

# Catalog de date în medii Cloud pentru dezastre naturale

Marius Cristian CÎRȚÎNĂ<sup>1</sup>, Florin POP<sup>1,2</sup>, Mădălina Cornelia ZAMFIR<sup>2</sup>,  
Vladimir FLORIAN<sup>2</sup>, Gabriel NEAGU<sup>2</sup>

<sup>1</sup> Universitatea Politehnica București

<sup>2</sup> Institutul Național de Cercetare-Dezvoltare în Informatică – ICI București

marius.cirtina@gmail.com, florin.pop@ici.ro, madalina.zamfir@ici.ro,

vladimir.florian@ici.ro, gabriel.neagu@ici.ro

**Rezumat:** În prezent se observă o creștere continuă a numărului de surse și a volumului de date geospațiale stocate și, în același timp, o creștere a cererii de accesare a acestor date. Realizarea unei platforme web cu rol în gestionarea și agregarea acestor date, precum și facilitarea accesului la informații structurate în funcție de diferite criterii selectate de utilizator, este de importanță majoră. Se impune atât elaborarea de sisteme informatice care să urmărească stocarea informațiilor, prelucrarea și vizualizarea datelor referitoare la dezastre și crize majore cu scopul monitorizării acestora și a diminuării efectelor, cât și dezvoltarea unor instrumente de înregistrare în sistem, vizualizare, prelucrare și raportare a evenimentelor, actuale și istorice. Lucrarea de față prezintă o soluție de tip platformă cloud de informare asupra dezastrelor și crizelor majore. Platforma este bazată pe un catalog de date și pe interogări ale bazei de date create, care poate fi dezvoltată prin completarea bazei de date cu informații de actualitate, prin integrarea unor noi surse de date și corelarea cu platforme externe. Sunt descrise arhitectura platformei propuse, fluxul de lucru, funcționalități de bază, detalii de implementare și rezultate experimentale.

**Cuvinte cheie:** platformă Cloud, modele de date, date geospațiale, Model View Controller, Web Crawling.

## Data catalog in Cloud environment for natural disasters

**Abstract:** At this moment there is a continuing development of the sources and the amount of geospatial data stored and, in the same time, a growing demand for access to these data. Creating a web platform with a role in managing and aggregating this data, as well as facilitating access to structured information based on different user-selected criteria, is of major importance. It is necessary to develop information systems for the storage of information, processing and visualization of data related to major disasters and crisis, with the purpose of monitoring and mitigating the effects, as well as development of system recording tools, visualization, processing and reporting of current and historical events. This paper presents a cloud-based information platform to support major disasters and crisis monitoring. The platform is based on a data catalogue and database queries, which can be developed by completing the current data base with up-to-date information, integrating new data sources and correlation with external platforms. The proposed platform architecture, workflow, basic functionalities, deployment details and experimental results are described.

**Keywords:** Cloud Platform, Data Models, Geospatial Data, Model View Controller, Web Crawling.

### 1. Introducere

Managementul dezastrelor include activitățile de organizare și gestionare a resurselor și responsabilităților ce urmăresc minimizarea pierderilor umane și materiale. Un management activ înseamnă găsirea de informații solide și în timp cvasi-real, pentru a sprijini intervenția autorităților administrative locale și centrale, a coordona activitățile umanitare și a înlătura amenințările la adresa securității civililor. Acest fapt impune utilizarea și analiza datelor satelitare și multi-sursă, precum și metode de cartografiere rapidă. După identificarea surselor de date de intrare, preluarea și procesarea în cadrul unui sistem informatic a informațiilor din diverse surse de date caracteristice diferitelor tipuri de dezastre presupun evaluarea posibilităților de implementare a sistemului de

gestiune a datelor și modelelor de date utilizate, precum și realizarea unei baze de date de referință. Informațiile stocate în baza de date de referință pot fi accesate rapid și vizualizate prin selectarea lor, în funcție de anumite criterii specificate de utilizatori.

Standardele web referitoare la datele geospațiale sunt deosebit de importante, deoarece în baza acestora au fost adoptate legi sau norme de reglementare cu privire la dezvoltarea de infrastructuri naționale de date spațiale. La nivel internațional, există o serie de instituții care se ocupă cu elaborarea și promovarea standardelor deschise în domeniul geospațial. La nivel național sunt coordonate în mod unitar și permanent activitățile de prevenire și gestionare a situațiilor de urgență, prin evaluarea riscurilor de dezastre la nivel național și prelucrarea datelor satelitare.

Trebuie dezvoltate instrumente pentru evaluarea unor evenimente, precum inundații și situația hidrologică în timp real, fenomene meteorologice, cutremure, secetă. Tehnologiile implicate în dezvoltarea platformei Cloud sunt: C#, ASP.Net, MVC, HTML, MySql, TypeScript, HTTP, XML, tehnologii folosite și în alte platforme de gestiune a datelor [5, 6].

Platforma propusă în această lucrare furnizează o soluție pentru colectarea și agregarea datelor referitoare la gestionarea dezastrelor și crizelor majore pornind de la o analiză generală a sistemelor existente atât la nivel național cât și internațional. De asemenea, platforma oferă posibilitatea accesării rapide a unor informații specifice evenimentelor actuale și istorice, stocate într-o bază de date de referință. În continuare, după secțiunea introductivă, lucrarea este structurată în alte patru secțiuni: în secțiunea 2 este prezentată situația soluțiilor folosite în prezent în gestionarea și colectarea datelor referitoare la dezastrele și crizele majore; secțiunea 3 este dedicată descrierii arhitecturii soluției propuse. Arhitectura este organizată pe trei nivele funcționale: nivelul surselor de date, nivelul de management al datelor și nivelul de vizualizare. De asemenea, sunt descrise componentele și fluxul de lucru din cadrul platformei; secțiunea 4 prezintă câteva detalii de implementare a soluției pentru gestionarea și colectarea datelor referitoare la dezastre și situații de criză majoră, cu accent pe componentele de colectare date, agregare și vizualizare; în secțiunea 5 sunt prezentate rezultatele experimentale prin prisma capabilității de vizualizare a datelor gestionate în cadrul platformei, precum și evidențierii performanțelor obținute în utilizarea acesteia și posibilitatea realizării unor îmbunătățiri și dezvoltări ulterioare a soluției propuse; în secțiunea 6 a lucrării sunt formulate concluziile și modul în care se poate dezvolta ulterior platforma. Lucrarea se încheie cu o bibliografie.

## 2. Soluții existente

Sunt descrise în continuare principalele **modele de date** care pot fi utilizate în scopul elaborării unei platforme de monitorizare a dezastrelor și crizelor majore, manifestate la nivel local, național și regional.

Cercetările și studiile realizate în domeniul tehnologiilor de stocare și management al datelor masive (Big Data) se pot clasifica pe trei niveluri, dispuse de la bază spre vârf, astfel: sisteme de fișiere, baze de date și modele de programare [3].

Sistemele de gestiune a bazelor de date sunt construite și mai apoi dezvoltate cu scopul de a oferi suportul pentru prelucrarea seturilor de date de diferite dimensiuni ce pot fi utilizate în cadrul aplicațiilor software.

Pentru gestionarea datelor de mari dimensiuni se folosesc bazele de date de tipul NoSQL (baze de date non-relaționale) [11, 12], care oferă suport pentru copierea datelor mult mai simplu și mai ușor, moduri flexibile de lucru, un API simplu și coerența eventuală a datelor. Cele mai utilizate tipuri de baze de date NoSQL sunt: baze de date cheie-valoare, baze de date orientate pe coloane și baze de date dedicate pentru documente.

