serviciul web a fost dezvoltat utilizând o arhitectură software pe mai multe niveluri respectând principiul Separation of Concerns. Astfel, potrivit acestei arhitecturi s-au construit componentele de tip front-end și back-end care pot fi vizualizate în figura următoare:

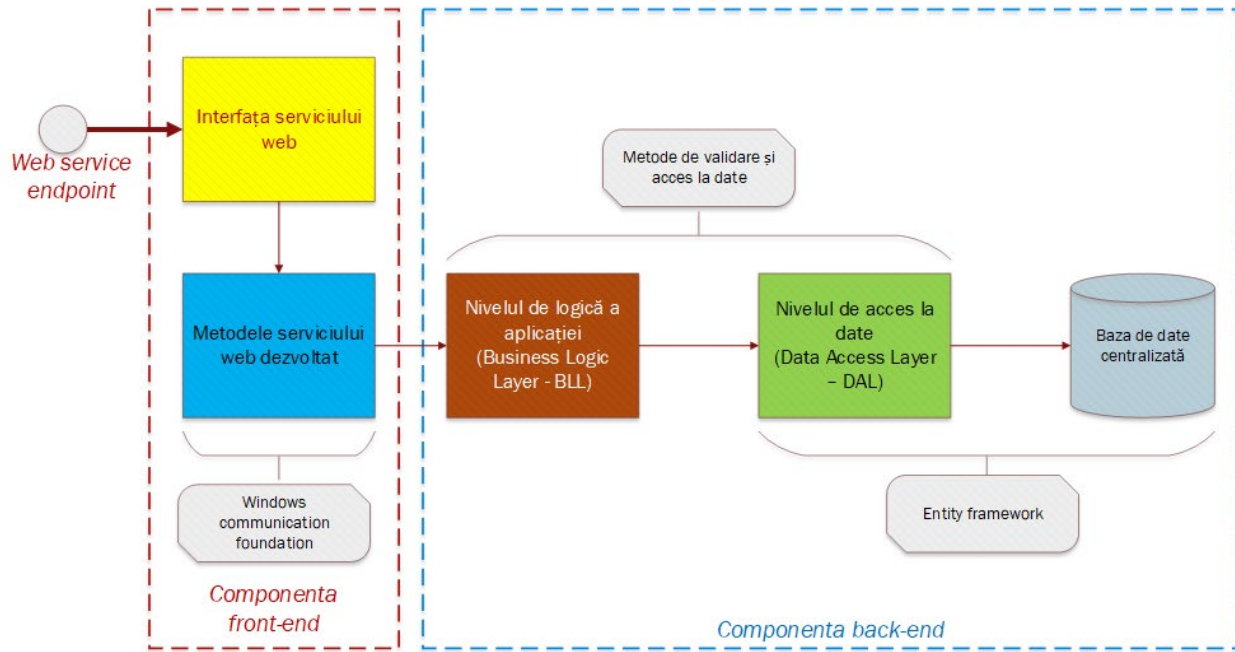


Figura 21. Nivelurile arhitecturale implementate pentru componenta de transfer date

Metodele expuse în cadrul serviciului web SipocaWebService/ImportService.svc sunt următoarele:

1. *GetLastTransmittedImport(long institutionId)*

Returnează un obiect cu proprietățile LastId, StartDate, EndDate. LastId reprezintă id-ul ultimului transfer al respectivei instituții. StartDate – ziua următoare datei de final până la care s-a realizat ultimul transfer. EndDate - ziua de ieri.

2. *GetLastTransmittedImportPerInstitution(string username, string password)*

La fel ca precedenta însă pornind de la nume de utilizator și parolă.

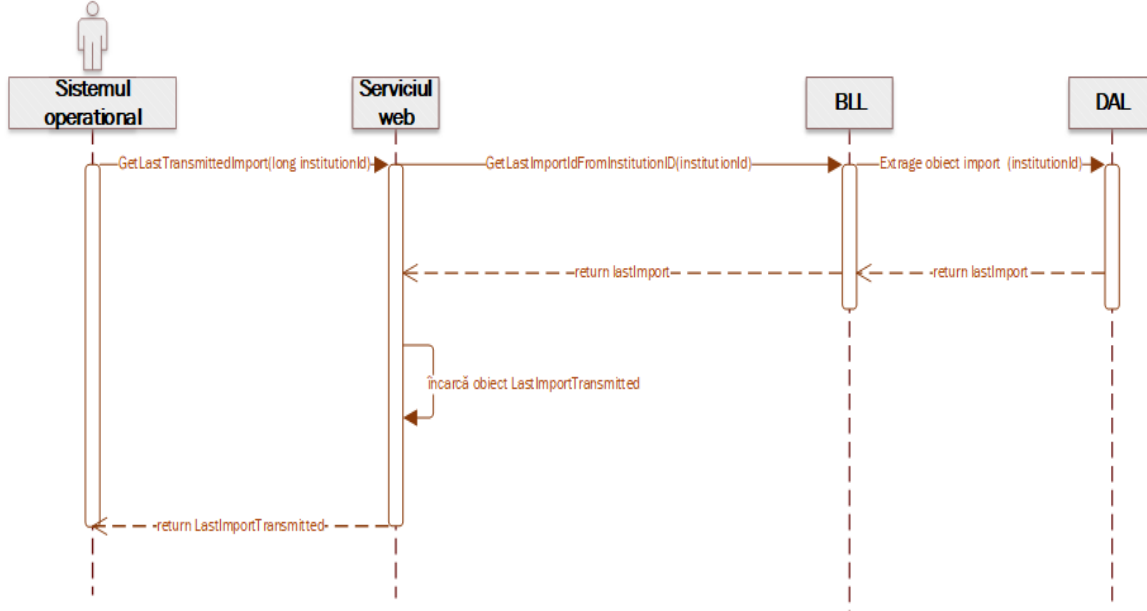


Figura 22. Diagrama de secvențe privind metoda web de extragere a ultimului import al unui UAT

3. *GetImportListByInstitutionID(string username, string password)*

Returnează un șir JSON cu toate înregistrările din baza de date raportate de către instituția cu acele date de conectare. Șirul JSON conține și o colecție de erori (dacă este cazul), care sunt identificate în etapa de procesare a acestuia. În cazul în care nu sunt erori de procesare, ProcessingStartDate și ProcessingEndDate vor fi completate cu data și ora la care a început, respectiv s-a finalizat procesarea șirului JSON și popularea tabelor din baza de date cu înregistrările raportate de acesta. Între momentul înregistrării raportării și momentul procesării șirului JSON, câmpurile ProcessingStartDate și ProcessingEndDate vor rămâne nule.

