

**ACADEMIA DE STUDII ECONOMICE DIN BUCUREȘTI**  
Consiliul pentru Studii Universitare de Doctorat  
Școala Doctorală de Informatică Economică

**PROCESAREA ȘI ANALIZA DATELOR ÎN CONTEXTUL  
IOT ȘI BIG DATA**

**Dragoș – Cătălin G. Barbu**

**Conducător de doctorat: Prof. Univ. Dr. Adela Bâra**

București, 2024

# Cuprins

<a href="#">Rezumat</a> .....	i
<a href="#">Multumiri</a> .....	ii
<b>1. INTRODUCERE</b> .....	<b>1</b>
<b>1.1. CONTEXT ȘI MOTIVAȚIE</b> .....	<b>1</b>
1.1.1. <a href="#">Context</a> .....	1
1.1.2. <a href="#">Motivare</a> .....	2
1.1.3. <a href="#">Metodologie</a> .....	3
1.1.4. <a href="#">Semnificația cercetării</a> .....	3
1.1.5. <a href="#">Structura tezei</a> .....	4
1.1.6. <a href="#">Obiectivele cercetării</a> .....	5
<b>1.2. SEMNIFICAȚIA STUDIULUI</b> .....	<b>6</b>
<b>2. STADIUL ACTUAL AL CERCETĂRII</b> .....	<b>8</b>
<b>2.1. BIG DATA: VIZIUNE GENERALĂ ȘI APLICAȚII</b> .....	<b>8</b>
2.1.1. <a href="#">Introducere</a> .....	8
2.1.2. <a href="#">Definiția și caracteristicile Big Data</a> .....	9
2.1.3. <a href="#">Tehnici și tehnologii de analiză a datelor mari</a> .....	13
<b>2.2. ANALIZA INFORMAȚIILOR DIGITALE DE NATURĂ FISCALĂ</b> .....	<b>13</b>
2.2.1. <a href="#">Cercetări și studii anterioare</a> .....	14
2.2.2. <a href="#">Big Data Analytics în context românesc</a> .....	16
2.2.3. <a href="#">Studii de caz în analiza datelor fiscale</a> .....	17
2.2.4. <a href="#">Caracteristicile datelor fiscale din casele de marcat electronice</a> .....	18
<b>2.3. CICLUL DE VIAȚĂ AL DATELOR</b> .....	<b>20</b>
<b>2.4. STUDIU ȘI PLANIFICARE ASUPRA INFORMAȚIILOR FISCALE DIN ECR-URI</b> .....	<b>23</b>
<b>2.5. COLECTAREA INFORMAȚIILOR DIGITALE DE NATURĂ FISCALĂ</b> .....	<b>24</b>
2.5.1. <a href="#">IOT – INTERNETUL OBIECTELOR (INTERNET OF THINGS)</a> .....	25
<b>2.6. PREPROCESAREA ȘI CURĂȚAREA DATELOR</b> .....	<b>28</b>
<b>2.7. ANALIZA DATELOR</b> .....	<b>28</b>
2.7.1. <a href="#">Analiză în timp real și flux de date</a> .....	28
2.7.2. <a href="#">Analiza exploratorie a datelor</a> .....	29
2.7.3. <a href="#">Tendințe în analiza informațiilor de tip Big Data. Tehnici și metode de extragere a datelor</a> .....	30
2.7.4. <a href="#">Modelul de regresie simplă</a> .....	31
2.7.5. <a href="#">Validarea modelului de regresie</a> .....	32
2.7.6. <a href="#">Machine learning</a> .....	34