**Bazele de date cheie-valoare** sunt constituite în baza unui model al datelor simplu, model potrivit căruia datele sunt stocate corespunzător unei perechi cheie-valoare. Fiecare cheie este unică, iar clienții pot introduce valori pentru fiecare cheie.

**Bazele de date orientate pe coloane** diferă de cazul sistemelor relaționale tradiționale, întrucât stochează și prelucrează datele în funcție de coloane și nu de rânduri. Atât coloanele cât și rândurile sunt segmentate în mai multe noduri pentru a suporta extinderea bazei de date. Bazele de date orientate pe coloane au fost inspirate în principal de BigTable, sistem ce a fost realizat de către Google.

**BigTable** se bazează pe mai multe componente fundamentale ale infrastructurii Google, inclusiv GFS, sistemul de management al clusterelor de calcul, formatul de fișiere SSTable și Chubby [2]. GFS este folosit pentru a stoca datele și fișierele jurnal. Sistemul de management al clusterelor este folosit la gestionarea programării încărcării cu aplicații, partajarea resurselor, procesarea erorilor sistemelor de calcul și monitorizarea stării sistemelor. Fișierele în format SSTable sunt folosite pentru a stoca intern datele BigTable.

**Cassandra** este un sistem de stocare distribuit, special proiectat pentru a gestiona o cantitate mare de date structurate cu ajutorul unor servere obișnuite [10]. Acest tip de instrument a fost dezvoltat de către Facebook și a devenit open source în anul 2008. Sistemul de stocare Cassandra adoptă ideile și conceptele Amazon Dynamo și Google BigTable, mai ales în domeniul integrării tehnologiei sistemelor distribuite utilizate de Dynamo cu modelul structurării datelor utilizat de BigTable.

Dintre bazele de date dedicate pentru documente care pot gestiona tipuri de date mult mai complexe fac parte:

**MongoDB**, care reprezintă un produs open-source pentru stocarea documentelor care utilizează obiecte de tipul Binary JSON (BSON) pentru stocarea datelor [4]. Este o bază de date orientată pe mai multe platforme, care oferă o disponibilitate și performanță ridicată, precum și o scalabilitate ușoară.

**SimpleDB** oferă o interfață bazată pe servicii web, dezvoltată de către Amazon și organizată sub formă de domenii în care pot fi stocate datele. Datele sunt copiate pe diferite mașini aflate în diferite centre de date cu scopul de a se asigura siguranța datelor și de a se îmbunătăți performanța.

Sunt descrise în continuare principalele organizații specializate în crearea de standarde pentru datele geospațiale, reglementări legislative în domeniu.

În baza standardelor web referitoare la datele geospațiale sunt adoptate legi sau norme de reglementare cu privire la dezvoltarea de infrastructuri naționale de date spațiale, de către următoarele organisme de standardizare:

- *Open Geospatial Consortium (OGC)* (<http://www.opengeospatial.org>) este o organizație non-profit specializată în dezvoltarea de standarde pentru datele și serviciile cu caracter geospațial care promovează continuu adoptarea standardelor prin intermediul site-urilor web, a publicațiilor digitale sau tipărite, prin organizarea de seminarii practice și conferințe.
- *European Committee for Standardization / Comitetul European de Standardizare (CEN)* (<https://goo.gl/FN7SKG>) se ocupă cu adoptarea standardelor cu caracter geospațial la nivelul Uniunii Europene.
- *International Organization for Standardization / Organizația Internațională de Standardizare (ISO)* (<https://committee.iso.org/home/tc211>) este o confederație internațională de stabilire a normelor în toate domeniile, cu excepția electricității și a electronicii, ce sunt reprezentate de IEC (International Electrotechnical Commission) și cu excepția telecomunicațiilor, reprezentate de ITU (International Telecommunication Union). Grupul de lucru tehnic care se ocupă de standardele cu caracter geospațial poartă numele de ISO/TC 211.
- Asociația de Standardizare din România (ASRO) (<http://www.asro.ro>) este împuternicită să gestioneze standardizarea în România și adoptă standardele ISO și CEN fără traducere sau adaptare.

Una dintre reglementările legislative în domeniu este reprezentată de Directiva Infrastructură for Spatial Information in Europe (INSPIRE) (<http://inspire.ec.europa.eu>), care are ca scop

armonizarea informației geo-spațiale între țările europene la nivel local, regional și național, precum și facilitarea accesului de către cetățeni și a întreprinderilor private la informația spațială oriunde pe întreg cuprinsul Uniunii Europene.

**Standardele INSPIRE** pentru date și servicii de date se bazează pe cele emise de OGC/ISO/CEN. INSPIRE oferă specificații pentru următoarele servicii de rețea:

- *Servicii de căutare* care permit recunoașterea seturilor și serviciilor de date spațiale pe baza conținutului metadatelor și afișarea conținutului metadatelor;
- *Servicii de vizualizare* care permit cel puțin una din operațiile de afișare, navigare, mărire/micșorare, rotire panoramică, suprapunere vizuală a datelor spațiale.
- *Servicii de descărcare* care permit formarea de copii ale seturilor de date spațiale sau părți ale acestora, precum și accesarea directă a acestora atunci când este posibil;
- *Servicii de prelucrare* în care sunt prelucrate seturile de date spațiale în vederea realizării interoperabilității;
- *Servicii* care permit apelarea la serviciile de date spațiale.

În majoritatea țărilor membre ale Uniunii Europene s-au dezvoltat sisteme aplicabile în managementul dezastrelor și crizelor majore, adoptate în legislația națională a țării respective și deja operaționalizate în prezent.

**Copernicus** (<http://www.copernicus.eu/>) este programul european de monitorizare a Pământului, coordonat de CE în colaborare cu ESA. Programul are servicii din următoarele domenii: monitorizare atmosferică, terestră și maritimă, securitate, schimbări climatice și managementul dezastrelor.

Sistemul operațional **Copernicus EMS** (Emergency Management Service <http://emergency.copernicus.eu/>) este un sistem care urmărește dezvoltarea unui serviciu European de informație geospațială bazat pe date EO (Earth Observation), asigurate de misiunile Sentinel și date in-situ, furnizate de EEA - European Environment Agency.

**GEODIM** este un geoportal pentru produse cartografice derivate prin prelucrarea datelor satelitare (în special în caz de inundații) pentru managementul situațiilor de urgență (<http://www.geodim.meteoromania.ro/>).

**GDACS** este un cadru de cooperare care include factori de decizie pentru managementul dezastrelor și sisteme globale de informare (<http://www.gdacs.org/>).

*Din punct de vedere tehnologic*, ASP.NET este una dintre tehnologiile care au stat la baza dezvoltării soluției propuse, care se livrează cu contoare de performanță pe care dezvoltatorii și administratorii de sistem le pot monitoriza pentru a testa noi aplicații și pentru a aduna metrici pe aplicațiile existente [7]. ASP.NET oferă trei structuri conceptuale pentru crearea de aplicații web: Web Forms, ASP.NET MVC și ASP.NET Web Pages.