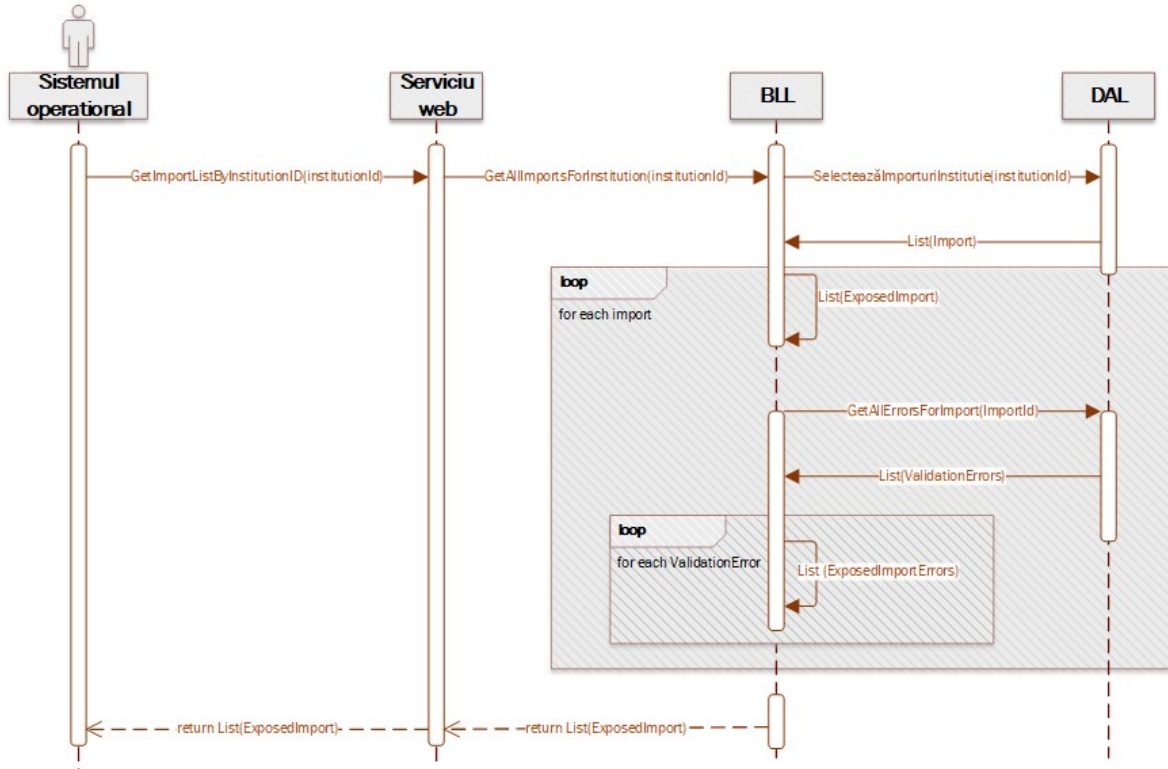


Figura 23. Diagrama de secvențe pentru metoda *GetAllImportsFromInstitution(institutionId)*

4. ***GetImportInfoById(long importId)***

Returnează un șir JSON cu înregistrarea din baza de date având id-ul importului precizat ca parametru. Șirul JSON este identic ca format cu cel aferent metodei *GetImportListByInstitutionID*, însă doar pentru un singur identificator de import.

5. ***GetTransferredJSONbyImportId(long importId)***

Afișează întregul șir JSON transferat în baza de date cu ID-ul de import precizat. ID-ul de import se poate vedea în rezultatul metodei *GetImportListByInstitutionID*.

6. ***TransferJSONforPeriod(string username, string password, string jsonString, string operation_Type, string startDate, string endDate)***

Salvează în baza de date JSON-ul transmis ca argument (*jsonString*).

Parametri:

- Username, Password – datele de conectare ale UAT-ului.
- Operation_Type – se va preciza, după caz, una din valorile: Add, Edit, Delete.

- startDate – data de început a intervalului raportat prin acest transfer (ziua următoare datei de final aferente ultimului transfer). Datele sunt de tip șiruri de caractere specificate în formatul **aaaallzz**.
- endDate – data de sfârșit a intervalului raportat – nu poate să depășească data afișată prin proprietatea EndDate din metodele *GetLastTransmittedImport* sau *GetLastTransmittedImportPerInsitution*. Datele sunt de tip șiruri de caractere specificate în formatul **aaaallzz**.

Dacă importul se realizează cu succes, va returna id-ul importului. Dacă sunt erori, va returna un obiect de tip JSON cu rezultatul transferului în care există lista de erori definită.

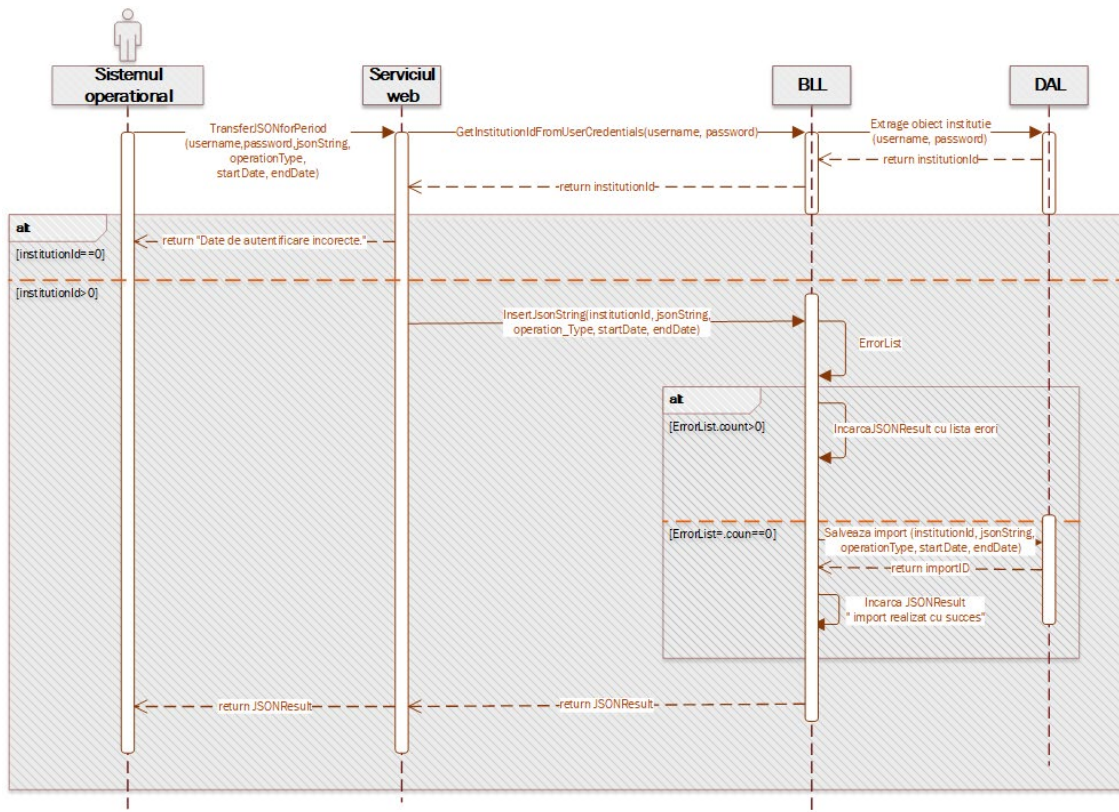


Figura 24. Diagrama de secvențe privind metoda web de transfer a unui set de date în baza de date centralizatoare

7. DeleteImport(long importId)

Șterge un import împreună cu șirul JSON transferat, doar dacă este ultimul transferat și doar dacă nu a fost deja procesat (Figura 19). De asemenea, metoda verifică dacă există importuri

de la aceeași instituție după cel care se dorește a fi șters și nu permite ștergerea dacă acestea există (vezi Figura 20)

Returnează un șir de caractere cu statusul operațiunii și, dacă este cazul, cu erorile identificate la încercarea de ștergere.