Modelul MVC oferă posibilitatea de a construi website-uri dinamice și crearea de aplicații care utilizează cele mai recente standarde web. MVC este structurat după modelul Model-View-Controller: componenta Model este responsabilă pentru datele din domeniul aplicației; componenta View prezintă modelul în interfața cu utilizatorul; componenta Controller este cea mai importantă componentă a MVC, intermediarul care leagă modelul și vizualul împreună, adică are nevoie de intrare de la utilizator, manipulează modelul și face ca actualizarea să fie preluată [15].

Cele mai importante beneficii ale framework-ului open-source ASP.NET MVC sunt: organizarea modulară, care permite izolarea părții de cod de interes de restul programului; folosirea tehnologiilor deja existente precum ORM-uri, framework-uri de logging, reducând efortul pentru implementarea de funcții de bază; testarea ușor accesibilă.

TypeScript este un limbaj compilat puternic, orientat obiect, proiectat în cadrul companiei Microsoft. TypeScript este JavaScript plus câteva caracteristici suplimentare, cum ar fi [9]: adoptă blocurile de bază ale programului de la JavaScript; este convertit în echivalentul său JavaScript în scopul executării; suportă alte biblioteci JS. Textul compilat poate fi consumat din orice cod

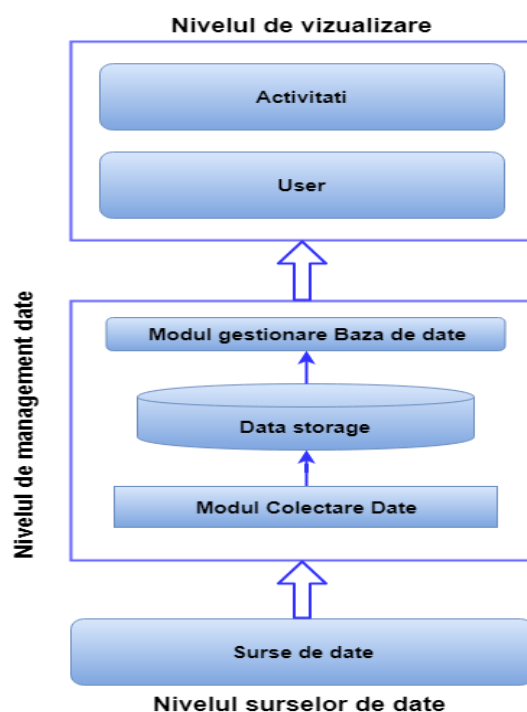
JavaScript; JavaScript este TypeScript. Aceasta înseamnă că orice fișier .js valabil poate fi redenumit în .ts și compilat cu alte fișiere tip; este portabil în browsere, dispozitive și sisteme de operare. Se poate executa pe orice mediu pe care rulează JavaScript.

### 3. Arhitectura soluției propuse

Arhitectura platformei propuse este organizată pe trei niveluri funcționale, după cum se observă în Figura 1:

- Nivelul surselor de date, care organizează și descrie sursele de date ce vor fi utilizate de către serviciile proiectului;
- Nivelul de management de date, care descrie etapele de achiziție de date și integrare;
- Nivelul de vizualizare, în care este inclusă platforma de livrare de date și de acces la date și servicii, precum și descrierea fluxului de execuție.

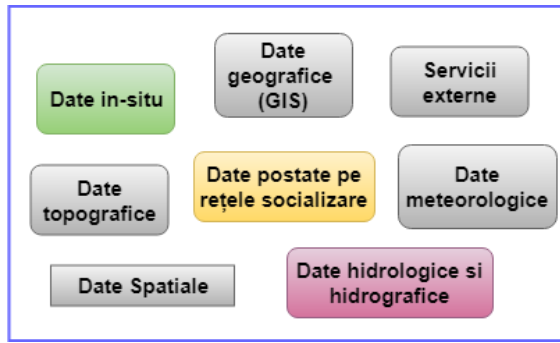
În modelul MVC, componenta Model va conține datele colectate într-un format specific.



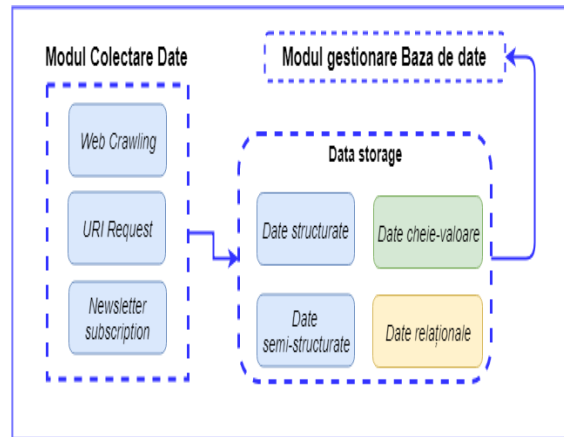
**Figura 1.** Arhitectura aplicației

Componenta View va conține nivelul de vizualizare și prezentare al aplicației construite folosind tehnologii web clasice, precum HTML și CSS, dar și structuri mai moderne cum ar fi Angular 5 și TypeScript. Componenta Controller se asigură prin implementarea nivelului de management de date, în care se vor apela servicii de preluare a datelor din sursele menționate în nivelul surselor de date, servicii de actualizare a datelor și metode de integrare a acestora în diferite unități structurate în funcție de diferite criterii.

**Nivelul surselor de date** reprezintă o caracteristică importantă a platformei, care definește varietatea surselor și tipurilor de date de intrare, organizează și descrie sursele de date ce vor fi utilizate de către serviciile proiectului și identifică anumite standarde ale datelor ce pot fi folosite în cadrul aplicației (Figura 2). Se descriu în continuare sursele de date ce vor fi utilizate de către serviciile platformei: Nomenclatorul național al gestionărilor de date geospațiale, Agenția Spațială Europeană și Institutul Național de Hidrologie și de Gospodărire a Apelor.



**Figura 2.** Nivelul surselor de date



**Figura 3.** Nivelul de management de date

În Nomenclatorul național al gestionărilor de date geospațiale (<http://geoportal.gov.ro/Geocatalog/Organization/Index>) au fost identificate sursele de date furnizate de activitățile instituțiilor înregistrate.

Informațiile relevante privind tipuri, surse și date asociate activităților desfășurate de instituțiile înscrise în acest nomenclator sunt centralizate.

Datele ESA folosite sunt reprezentate prin Date de tip Sentinel-1, Sentinel-2 și Sentinel-3.

Un produs de tip Sentinel se referă la un director care conține o colecție de informații și conține un fișier manifest.safe cu informațiile generale despre produs din sub-directoarele XML pentru seturi de date de măsurare cu date de imagine în diferite formate binare, un director de previzualizare, care conține fișiere quicklook în format PNG, suprapuneri Google Earth format KML și fișiere de previzualizare HTML, un director de adnotări cu metadatele produsului în XML, precum și date de calibrare și un director suport care conține schemele XML care descriu produsul XML.

Datele Sentinel-1 constau în date de nivel 0, care includ zgomotul, calibrarea internă și pachetele sursă ecou, precum și informații despre orbită și altitudine.

Datele concentrate nivel 1 sunt produse accesibile pentru majoritatea utilizatorilor de date. Datele de nivel 0 sunt transformate într-un produs de nivel 1, prin aplicarea algoritmilor și a datelor de calibrare, astfel încât să se obțină un produs de inginerie de bază de la care derivă niveluri mai ridicate.

Nivelul 2 constă în produse geofizice geografice derivate din nivelul 1, existând un singur produs standard de nivel 2 pentru aplicațiile de vânt, valuri și curenți și anume produsul nivel 2 Ocean (OCN). Pentru datele Sentinel-2, nivelul 0 este constituit de date brute comprimate și conține toate informațiile necesare pentru a genera datele produsului de nivel 1 (și superioare).

Datele Sentinel-3 constau în produse utilizator Level-1 și Level-2 (Land) pentru instrumente precum OLCI, care are ca obiectiv principal de a examina oceanul și suprafața terenului, pentru a colecta informații legate de biologie.

Datele hidrologice și hidrografice sunt gestionate prin intermediul aplicațiilor și programelor informatice specifice pe baza Sistemelor Relaționale de Gestionare a Bazelor de Date, fiind proprietate intelectuală a Institutului Național de Hidrologie și de Gospodărire a Apelor. Administrarea datelor se realizează atât prin introducerea, validarea și corectarea șirurilor de date și metadate hidrologice și hidrogeologice curente și istorice.

Datele se constituie în structură relațională de baze de date în fluxul lent, dar și prin administrarea serverului utilizat pentru stocarea acestora.

Datele administrate în flux lent sunt:

- date hidrologice și hidrogeologice obținute din activitățile de observare și măsurare, fiind considerate date brute;
- date hidrologice și hidrogeologice rezultate în urma validărilor definitive a datelor brute, denumite date primare;
- date hidrologice și hidrogeologice care au la baza datele primare, dar care sunt obținute prin analize specifice (algoritmi specializați domeniului de analiză), și se numesc date derivate sau prelucrate.

Sistemul de informații geografice (GIS) reprezintă un sistem informatic destinat captării, stocării, gestionării și afișării tuturor tipurilor de date spațiale sau geografice. Este un ansamblu de instrumente care permite utilizatorilor să efectueze interogări spațiale, să analizeze și să editeze date spațiale, fiind un sistem folosit pentru a crea și imprima hărți. Pentru a îndeplini sarcina de bază în GIS, straturile sunt combinate, editate și proiectate [13]. O caracteristică importantă a GIS este capacitatea de a combina diferite straturi pentru a afișa noi informații.

**Nivelul de management al datelor** este nivelul la care are loc preluarea datelor, examinarea și încapsularea acestora pentru a putea fi gestionate în cadrul unei baze de date și vizualizate în paginile web construite în cadrul platformei (Figura 3). Nivelul de management al datelor este componenta principală a platformei, fiind responsabilă, prin modulul de colectare a datelor, de preluarea în cadrul platformei a datelor, accesibile doar prin intermediul unor portaluri web, fără nicio altă expunere structurată sau vreo altă modalitate de descărcare cu ajutorul unor script-uri de web de tip crawling, specifice fiecărei surse de date în parte. Identifică modele de date ce pot fi folosite în cadrul platformei. Conține și un modul de gestionare a bazei de date.

**Modelele de date** reprezintă componentele de bază ale aplicațiilor, având un impact major asupra performanței și capabilităților acestora. Mai mult decât atât, aceste modele de date pot fi impuse de anumite instrumente folosite în cadrul prelucrării lor.

Sunt prezentate în continuare diferitele modele de date care există în contextul Big Data: date structurate, date în fișiere text, date semi-structurate și date cheie-valoare.

Modelele de date structurate se referă la datele conținute în fișiere de baze de date sau foi de calcul. Sursele de date structurate într-o Cyber-Infrastructură sunt senzori de măsurare (de exemplu, parametrii de apă precum conductivitatea, salinitatea, temperatura etc.), sisteme de informații geografice (de exemplu, geo-baze de date care conțin date spațiale) și surse de flux de date (de exemplu, date generate de intervenția umană).

Modelele de date tip document se află la capătul opus al datelor structurate, deoarece acest tip de date nu are o structură și un sens bine definit. Aceste date se pot regăsi în diverse documente legate de reglementări emise de instituții de reglementare în domeniul infrastructurilor cibernetice, spre exemplu managementul apei.

Valorile în acest model de date sunt reprezentate de "documente" care au anumite structuri și modele de codificare (de exemplu, XML, JSON, BSON, etc.) al datelor gestionate.

Datele semi-structurate reprezintă date care au o structură, dar nu sunt relaționale. Aparatura de instrumentație, cum ar fi senzorii, generează aceste date. Pentru a fi stocate într-o bază de date relațională, datele trebuie transformate. Un avantaj major al datelor semi-structurate este reprezentat de faptul că pot fi încărcate direct într-un sistem de fișiere Hadoop HDFS și prelucrate în forma inițială.

Acest tip de date are o structură de auto-descriere care conține etichete pentru a separa elementele și a stabili ierarhii de înregistrări și câmpuri ale datelor. În general, acest tip de date sunt stocate în tipuri speciale de baze de date HBase fiind unul dintre cele mai cunoscute exemple.

Modelul de date cheie-valoare are următoarele caracteristici: structura simplă, viteza interogării mai mare decât în cazul bazei de date relaționale, suport pentru stocare în masă, interogarea și modificarea operațiilor pentru date prin cheia primară fiind acceptate cu succes.

Bazele de date cheie-valoare sunt de tip NoSql, în care valorile sunt asociate cu cheile reprezentate de șiruri de caractere. Operațiile de bază în cadrul bazelor de date NoSQL sunt:

- asocierea unei valori cu cheia corespunzătoare - Put(key, value);
- eliminarea tuturor valorilor asociate cu o cheie furnizată - Delete(key);
- returnarea valorilor pentru o cheie furnizată - Get(key);
- returnarea listei de valori pentru mai multe chei furnizate – MultiGet (key1, key2, ..., keyn).

Deși bazele de date NoSQL pot fi utilizate eficient pentru a stoca date rezultate din algoritmi, cum ar fi numărarea frazelor, acestea au și dezavantaje, cum ar fi: dificultatea generării de șiruri de caractere unic pentru chei. În al doilea rând, acest model nu oferă capacități, cum ar fi logica pentru mai multe tranzacții executate simultan, acestea fiind furnizate de nivelul aplicației.

Sistemul de gestiune a bazei de date este dezvoltat cu scopul de a oferi suport pentru administrarea unor seturi de date de diferite dimensiuni ce sunt utilizate în cadrul aplicației.

Asigurarea sistemului se face prin implementarea unei baze de date **MySQL** și utilizând un ORM - Object-Relational Mapping (de ex. soluția NHibernate a Microsoft) pentru a facilita legătura între obiectele din cadrul aplicației și tabele din baza de date.

A fost folosit MySQL pentru implementarea bazei de date deoarece oferă anumite avantaje, precum oferirea unui cadru distinct pentru un motor de stocare a cărui configurație este realizată de administratorul de sistem, asigurând o performanță impecabilă atât pentru o dimensiune mare a bazei de date cât și pentru un sistem de procesare a tranzacțiilor de mare viteză.

De asemenea, flexibilitatea la cerere este o caracteristică cheie a MySQL. Această soluție open source permite personalizarea completă a activităților efectuate și ajută la implementarea unor cerințe unice ale server-ului de baze de date. MySQL are un timp de funcționare continuu, de 24 de ore pe zi, în fiecare zi a săptămânii, oferă o gamă largă de soluții de disponibilitate, inclusiv servere cluster specializate și configurații master/slave. Acesta se află pe primul loc în lista motoarelor de stocare tranzacționale robuste de pe piață datorită suportului atomic, consistent și multi-versiune al tranzacțiilor precum și blocarea fără restricții la nivelul liniilor, fiind o soluție foarte bună pentru asigurarea integrității datelor, garantând identificarea rapidă a situațiilor blocante prin integritatea referențială impusă de server.