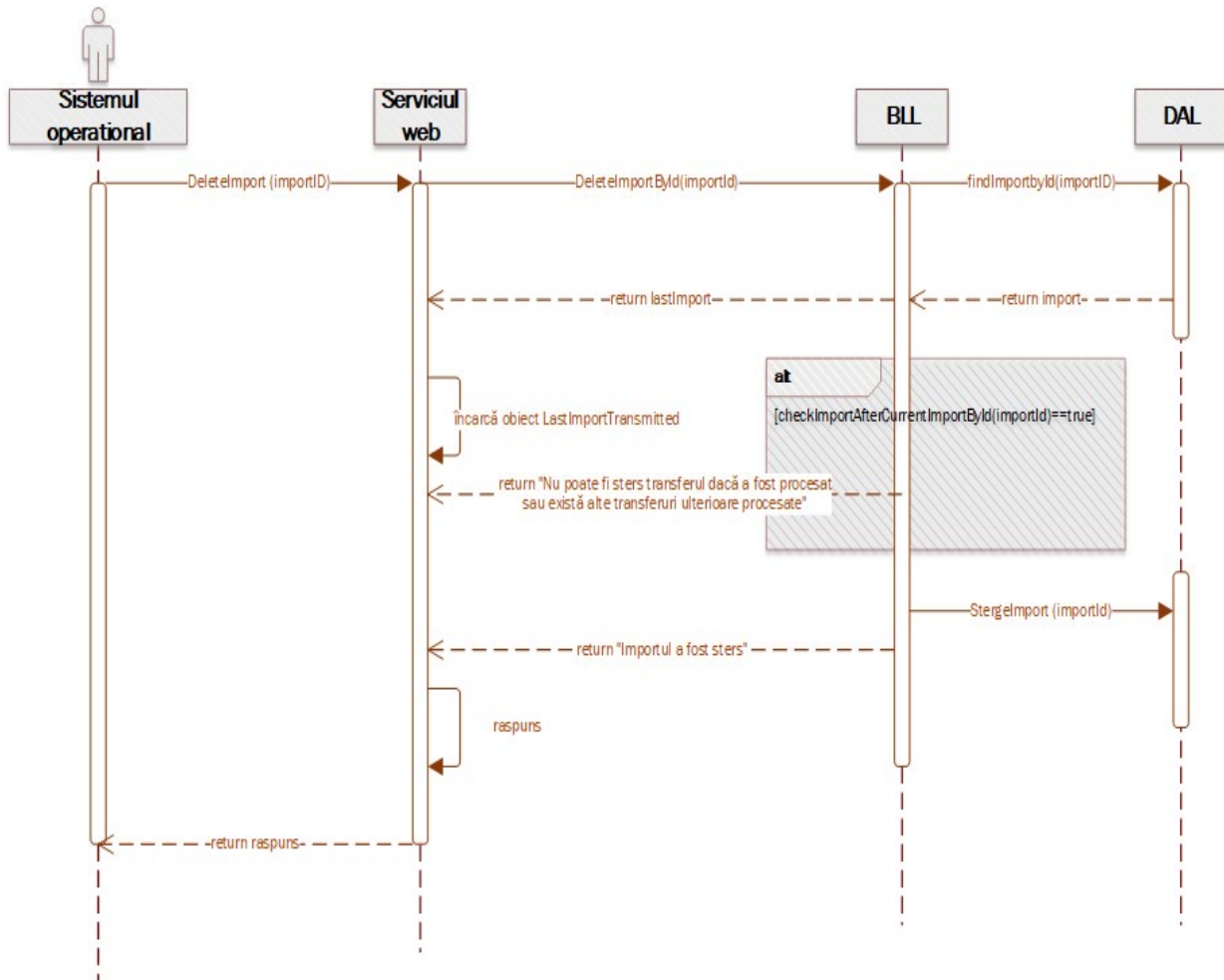


Figura 25. Diagrama de secvențe privind metoda de ștergere a unui import existent din partea unui UAT

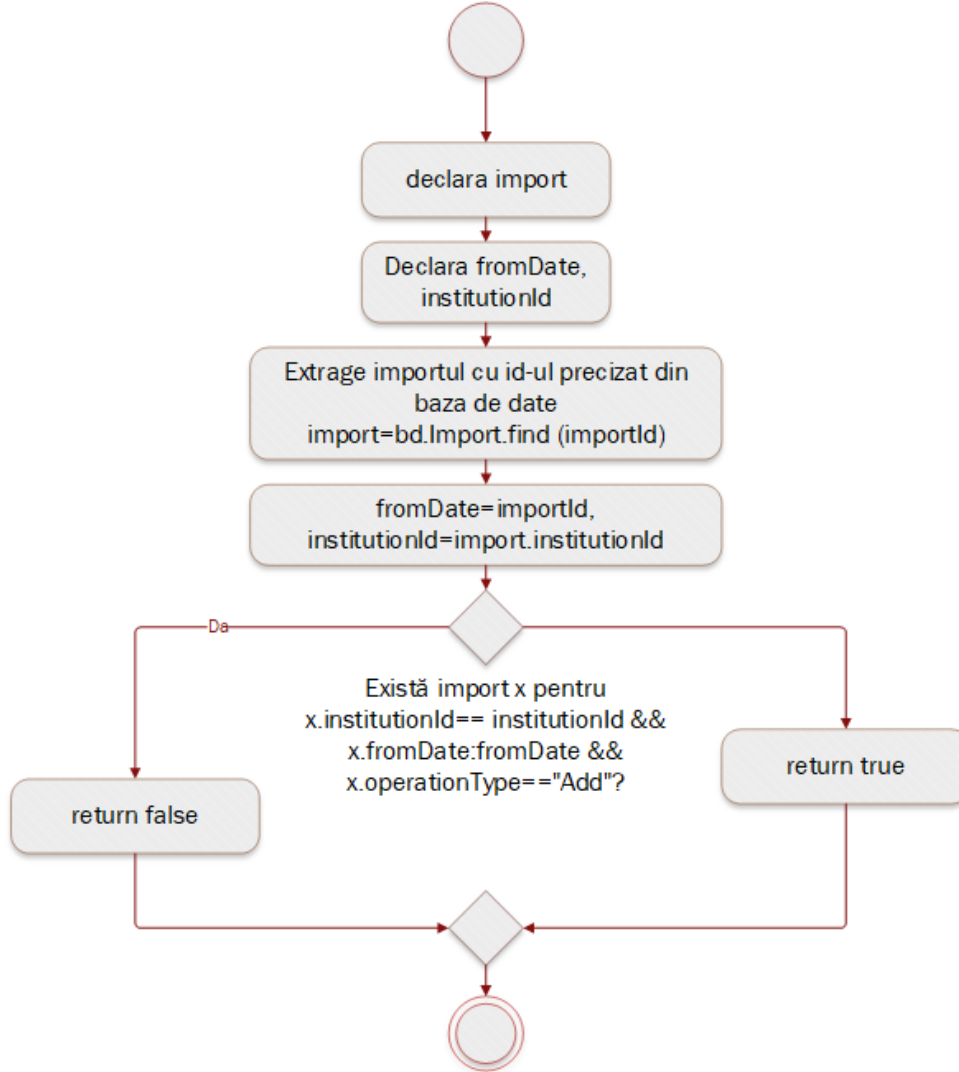


Figura 26. Diagrama de activități pentru verificarea existenței transferurilor după cel care urmează a fi șters

8. **AddNewUAT(string uatName,string uatusername,string uatpassword)**

Adaugă în nou UAT (institution) prin specificarea numelui acestuia, unui nume de utilizator care nu a mai fost folosit la un alt UAT și a unei parole specifice doar aceluia UAT. Returnează un șir în format JSON cu starea operațiunii de adăugare, eventual lista de erori identificate.

9. **DeleteUAT(long uatID)**

Șterge un UAT adăugat anterior. Parametrul UatID permite specificare UAT-ului de șters. Nu poate fi șters un UAT dacă există importuri pentru acesta. Returnează un șir de tip JSON cu starea operațiunii de ștergere (succes sau lista de erori survenite la încercarea de ștergere).

3.4 Sisteme de asistare a deciziei la nivel tactic și strategic pentru programul PPCA

Colectarea deșeurilor menajere constituie un pilon critic în cadrul mai larg al gestionării deșeurilor și al durabilității mediului. Pe măsură ce populațiile urbane continuă să crească, practicile eficiente de gestionare a deșeurilor devin esențiale pentru sănătatea publică, dezvoltarea sustenabilă a localităților, conservarea mediului și resurselor.