MySQL este proiectat pentru a satisface chiar și cele mai exigente aplicații, asigurând în același timp o viteză optimă, indexuri full-text și cache-uri de memorie unice pentru o performanță îmbunătățită [1].

O altă posibilă implementare a sistemului de gestiune a bazei de date este **Apache Cassandra**, o bază de date extrem de scalabilă, de înaltă performanță, proiectată să gestioneze cantități mari de date pe mai multe servere, oferind disponibilitate ridicată. Este un tip de bază de date NoSQL, care oferă un mecanism pentru stocarea și preluarea datelor diferit față de relațiile tabelare utilizate în bazele de date relaționale [14].

Cassandra a fost proiectată cu scopul de a gestiona încărcările mari de date din cadrul mai multor noduri fără niciun punct de eșec. Cassandra folosește un sistem distribuit peer-to-peer de-a lungul nodurilor sale, iar datele sunt distribuite între toate nodurile dintr-un cluster. Fiecare nod este independent și în același timp interconectat cu alte noduri având același rol și poate accepta cereri de citire și scriere, indiferent de locul în care datele sunt de fapt localizate în cluster. În momentul în care un nod este scos din rețea, cererile de citire/scriere pot fi difuzate de la alte noduri din rețea [8].

Modelul de date folosit în Cassandra este semnificativ diferit de cel dintr-un sistem de management al bazelor de date relaționale (RDBMS). Baza de date este distribuită pe mai multe mașini care funcționează împreună. Cel mai exterior container este cunoscut sub numele de Cluster.



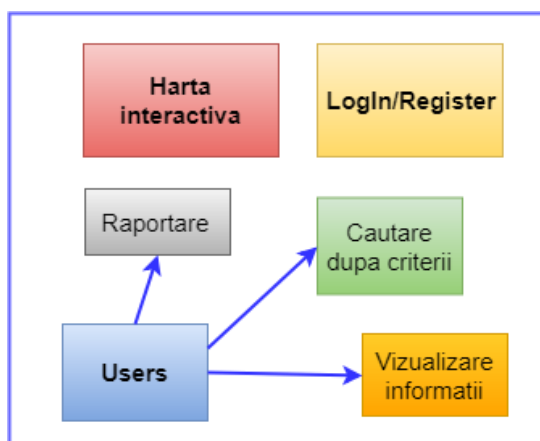
Pentru evitarea erorilor, fiecare nod conține o replică, iar în caz de eșec, replica îi va lua locul. Nodurile sunt aranjate într-un cluster, acesta are un format inelar și le atribuie date [16].

Arhitectura peer-to-peer, scalabilitatea elastică prin adăugarea sau ștergerea nodurilor, fără a fi necesară repornirea cluster-ului, modelul de date orientat pe coloane, valorificarea mai multor mașini pentru o performanță foarte bună sunt doar câteva din motivele pentru care Cassandra este folosită în numeroase aplicații și solicitată în rândul multor companii.

Conexiunea cu baza de date este realizată prin intermediul NHibernate, reprezentând un motor de persistență sub forma de framework, care încarcă informațiile din baza de date în obiecte la nivelul aplicației și salvează eventualele modificări ale acestor obiecte înapoi în bază. Traducerea datelor se poate realiza prin configurarea unor fișiere de mapare sau prin utilizarea unor atribute specifice peste clase și proprietățile acestora.

NHibernate se află printre cele mai populare ORM-uri, datorită faptului că este un descendent al Hibernate. ORM este foarte popular în mediul de dezvoltare Java. NHibernate acceptă o gamă de baze de date diferite. Orice bază de date relațională existentă poate fi accesată la NHibernate. Serverul SQL este baza de date primară acceptată, majoritatea dezvoltatorilor o utilizează, fiind de aceea și cea mai comună.

**Nivelul de vizualizare** oferă acces la datele disponibile public, fiind necesar ca aplicația să posede mai întâi capacitatea de a stoca date în propriul sistem de management, precum și de a oferi utilizatorilor opțiuni de bază de vizualizare sau cel puțin să gestioneze metadatele împreună cu linkuri către setul de date complet stocat pe alte servere (Figura 4).



**Figura 4.** Nivelul de vizualizare

Acest nivel urmărește oferirea unei experiențe plăcute utilizatorilor în folosirea aplicației prin instrumente de accesare, explorare și implicare în manipularea datelor: realizarea unui modul de căutare full-text, căutare într-o zonă geografică definită, filtrare și sortare a rezultatelor căutărilor, posibilitatea de vizualizare a datelor prin diagrame și hărți.

**Fluxul de lucru** reprezintă un ansamblu de pași care pot fi efectuați în cadrul platformei precum și succesiunea acestora în scopul obținerii rezultatelor propuse. În cadrul aplicației se pot identifica două fluxuri de lucru: pentru utilizatorii aplicației și pentru componenta de colectare a datelor.

**Fluxul de lucru pentru utilizatori** este reprezentat în diagrama din Figura 5. Utilizatorul intră pe pagina principală a aplicației, pe care poate parcurge informațiile oferite sau poate căuta informații după diferite criterii. Se trimite o cerere de interogare a server-ului de date și se așteaptă răspunsul, care în caz de succes va oferi informațiile căutate de către utilizator.

Aplicația oferă utilizatorului și posibilitatea de raportare evenimente, cu condiția ca utilizatorul să fie înregistrat și conectat în sesiunea curentă în sistem.

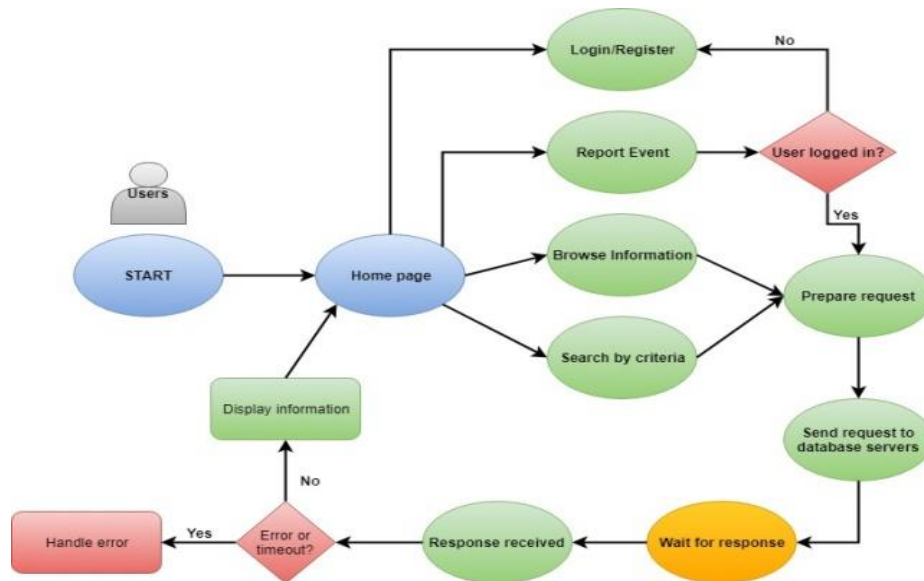


Figura 5. Fluxul de lucru utilizator

Fluxul de lucru pentru modulul de colectare date este prezentat în Figura 6 și arată modul în care sunt preluate datele de pe o sursă prin intermediul unui URL. Din pagina principală se va iniția preluarea datelor, se va trimite o solicitare către o pagină web, ale cărei informații vor fi preluate complet în cadrul aplicației, în cazul în care nu se va întâmpina o eroare. Pagina preluată va fi analizată și vor fi selectate din aceasta alte resurse care vor fi preluate ulterior. După preluarea tuturor paginilor disponibile și a resurselor pe care acestea le conțin, se completează baza de date cu informațiile colectate.

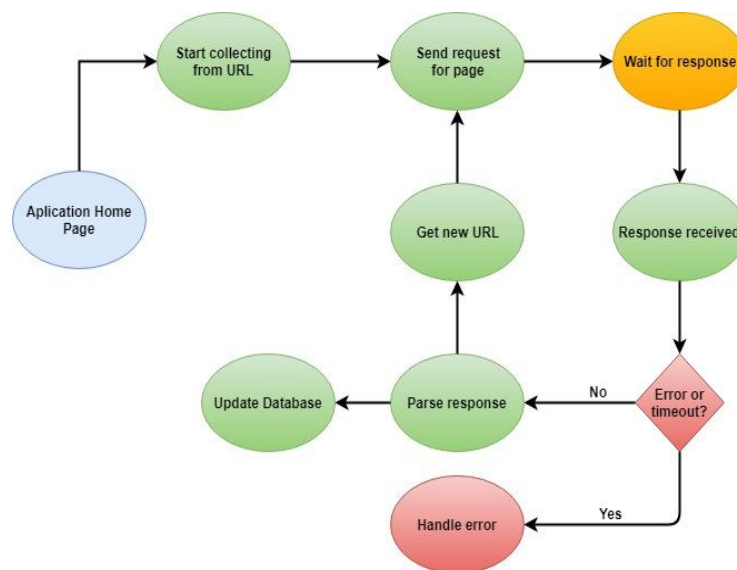


Figura 6. Fluxul de lucru colectare date

#### 4. Detalii de implementare

Aplicația s-a realizat folosind **paradigma MVC** (Model View Controller):

- Componenta Model conține obiectele asociate modelului de date;
- Componenta Controller asigură implementarea nivelului de management de date în care se apelează servicii de preluare a datelor;
- Componenta View asigură vizualizarea și prezentarea aplicației.

**Tehnologii folosite** în cadrul aplicației sunt:

- SQL 12.0 – baza de date de stocare a informațiilor;
- Nhibernate – conexiunea între nivelul de vizualizare și baza de date;
- HTML, CSS și JavaScript – pentru crearea paginilor web.

**Componenta de colectare.** Colectarea este un proces care se bazează pe aplicarea unor tehnici specifice necesare pentru obținerea datelor neprelucrate sau brute preluate din diferite surse. În aplicația propusă, componenta de colectare a datelor corespunzătoare nivelului de management al datelor este realizată printr-un algoritm care parsează HTML-ul paginilor web și preia informațiile necesare din acesta, astfel implementându-se printr-o metodă de Web Crawling. Avantajul acestei metode este acela de a oferi posibilitatea de preluare a tuturor informațiilor.

Metoda aplicată permite colectarea datelor prin intermediul unor instrumente auxiliare, fiind realizată în mod specific pentru preluarea de date din cadrul Nomenclatorului național al gestionării de date geospațiale <http://geoportal.gov.ro/Geocatalog/Organization/Index/>, prin preluarea informațiilor existente pe fiecare pagină a acestui site. Pagina este preluată în totalitate, întregul cod HTML care o descrie, și se realizează o analiză și identificare a informațiilor cu adevărat necesare și relevante pentru obiectivele propuse.

S-a folosit o clasă de tipul StreamReader pentru a realiza citirea fluxului de date, care a fost ulterior trimis către o funcție de parsare în care se vor prelua efectiv informațiile din acesta. Acest serviciu de preluare a datelor va fi rulat periodic în scopul de a actualiza informațiile oferite utilizatorilor. În figura 7 se prezintă o listare a codului specific modului de colectare a datelor.

```
public void GetDataFromURL(string URL)
{
    HttpRequest request = (HttpRequest)WebRequest.Create(URL);
    HttpResponse response = (HttpResponse)request.GetResponse();
    string responseString = null;

    if (response.StatusCode == HttpStatusCode.OK)
    {
        Stream receiveStream = response.GetResponseStream();
        StreamReader readStream = null;

        if (response.CharacterSet == null)
        {
            readStream = new StreamReader(receiveStream);
        }
        else
        {
            readStream = new StreamReader(receiveStream,
                Encoding.GetEncoding(response.CharacterSet));
        }

        responseString = readStream.ReadToEnd();

        response.Close();
        readStream.Close();
    }
    ParseResponse(responseString);
}
```

**Figura 7.** Modul de colectare date

**Componenta de agregare.** După finalizarea procesului de colectare a datelor, acestea vor fi transferate în cadrul unui sistem de stocare cu scopul de a fi ulterior procesate și analizate.

Un număr foarte de mare al surselor și seturilor de date oferite de modulul de colectare poate prezenta anumite inconveniențe, precum elemente nerelevante pentru obiectivul propus, care pot să producă o gestionare defectuoasă a spațiului de stocare, în sensul ocupării bazei de date cu informații inutile. În acest sens, este necesar ca etapa de vizualizare a aplicației să fie precedată de

implementarea unei componente de agregare care conduce la reducerea spațiului de stocare și la creșterea acurateții analizei datelor.

Componenta de agregare este realizată în nivelul de management al datelor, mai specific pe server-ul bazei de date. Aceasta este implementată prin folosirea unor proceduri stocate speciale, care se execută automat în cazul apariției unui eveniment pe serverul bazei de date (Funcții Trigger).

Aceste proceduri oferă execuția rapidă a unor tranzații de date rulate din cadrul aplicației. Componenta de agregare este responsabilă de procesarea datelor ce provin din mai multe surse, bazându-se pe combinarea informațiilor, reunirea acestora în cadrul unui singur model și prezentarea unei imagini unitare asupra seturilor de date. Astfel, se obțin informații îmbunătățite, mai puțin costisitoare, de o calitate superioară sau mai relevante în analiza și procesarea datelor. Unul dintre avantajele componente de agregare îl constituie, în principal, îmbunătățirea autenticității sau disponibilității datelor. Componenta de agregare are ca scop popularea unui model de date tabelare care conține seturi de date și caracteristici specifice seturilor de date, provenite de la un model ierarhic așa cum se prezintă în cadrul Nomenclatorului național al gestionărilor de date geospațiale. Acesta din urmă este organizat pe instituții care conțin seturi de date și servicii aferente.

**Componenta de vizualizare.** Componenta de vizualizare este realizată prin partea de View din Modelul MVC, prin fișiere de tip csHtml folosite de motorul ASP.NET pentru a construi paginile web ce vor fi afișate în browser.