Colectarea eficientă a deșeurilor menajere este în strânsă conexiune cu:

- **Sănătate publică și implicații asupra mediului**
Colectarea eficientă a deșeurilor menajere este unul din factorii fundamentali pentru protejarea sănătății publice. Previne poluarea, diminuează riscul apariției unor epidemii și asigură locuitorilor un mediu de viață curat, sănătos și sigur.
- **Gestionarea eficientă a mediului înconjurător și conservarea resurselor**
Colectarea responsabilă a deșeurilor menajere joacă un rol esențial în conservarea resurselor naturale. Prin separarea deșeurilor organice, reciclabile și nereciclabile, gospodăriile contribuie la reducerea presiunii asupra depozitelor de gunoi și a instalațiilor de incinerare.
- **Bunăstarea comunității și atractivitatea estetică a localităților**
Sistemele de colectare a deșeurilor menajere bine organizate sporesc atractivitatea estetică a cartierelor și stimulează un sentiment de mândrie comunitară, iar străzile și spațiile publice mai curate contribuie la o calitate mai ridicată a vieții rezidenților.

Arhitectura depozitului de date este de tip constelație și este prezentată în Figura 10 și Figura 11.

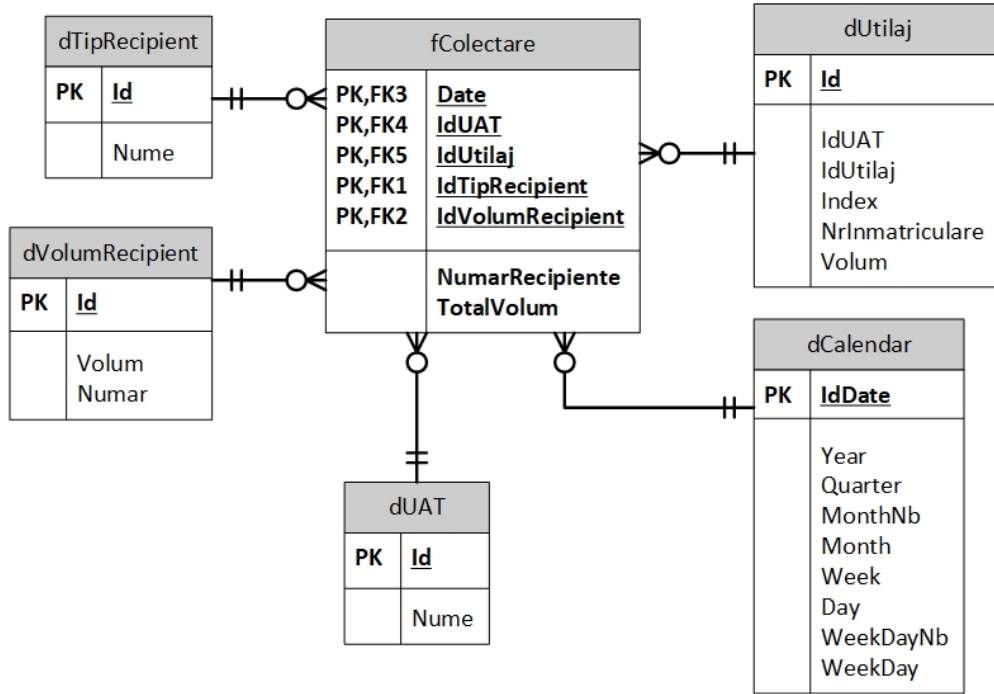


Figura 10 Secvență din arhitectura depozitului de date

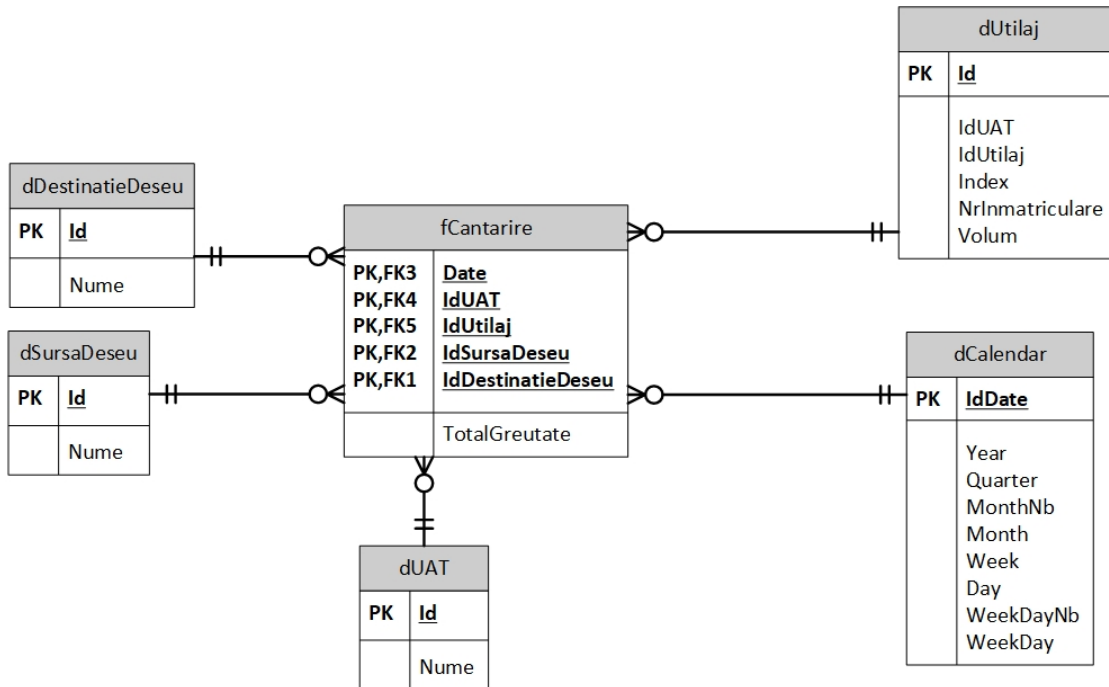


Figura 11 Secvență din arhitectura depozitului de date

Utilizarea tablourilor de bord în colectarea deșeurilor menajere

- Monitorizarea și raportarea în timp real: tablourile de bord oferă gospodăriilor și companiilor implicate în gestionarea deșeurilor menajere informații în timp real asupra operațiunilor de colectare a deșeurilor. Senzorii din containerele de deșuri, urmărirea GPS pe vehiculele de colectare și tehnologia Internet of Things (IoT) converg pentru furnizarea de date și informații în timp real referitoare la procesul colectării deșeurilor.
- Optimizarea și planificarea rutelor: integrarea sistemelor de informații geografice (Geographic Information Systems - GIS) și a algoritmilor avansați de inteligență artificială oferă noi modalități de optimizare a rutelor de colectare a deșeurilor, fapt care conduce la un consum mai redus de combustibil, emisii mai scăzute și o eficiență operațională sporită.
- Analiza predictivă pentru estimarea cantităților colectate: prin utilizarea datelor istorice și algoritmii de învățare automată (machine learning), tablourile de bord pot propune modelele care să estimeze cantitățile de deșuri care vor trebui colectate, fapt care permite alocarea eficientă de resurse din partea companiilor care asigură colectarea deșeurilor.
- Tablouri de bord pentru conformitatea cu reglementările: tablourile de bord urmăresc respectarea reglementărilor de gestionare a deșeurilor și conformitatea cu standardele de mediu existente la nivel național și/sau european, "semnalizând" alinierea sau nu a practicilor de colectare a deșeurilor cu cerințele legale.
- Tablouri de bord pentru evaluarea performanței: tablourile de bord cu indicatori cheie de performanță (Key Performance Indicator - KPI) măsoară diverse metrici de performanță, cum ar fi eficiența colectării, timpul de răspuns, satisfacția clienților etc., oferind o imagine de ansamblu cuprinzătoare asupra eficienței operaționale a autorităților locale sau centrale.
- Tablouri de bord pentru evaluarea impactului asupra mediului: tablourile de bord se pot concentra pe valori cheie legate de eforturile de reciclare, prin urmărirea procentului de deșuri "deturnate" de la gropile de gunoi, monitorizarea ratelor de reciclare și evaluarea impactului asupra mediului al programelor de reciclare.

- Tablouri de bord pentru educarea și implicarea comunității: tablourile de bord pot monitoriza eficiența inițiativelor educaționale care vizează promovarea practicilor de gestionare responsabilă a deșeurilor în cadrul comunităților și oferă suport în cuantificarea participării locuitorilor, nivelurile de conștientizare și impactul programelor de informare.
- Tablouri de bord pentru răspunsul la situațiile de urgență: aceste tablouri de bord sunt concepute pentru a face față situațiilor de urgență, cum ar fi dezastrele naturale, crizele de sănătate publică etc., oferind informații în timp real despre nevoile de colectare a deșeurilor în timpul situațiilor de urgență.
- Tablouri de bord de evaluare comparativă a performanței: aceste tablouri de bord permit ministerului să compare performanța de colectare a deșeurilor menajere cu valorile de referință, cu regiunile învecinate cu media la nivel național etc, fapt care ajută la identificarea zonelor de îmbunătățire și la implementarea celor mai bune practici.
- Tablouri de bord pentru raportarea către stakeholderi: aceste tablouri de bord facilitează raportarea datelor privind colectarea deșeurilor către diferite părți interesate, inclusiv publicul, organismele de reglementare și alte agenții guvernamentale, oferind date și informații în mod transparent despre eforturile ministerului de gestionare a deșeurilor menajere.

În figurile următoare sunt prezentate o serie de tablouri de bord pentru analiza colectărilor la nivel de UAT sau grup de UAT-uri, zilnice, lunare sau anuale.