Seturi de date disponibile Preluare seturi de date

Select denumire set date:  Select categorie tematica:  Select punct de contact:  Select institutie de gestionare:  Search

Denumire Set de Date	Categorie Tematica	Tip Resursa	Punct de Contact	Institutie Gestionare	Link Resursa Externa
Serviciu de descarcare Tema Geologie, Anexa II, scara 50k, IGR	-	Serviciu de date spațiale	Institutul Geologic al Romaniei	Institutul Geologic al României	<a href="#">Resurse Externe</a>
Serviciu de vizualizare Tema Geologie, Anexa II, scara 1M, IGR	-	Serviciu de date spațiale	Institutul Geologic al României	Institutul Geologic al României	<a href="#">Resurse Externe</a>
Serviciu de vizualizare Tema Geologie, Anexa II, scara 200k, IGR	-	Serviciu de date spațiale	Institutul Geologic al Romaniei	Institutul Geologic al României	<a href="#">Resurse Externe</a>
Serviciu de descarcare Tema Geologie, Anexa II, scara 1M, IGR	-	Serviciu de date spațiale	Institutul Geologic al Romaniei	Institutul Geologic al României	<a href="#">Resurse Externe</a>
Serviciul de vizualizare pentru tema Geologie 50k, anexa II, INSPIRE, IGR	-	Serviciu de date spațiale	Institutul Geologic al Romaniei	Institutul Geologic al României	<a href="#">Resurse Externe</a>

1 - 5 din 15 elemente 5 elemente per pagină Previous 1 2 Next

**Figura 8.** Funcționalități de bază – Vizualizare

Prin aceasta metodă se va oferi utilizatorilor o interfață stilizată și un aspect vizual plăcut prin folosirea directivelor CSS, precum și interactivă prin utilizarea unor scripturi JavaScript care ajută la obținerea efectului dinamic al paginilor.

Legătura dintre modelul de date și componenta de vizualizare este gestionată din cadrul aplicației prin folosirea framework-ului NHibernate pentru care s-au construit fișiere de configurare (hbm.xml) care permit manipularea tabelelor din baza de date cu ușurință și acuratețe. Fișierele de configurare au o structură strictă de implementare impusă de o serie de proprietăți XML.

Componenta de vizualizare oferă posibilitatea utilizatorului de a realiza diferite acțiuni printre care vizualizarea informațiilor legate de dezastru și crize majore în cadrul unui catalog de date prezentat sub forma unui tabel, căutarea după criterii alese de utilizatori astfel încât să vină în întâmpinarea necesităților acestora.

## 5. Rezultate experimentale

Sunt prezentate în continuare rezultatele experimentale obținute privind capabilitățile de vizualizare, funcționalitățile de colectare și raportare (Figurile 9 și 10), precum și performanțele obținute.

**Capabilități de vizualizare.** Experiența utilizatorului sau UX (User eXperience) constituie un concept capabil de a reda și măsura rezultatul perceput de utilizator în timpul unei utilizări temporare sau repetate a unui obiect funcțional sau a unei interfețe de interacțiune.

Interfața cu utilizatorul reprezintă, pe scurt, partea vizibilă a sistemului facilitând introducerea informațiilor, execuția unor comenzi și vizualizarea unor modele având rolul de a interacționa cu utilizatorul.

Un aspect important îl reprezintă designul interfeței care asigură ușurință de accesare a interfeței și de a introduce date în sistem, de a interpreta ieșirile din sistem precum și timpul necesar pentru ca utilizatorul să se familiarizeze cu interfața respectivă.

Pe platforma propusă se vor putea vizualiza informații sub formă tabelară, cu posibilitatea de selectare a numărului de intrări în tabel, de filtrare a elementelor în funcție de coloane și căutare full-text în cadrul informațiilor din tabel.

**Performanțele platformei.** Performanța unei platforme se traduce în flexibilitatea acesteia, consistența informațiilor oferite, simplitate și adaptabilitate în utilizare, operare rapidă și eficientă.

Performanța platformei propuse se poate evidenția prin cantitatea de date care poate fi stocată în baza de date, preluarea datelor dintr-un număr relativ mare de surse existente, timpul scurt de răspuns al interogărilor în căutarea datelor, acuratețea în prelucrarea și agregarea datelor, toleranța la defecte a sistemelor de gestiune a datelor.

Capabilitatea de a permite dezvoltări ulterioare prin preluarea informațiilor din alte surse de date și corelarea cu servicii externe, precum cele de raportare evenimente, reprezintă de asemenea o caracteristică de performanță a platformei propuse.

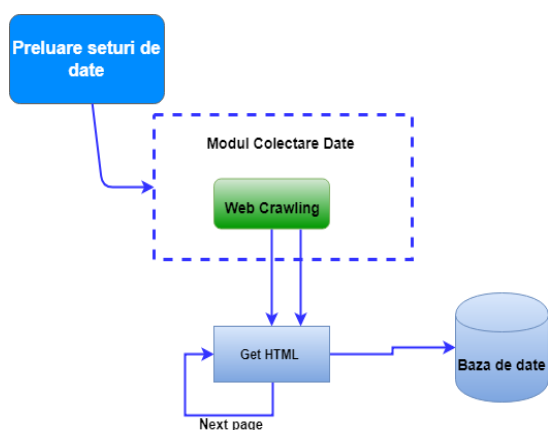


Figura 9. Funcționalități de bază – Colectare

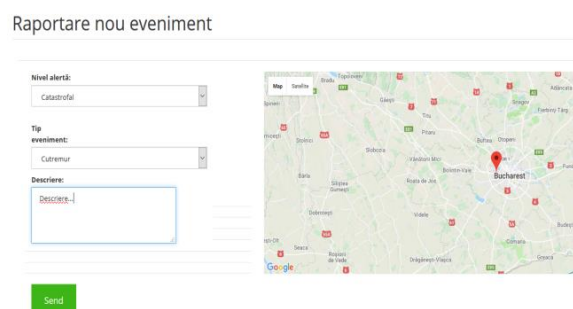


Figura 10. Funcționalități de bază – Raportare

**Detaliile experimentale** obținute sunt următoarele: timpul de preluare a tuturor datelor înregistrate în Nomenclatorul național al gestionărilor de date geospațiale este de aproximativ 2 minute și 20 de secunde, timp în care se vor parcurge 182 de pagini HTML și se vor introduce 78 de seturi de date din baze asociate unui număr de 38 de instituții. Timpul necesar aducerii celor 78 de seturi de date în browser este de aproximativ 8,7 secunde.

## 6. Concluzii

În lucrarea de față este prezentată o platformă caracterizată prin capacități de vizualizare precum un catalog de date, o hartă interactivă, o interfață care să răspundă cerințelor unui număr mare de utilizatori, precum și alte instrumente și servicii web standardizate. Spre deosebire de soluțiile existente în prezent, platforma propusă reunește un număr mare de surse de date și dezvoltă un serviciu specializat pentru gestionarea și arhivarea datelor, având o toleranță bună la defecte cu scopul de a îmbunătăți nivelul de cunoștințe privitoare la monitorizarea dezastrelor și crizelor majore. Platforma a fost dezvoltată utilizând tehnologiile C#, ASP.Net, MVC, HTML, MySQL, TypeScript, HTTP, XML. Au fost prezentate arhitectura platformei propuse, fluxul de lucru, funcționalitățile de bază și câteva detalii de implementare pentru gestionarea și colectarea datelor și rezultatele experimentale obținute. De asemenea, au fost descrise modelele de date care pot fi utilizate în scopul elaborării unei platforme de monitorizare a dezastrelor și crizelor majore, organismele de standardizare specializate în dezvoltarea de standarde pentru datele și serviciile cu caracter geospațial și care oferă specificații pentru serviciile de rețea. În ceea ce privește dezvoltările ulterioare, platforma web creată poate fi dezvoltată prin completarea bazei de date cu informații de actualitate, pot fi realizate anumite validări ale datelor înregistrate în sistem, pot fi integrate noi surse de date și, se poate realiza corelarea cu platforme și servicii externe. De asemenea, este posibilă realizarea unor algoritmi de predicție a evenimentelor pe baza datelor preluate.