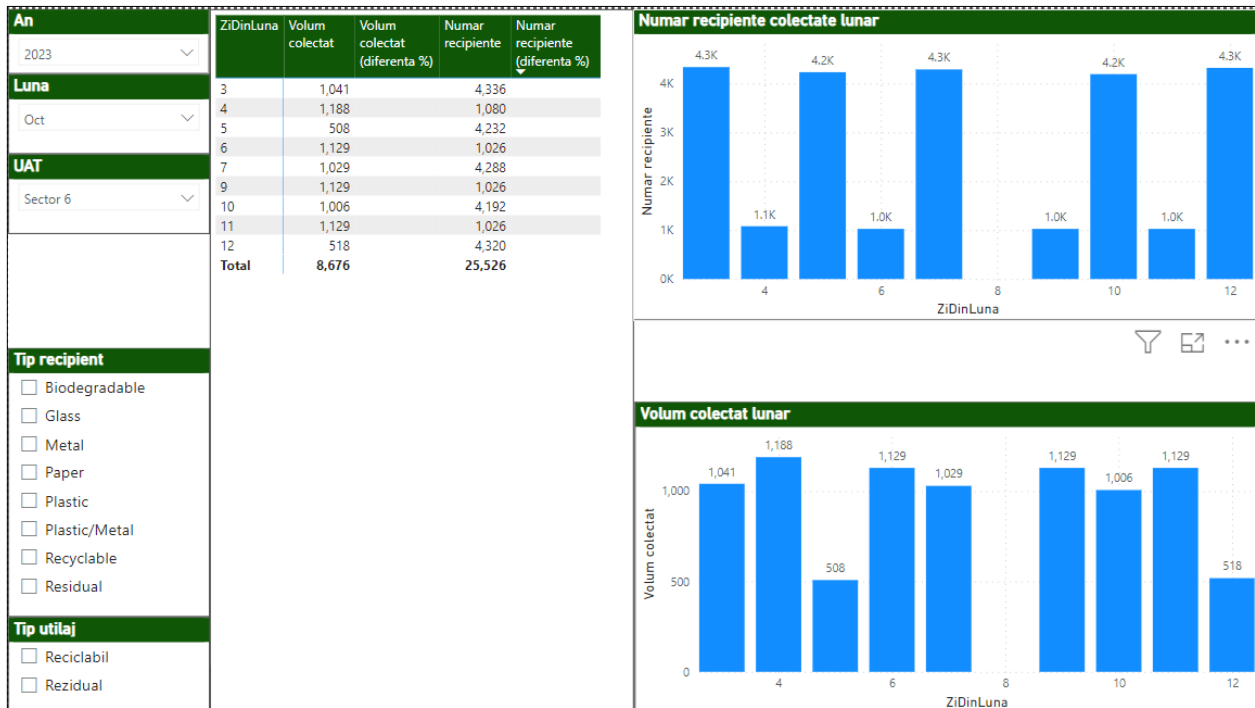


Figura 27. Colectari zilnice la nivel de UAT

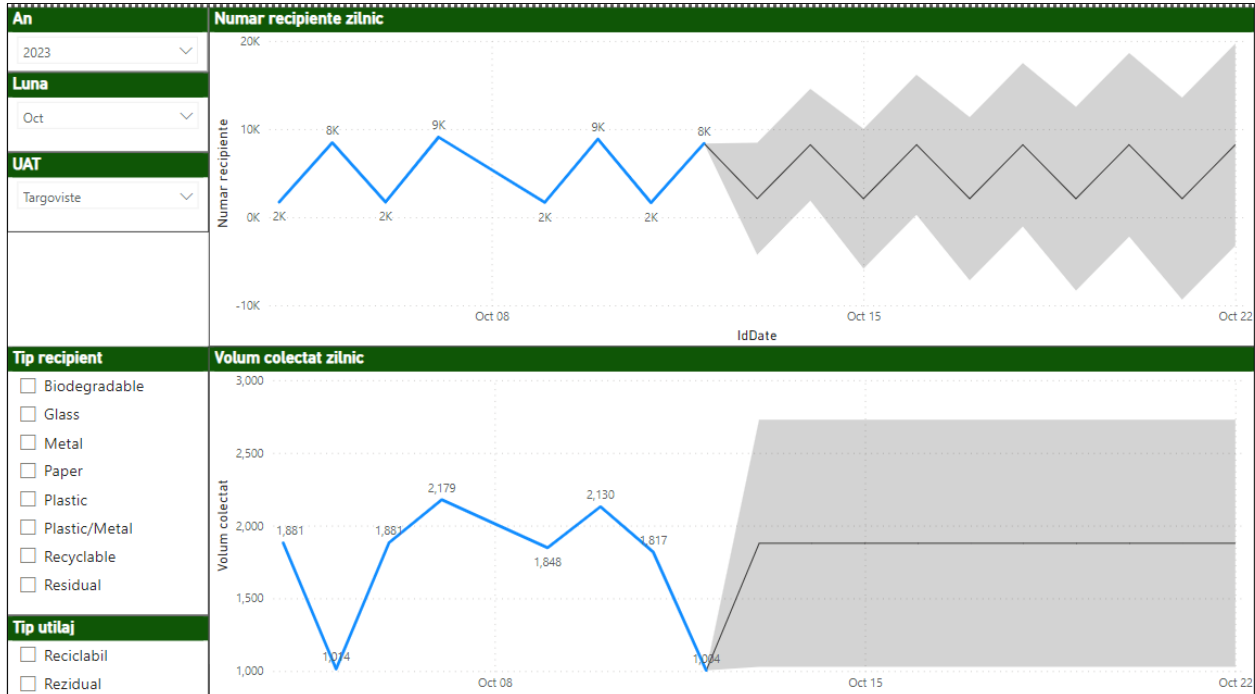


Figura 28. Estimări colectări zilnice la nivel de UAT

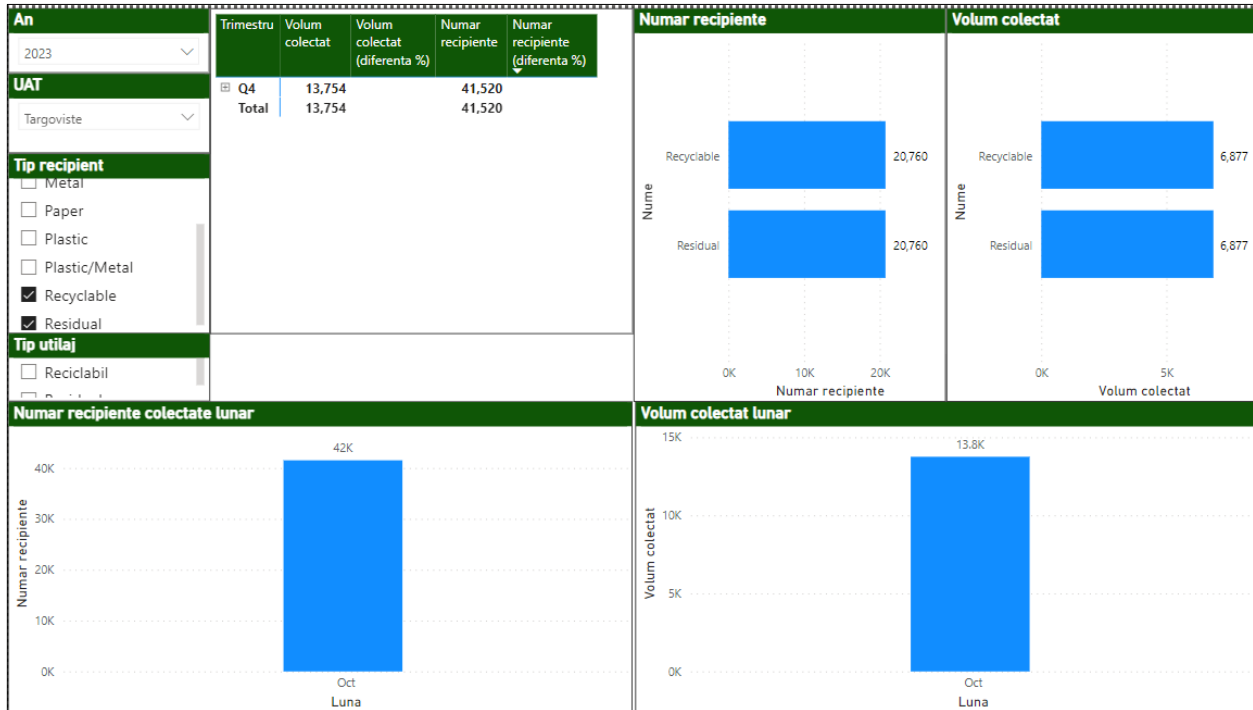


Figura 29. Colectări lunare la nivel de UAT

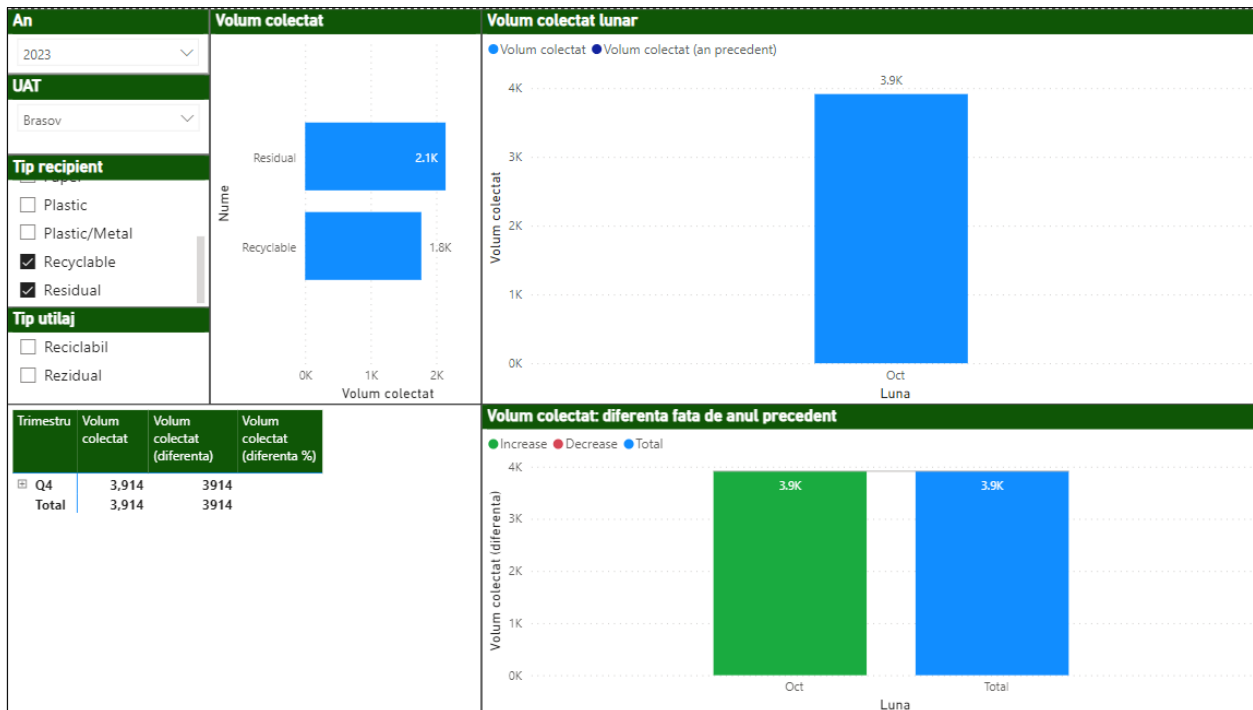


Figura 30. Colectări lunare la nivel de UAT (comparație cu anul precedent)