## Mențiuni

Soluția propusă în acest articol face parte din proiectul SPERO – Tehnologii spațiale în managementul dezastrelor și crizelor majore, manifestate la nivel local, național și regional este finanțat de Ministerul Cercetării și Inovării, UEFISCDI, având codul PN-III-P2-2.1-SOL-2016-03-0046, contract numărul 3Sol/2017, în cadrul programului PNCDI III, perioada de derulare: Aprilie 2017 – Septembrie 2020.

## BIBLIOGRAFIE

1. Challawala, S. et al. (2017). *MySQL 8 for Big Data: Effective data processing with MySQL 8, Hadoop, NoSQL APIs, and other Big Data tools*;
2. Chang, F. et al. (2008). *Bigtable*. *ACM Transactions on Computer Systems*, 26(2), pp.1-26;
3. C, M., Mao, S. & Liu, Y. (2014). *Big Data: A survey*. *Mobile Networks and Applications*, 19(2), pp.17-1209;
4. Chodorow, K. & Dirolf, M. (2013). *MongoDB: The Definitive Guide: Powerful and Scalable Data Storage*;
5. Dumitrache, M., Sandu, I., Barbu, D. (2017). *An Integrated Cloud Computing Solution for Romanian Public-Sector Entities: ICIPRO Project*. *Studies in Informatics and Control* 26, no. 4: 481-487;
6. George, I. et al. (2017). *Service Level Agreement Characteristics of Monitoring Wireless Sensor Networks for Water Resource Management (SLAs4Water)*. *Studies in Informatics and Control* 26, no. 4: 379-386;
7. Japikse, T. et al. (2017). *"Introducing ASP. NET Web API."* *In Pro C# 7*, pp. 1223-1242. Apress, Berkeley, CA;
8. Khanh, D. et al. (2017). *Geospatial Database*;
9. Kristensen, E. K., Møller, A. (2017). *Type test scripts for TypeScript testing*. *Proceedings of the ACM on Programming Languages* 1, no. OOPSLA:90;
10. Lakshman, A. & Malik, P. (2009). *Cassandra: Structured storage system on a P2P network*. *In Proceedings of the 28th ACM Symposium on Principles of Distributed Computing - PODC '09*. The 28th ACM Symposium. New York, New York, USA: ACM Press, P. 5;



11. Nicolau, D. (2018). *Considerații asupra bazelor de date NoSQL*, Romanian Journal of Information Technology and Automatic Control - Revista Română de Informatică și Automatică, ISSN 1220-1758, vol. 28(3), pp. 53-62;
12. Nicolau, D. et al. (2017). *Servicii Cloud de versionare cu suport baze de date No SQL*, Romanian Journal of Information Technology and Automatic Control - Revista Română de Informatică și Automatică, ISSN 1220-1758, vol. 27(3), pp. 47-58, 2017;
13. Quattrochi, D. A. et al. (2017). *Integrating Scale in Remote Sensing and GIS*. CRC Press;
14. Saurabh, A. et al. (2018). "Working with Cassandra Database." In *Information and Decision Sciences*, pp. 531-538. Springer, Singapore;
15. Sinha, Sanjib (2017). "Model, View, Controller Workflow." In *Beginning Laravel*, pp. 59-61. Apress, Berkeley, CA;
16. Skylab, R., Pethuru, R. (2017). *Hosting and Delivering Cassandra NoSQL Database via Cloud Environments*. In *NoSQL*, pp. 429-448. Chapman and Hall/CRC.



**Marius Cristian CÎRȚINĂ** este student la master la Departamentul de Calculatoare și Tehnologia Informației la Universitatea *Politehnica* din București. Interesele sale generale de cercetare sunt: dezvoltarea de software, proiectarea bazelor de date, regăsirea informațiilor și agregarea datelor.

**Marius Cristian CÎRȚINĂ** is Master's Student at the Department of Computer Science and Engineering at the Politehnica University of Bucharest. His general research interests include: software development, database design, information retrieval and data aggregation.



**Dr. Florin POP** este profesor la Departamentul de Calculatoare și Tehnologia Informației la Universitatea *Politehnica* din București. De asemenea, lucrează ca și cercetător științific gradul I la ICI București. Interesele sale generale de cercetare sunt: sisteme distribuite pe scară largă (proiectare și performanță), grid computing și cloud computing, sisteme peer-to-peer, gestionarea Big Data, agregarea datelor, tehnici de regăsire și clasificare a informațiilor, metode de optimizare Bio-Inspired.

**Dr. Florin POP** is professor at the Department of Computer and Information Technology, the Politehnica University of Bucharest. He also works as a 1<sup>st</sup> degree scientific researcher at ICI Bucharest. His general research interests are: distributed systems (design and performance), grid computing and cloud computing, peer-to-peer systems, Big Data management, data aggregation, information retrieval and classification techniques, Bio-inspired optimization methods.



**Mădălina ZAMFIR** este cercetător științific în cadrul Departamentului „Sisteme inteligente distribuite, intensive ca date“ din ICI București. Este doctorand în domeniul suportului IoT pentru sisteme de afaceri, la Universitatea *Politehnica* București. Subiectele de interes în activitatea de cercetare acoperă infrastructurile de tip Cloud, suportul IoT pentru soluțiile de tip Big Data, data analytics, modele de afaceri pentru IoT și Big Data.

**Mădălina ZAMFIR** is a scientific researcher in the „Distributed and Data Intensive Intelligent Systems“ Department at ICI Bucharest. She is a PhD student in IoT support for Business Systems at the *Politehnica* University of Bucharest. Topics of interest in the research activity include Cloud infrastructures, IoT support for Big Data solutions, data analytics, business models for IoT and Big Data.



**Vladimir FLORIAN** este inginer dezvoltare tehnologică gradul I în Departamentul „Sisteme inteligente distribuite, intensive ca date“ din ICI București. A obținut titlul de doctor în Știința Calculatoarelor la Universitatea *Politehnica* București, în anul 2006. Domeniile sale de interes includ: arhitecturi și sisteme distribuite, BigData, Big Data Analytics, IoT.

**Vladimir FLORIAN** is a 1st degree technology development engineer in the „Distributed and Data Intensive Intelligent Systems“ Department at ICI Bucharest. He holds a PhD in Computer Science at the *Politehnica* University of Bucharest, in 2006. His topics of interest include: distributed architectures and systems, Big Data, Big Data Analytics, IOT.



**Gabriel NEAGU** este cercetător gradul I și director științific în ICI București. A obținut titlul de doctor în Informatică Aplicată la Universitatea *Politehnica* București, în anul 1998. Principalele domenii de interes pentru activitatea de cercetare includ: arhitecturi de sisteme distribuite, analiza avansată a datelor masive, servicii IoT-Cloud, tehnologia blockchain.

**Gabriel NEAGU** is 1st degree senior researcher and Scientific Director at ICI Bucharest. He received a PhD in Applied Informatics at the *Politehnica* University of Bucharest, in 1998. His main topics of interest for research activity include: distributed system architectures, data analytics, Cloud- IoT services, blockchain technology.