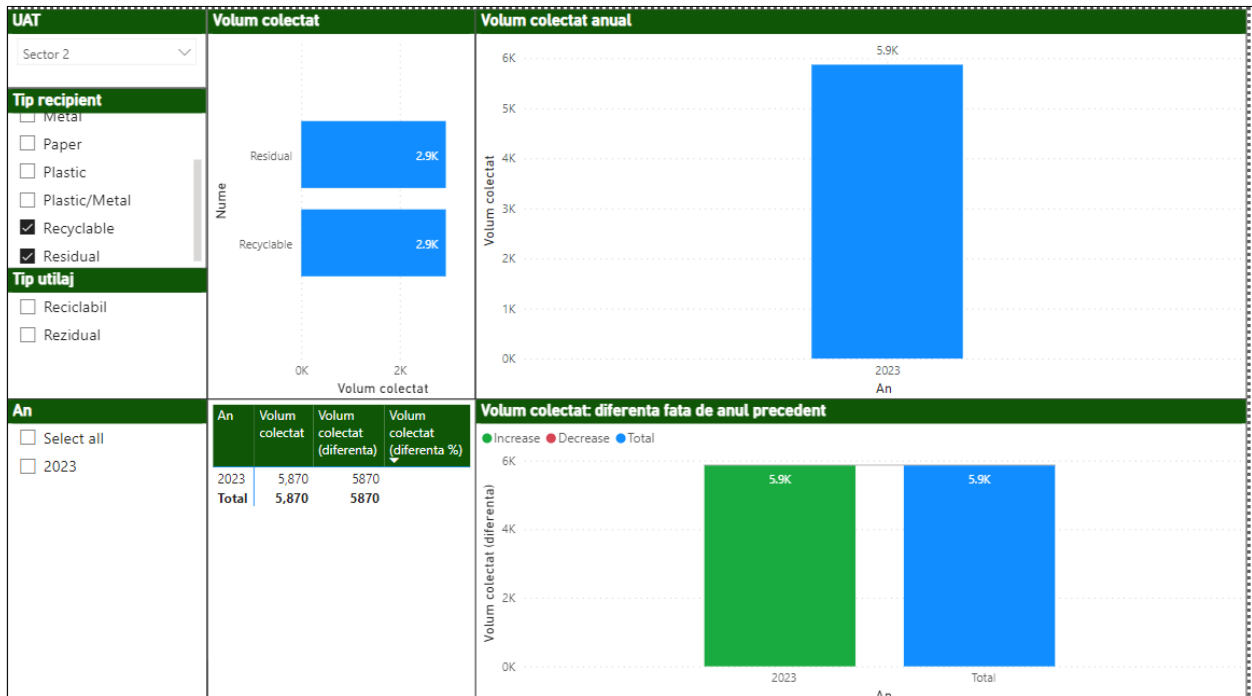


Figura 31. Colectări anuale la nivel de UAT

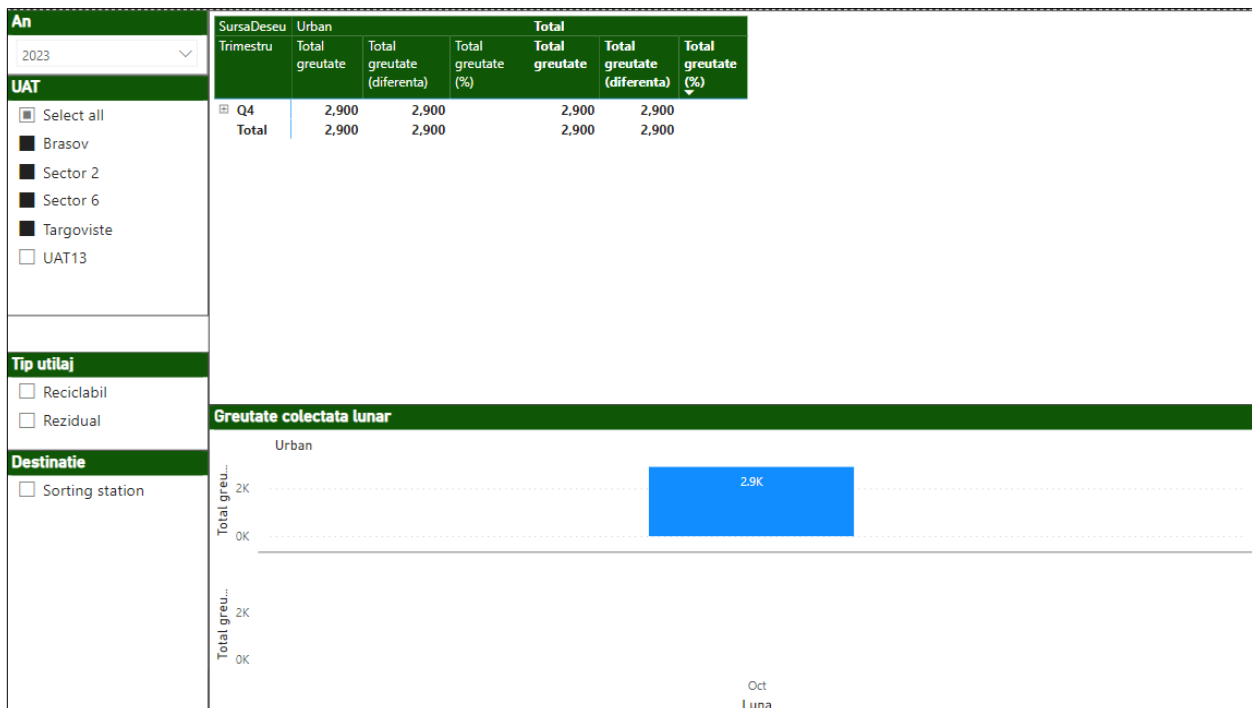


Figura 32. Colectări lunare (selecție multiplă la nivel de UAT)

4 Concluzii

Un depozit de date la nivel național este un element fundamental pentru stabilirea unui sistem eficient de tipul „Plătește pe măsură ce arunci”. Permite adoptarea deciziilor bazate pe date (data-driven decisions), facilitează conformitatea cu reglementările, sprijină monitorizarea performanței și contribuie la dezvoltarea unor practici de gestionare a deșeurilor menajere mai durabile și mai responsabile din punct de vedere ecologic la nivel național.

Procesul de extragere, transformare și încărcare a datelor reprezintă un element esențial în domeniul managementului și analizei datelor. Capacitatea sa de a integra perfect surse de date eterogene, de a curăța și transforma datele brute și de a le încărca în depozitele de date permite organizațiilor să gestioneze mult mai eficient procesele decizionale și implicit să adopte decizii mai bune, bazate pe date și informații superioare calitativ. În ciuda provocărilor sale, implementarea celor mai bune practici ETL poate atenua riscurile și poate asigura fluxul fără probleme al datelor. Pe măsură ce organizațiile continuă să valorifice puterea datelor, procesul ETL rămâne un instrument indispensabil în contribuind la menținerea avantajului competitiv.

Capacitatea tablourilor de bord de a furniza date în timp real, de a optimiza rutele de colectare, de a estima generarea de deșeurii menajere și de a sprijini inițiativele de durabilitate le face instrumente esențiale pentru a avea comunități mai curate și mai durabile.



Bibliografie

Abba S., Light C.I. IoT-Based Framework for Smart Waste Monitoring and Control System: A Case Study for Smart Cities (2020). *Engineering Proceedings.*, 2(1):90. Disponibil online la: <<https://doi.org/10.3390/ecsa-7-08224>> [Accesat la 22.08.2023]

Alexandru A., Coardos D.(2017). Big Data–Concepte, Arhitecturi și Tehnologii. Disponibil online: <https://rria.ici.ro/wp-content/uploads/2017/03/04-ART2-RRIA-1-2017-Alexandru-Coardos-big-data-2-1.pdf> [Accesat la 17.08.2023]

Bilitewski, B.(2008) .From traditional to modern fee systems Waste Management, 28 (12), pp. 2760-2766

Berglund,C. (2005). Burning in moral, drowning in rationality? Ethical considerations in forming environmental policy. *Minerals & Energy Raw Materials Report*,20 (1), pp. 16-22

Cox,M. , Ellsworth, D. (1997). Application-controlled demand paging for out-of-core visualization. *Proceedings. Visualization '97* (Cat. No. 97CB36155), Phoenix, AZ, USA, 1997, pp. 235-244, doi: 10.1109/VISUAL.1997.663888.

Card,D.,Schweitzer E.-P. (2016). Pay-As-You-Throw schemes in the Benelux countries.Disponibil online la: <<https://ieep.eu/wp-content/uploads/2022/12/BE-NL-LU-PAYT-final.pdf>> [Accesat la:13.08.2023]

De Jaeger,S.,Eyckmans,J.From pay-per-bag to pay-per-kg: the case of Flanders revisited. *Waste Management&Research*, 2015, 33 (12), pp. 1103-1111

Comisia Europeană (2022). Evaluarea din 2022 a punerii în aplicare a politicilor de mediu. Raport de țară-România. Disponibil online la: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=comnat%3ASWD_2022_0271_FIN [Accesat la 19.08.2023]

Dey,S. (2023). 4 Smart Waste Management Solutions That Are Revolutionising the Industry. Disponibil online la: <https://earth.org/smart-waste-management/> [Accesat la 7.05.2023]

European Commission, 2011. Report on the thematic strategy on waste prevention and recycling. <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX:52011DC0013> [Accesat la 5.08.2023]



European Commission (2023). Waste Early Warning Report. Disponibil online la: https://environment.ec.europa.eu/publications/waste-early-warning-report_en [Accesat la 28 august 2023]

European Parliament (2018). European Parliament and Council Directive (EU) 2018/851 amending directive 2008/98/EC on waste. Official Journal of the European Union

European Environment Agency (2023). Economic instruments and separate collection systems key strategies to increase recycling. Disponibil online la: <https://www.eea.europa.eu/publications/economic-instruments-and-separate-collection/economic-instruments-and-separatecollection/download.pdf.static> [Accesat 12.08.2023]

Eurostat (2023a). Municipal waste generated 2006 and 2021. Disponibil online la: https://ec.europa.eu/eurostat/statistics-explained/index.php?title=Municipal_waste_statistics [Accesat la 29.08.2023]

Eurostat (2023b). Waste generation 2020. Disponibil online la: https://ec.europa.eu/eurostat/statisticsexplained/index.php?title=Waste_statistics#Total_waste_generation. [Accesat la 2.08.2023]

Foundation Operate.it (2019). Pay as you throw (PAYT) toolkit for european cities. Action 12 – circular economy's EU urban agenda partnership. Disponibil online la: <https://ec.europa.eu/futurium/en/circular-economy/toolkit-establishing-pay-you-throw-schemes-cities.html> [Accesat la 14.08.2023]

Giffinger, R. (2017). Smart cities Ranking of European medium sized cities. *Centre of Regional Science, Vienna University of Technology*. 16, 13–18.

Halevi, G. and Moed, Henk F. (2012) .The evolution of big data as a research and scientific topic: Overview of the literature," *Research Trends: Vol. 1 : Iss. 30, Article 2*. Available at: <https://www.researchtrends.com/researchtrends/vol1/iss30/2>

Hill, J; Begin, A și Shaw, B (2002). Creative policy packages for waste – lessons for the UK. Disponibil online la: https://issuu.com/greenallianceuk/docs/creative_policy_packages_for_waste_overview [Accesat la:11.08.2023]



Hogg, D (2002). Financing and Incentive Schemes for Municipal Waste Management: Final Report to Directorate General Environment, European Commission.

Hogg, D; Sherington, C and Vergunst, T (2011). A Comparative Study on Economic Instruments Promoting Waste Prevention: Final Report to Bruxelles Environnement.

IPCC (2018). Global warming of 1.5 °C. Disponibil online la <<https://www.ipcc.ch/sr15>> [Accesat la 3.08.2023]

Kaza, S., Yao, L.C., Bhada-Tata, P., Van Woerden, F. (2018). What a Waste 2.0: A Global Snapshot of Solid Waste Management to 2050. Urban Development;. © Washington, DC: World Bank. <http://hdl.handle.net/10986/30317>

Linderhof, V. G. M., Kooreman, P., Allers, M., & Wiersma, D. (2001). Weight-based pricing in the collection of household waste: the Oostzaan case. *Resource and Energy Economics*, 23(4), 359-371. [https://doi.org/10.1016/S0928-7655\(01\)00044-6](https://doi.org/10.1016/S0928-7655(01)00044-6)

OECD (2006). Working Group on Waste Prevention and Recycling. Impacts of Unit-based Waste Collection Charges

OECD (2010). Environmental Performance Reviews: Luxembourg 2010. OECD Publishing, pp77-78

Păvălasc, M., Burlă, V. (2023). România sufocată de gunoaie. Gropi ilegale, bani pierduți, procese cu Uniunea Europeană. Disponibil online la: <https://romania.europalibera.org/a/romania-sufocata-de-gunoaie-gropi-ilegale-bani-pierduti-procese-la-cjue/32342094.html> [Accesat la 19.08.2023]

Parlamentul European(2023). Gestionarea deșeurilor în UE în date și cifre (infografic). Disponibil online la: https://www.europarl.europa.eu/pdfs/news/expert/2020/11/story/20180328STO00751/20180328STO00751_ro.pdf [Accesat 12.08.2023]

Radu, C. (2022). Munții de gunoaie din România, miză de 1,2 miliarde de euro. Cum putem ajunge de la o țară a importurilor de gunoaie la „România Curată”. Disponibil online la: <https://economedia.ro/muntii-de-gunoaie-din-romania-miza-de-12-miliarde-de-euro-cum-putem-ajunge-de-la-o-tara-a-importurilor-de-gunoaie-la-romania-curata.html> [Accesat la 10.08.2022]



Reveiu,A.(2003). Modelul semistructurat al datelor. *Revista Informatica Economica*, nr. 2(26)/2003 75

Reichenbach, J. (2008). Status and prospects of pay-as-you-throw in Europe – a review of pilot research and implementation studies. *Waste Management*. 28 (12), pp. 2809-2814

Smart City Magazine (2016). Bigbelly, colectarea deșeurilor ca parte integrantă a unei economii circulare. Disponibil online la: <https://smartcitymagazine.ro/bigbelly-colectarea-deseurilor-ca-parte-integrantaa-a-unei-economii-circulare/> [Accesat la 23.08.2023]

Statista (2023). Generation of municipal waste worldwide as of 2021, by select country (in million metric tons). Disponibil online la: <https://www.statista.com/statistics/916749/global-generation-of-municipal-solid-waste-by-country/> [Accesat la 21.08.2023]

Ukkonen,A., Sahimaa,O.(2021).Weight-based pay-as-you-throw pricing model: Encouraging sorting in households through waste fees.*Waste Management*, Vol.135,pp. 372-380. DOI: <https://doi.org/10.1016/j.wasman.2021.09.011>

United Nations (2018). World urbanization prospects: the 2018 revision. Disponibil la: <https://www.un.org/en/development/desa/population/theme/urbanization/index.asp>. [Accesat la 10.07.2023]

Vegacomp Consulting (2022). Smart City Scan in Romania 6th edition, June 2022. Disponibil online la: https://vegacomp.ro/wpr/wp-content/uploads/2022/08/radiografia-smart-city_2022_english_final.pdf [Accesat la 28.08.2023]

World Bank Report, Washington. 20 September 2018. Disponibil online: <https://www.worldbank.org/en/news/press-release/2018/09/20/global-waste-to-grow-by-70-percent-by-2050-unless-urgent-action-is-taken-world-bank-report>, [Accesat 12.07.2023]

Amazon Web Services [AWS], 2023. What's the Difference Between a Data Warehouse, Data Lake, and Data Mart? Disponibil la: <https://aws.amazon.com/compare/the-difference-between-a-data-warehouse-data-lake-and-data-mart/#:~:text=A%20data%20warehouse%20stores%20data,%2C%20marketing%2C%20or%20sales%20department>. Accesat: ianuarie, 2023



Bellini, P., Palesi, L.A.I., Giovannoni, A. and Nesi, P., 2023. Managing Complexity of Data Models and Performance in Broker Based Internet/Web of Things Architectures. *Internet of Things*, 23, p.100834. 10.1016/j.iot.2023.100834

Coursera, 2023. Data Lake vs. Data Warehouse: What's the Difference? Disponibil la: <https://www.coursera.org/articles/data-lake-vs-data-warehouse>. Accesat: iulie, 2023

Gorelik, A., 2023. The Enterprise Big Data Lake. O'Reilly Media, Inc. Disponibil la: <https://www.automationworld.com/analytics/article/22144680/4-ways-to-ensure-your-data-lake-doesnt-become-a-data-swamp>. Accesat: septembrie, 2023

Greenfield, D., 2022. 4 Ways to Ensure Your Data Lake Doesn't Become a Data Swamp. Disponibil la: <https://www.automationworld.com/analytics/article/22144680/4-ways-to-ensure-your-data-lake-doesnt-become-a-data-swamp>. Accesat: februarie, 2023

Gruenwald, C., Hustvedt, A., Beach, A. and Han, R., 2007, May. SWARMS: A sensornet wide area remote management system. In *2007 3rd International Conference on Testbeds and Research Infrastructure for the Development of Networks and Communities*, Lake Buena Vista, FL, USA, pp. 1-10. *IEEE*. 10.1109/TRIDENTCOM.2007.4444657

Hagstroem, M., Roggendorf, M., Saleh, T., and Sharma, J., 2017. A smarter way to jump into Data Lakes. Disponibil la: <https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/a-smarter-way-to-jump-into-data-lakes>. Accesat: martie, 2023

Hewlett Packard Enterprise Development LP [HPE], 2023. Data Lakehouse. Disponibil la: <https://www.hpe.com/us/en/what-is/data-lakehouse.html>. Accesat: septembrie, 2023