2.7.7.	<a href="#">Paradigme ale învățării nesupravegheate</a>	35
2.7.8.	<a href="#">Clusterizare</a>	38
2.7.9.	<a href="#">Estimarea densității</a>	38
2.8.	<a href="#">VIZUALIZARE ȘI RAPORTARE</a>	38
2.9.	<a href="#">CONFIDENȚIALITATE ȘI PROVOCĂRI DE SECURITATE</a>	39
2.10.	<a href="#">CONFIDENȚIALITATE ȘI CONSIDERAȚII ETICE</a>	39
2.11.	<a href="#">IDENTIFICAREA LACUNELOR DE CERCETARE</a>	40
2.12.	<a href="#">REZUMAT ȘI CONTRIBUȚII PERSONALE</a>	41
3.	<a href="#">METODOLOGIA</a>	43
3.1.	<a href="#">ETAPELE ȘI OBIECTIVELE METODOLOGIEI</a>	43
3.2.	<a href="#">COLECTAREA DATELOR</a>	45
3.2.1.	<a href="#">Sursele datelor</a>	45
3.2.2.	<a href="#">Procedura de colectare</a>	46
3.2.3.	<a href="#">Considerații tehnice și de securitate</a>	49
3.3.	<a href="#">PRELUCRAREA ȘI ANALIZA DATELOR</a>	49
3.3.1.	<a href="#">Curățarea datelor</a>	49
3.3.2.	<a href="#">Instrumente și tehnici de analiză</a>	50
3.3.3.	<a href="#">Interpretarea rezultatelor</a>	52
3.4.	<a href="#">LIMITĂRILE METODOLOGIEI</a>	53
3.4.1.	<a href="#">Limitările legate de colectarea datelor</a>	53
3.4.2.	<a href="#">Limitările legate de prelucrarea datelor</a>	53
3.4.3.	<a href="#">Limitări legate de analiza datelor</a>	53
3.4.4.	<a href="#">Limitări legate de interpretarea și utilizarea datelor</a>	54
3.4.5.	<a href="#">Limitările contextului românesc</a>	54
3.5.	<a href="#">EVALUAREA ȘI VALIDAREA METODOLOGIEI</a>	54
3.5.1.	<a href="#">Validitatea datelor</a>	54
3.5.2.	<a href="#">Fiabilitatea datelor</a>	55
3.6.	<a href="#">REZUMAT ȘI CONTRIBUȚII PERSONALE</a>	57
4.	<a href="#">ARHITECTURA ȘI SPECIFICAȚIILE DE PROIECTARE ALE PLATFORMEI DE COLECTARE ȘI ANALIZĂ A DATELOR</a>	59
4.1.	<a href="#">TIPURILE DE INFORMAȚII</a>	60
4.2.	<a href="#">TEHNOLOGII SI ARHITECTURI</a>	63
4.3.	<a href="#">Colectarea de date</a>	65
4.3.1.	<a href="#">Transmiterea datelor</a>	65
4.3.2.	<a href="#">Transmiterea fișierelor XML</a>	67
4.4.	<a href="#">Preprocesarea datelor</a>	67

<a href="#">4.5. INTEGRAREA DATELOR</a>	68
<a href="#">4.6. VALIDAREA ȘI PREZENTAREA DATELOR</a>	68
<a href="#">4.7. ANALIZA DATELOR</a>	69
<a href="#">4.8. PUBLICAREA DATELOR</a>	69
<a href="#">4.9. REZUMAT ȘI CONTRIBUTII PERSONALE</a>	69
<a href="#">5. IMPLEMENTAREA PLATFORMEI</a>	71
<a href="#">5.1. MODUL PRIMIRE DATE</a>	71
<a href="#">5.2. MODUL PREPROCESARE DATE</a>	72
<a href="#">5.3. MODUL INTEGRAREA DATELOR</a>	74
<a href="#">5.4. MODUL PREPARARE/ PREGĂTIRE A DATELOR</a>	76
<a href="#">5.5. MODUL ANALIZA DATELOR (DATA ANALYTICS)</a>	77
<a href="#">5.6. MODULUL PUBLICARE REZULTATE</a>	82
<a href="#">5.7. REZUMAT ȘI CONTRIBUTII PERSONALE</a>	84
<a href="#">Comentariile</a>	86
<a href="#">Concluziile</a>	87
<a href="#">CONTRIBUȚII ORIGINALE</a>	89
<a href="#">LISTA LUCRĂRILOR ORIGINALE</a>	89
<a href="#">Referințele bibliografice</a>	92
<a href="#">Lista abrevierilor și acronimelor</a>	100
<a href="#">Lista figurilor/ graficelor</a>	100
<a href="#">Lista tabelor</a>	100
<a href="#">Lista anexelor</a>	101

**Cuvinte cheie:** Tehnici Big Data, Cloud computing, Data aggregation, Data enrichment, Date digitale fiscale, Internet of Thing, Date semi-structurate, Data mining, Data Storage, Big Data processing, Big Data Analytics, Machine learning, Agregarea datelor, Case de marcat electronice

## **Sinteza lucrării:**

Teza de doctorat explorează utilizarea Big Data în analiza informațiilor fiscale colectate de la casele de marcat electronice (ECR-uri), având ca scop îmbunătățirea conformității fiscale, optimizarea veniturilor și sprijinirea formulării informate a politicilor fiscale. Lucrarea debutează cu o **Introducere** care prezintă contextul și motivația cercetării, subliniind semnificația utilizării Big Data pentru progresul cunoștințelor în domeniul fiscal și tehnologic. Obiectivele tezei sunt clar definite, vizând dezvoltarea unei platforme eficiente de colectare și analiză a datelor fiscale.

Capitolul **Stadiul actual al cercetării** oferă o viziune generală asupra conceptului de Big Data, tehnicilor și tehnologiilor de analiză a datelor mari, și detaliază aplicațiile acestora în contextul românesc al informațiilor fiscale. Sunt revizuite studii de caz și caracteristicile specifice datelor colectate de la ECR-uri, evidențiind provocările și oportunitățile existente. Capitolul include și o prezentare detaliată a ciclului de viață al datelor fiscale, de la colectare, preprocesare și curățare, până la analiză, vizualizare și raportare.

**Metodologia** descrisă în capitolul al treilea detaliază etapele de colectare, prelucrare și analiză a datelor, subliniind instrumentele și tehnicile utilizate, precum și limitările metodologice întâlnite, inclusiv cele legate de contextul românesc. Evaluarea și validarea metodologiei se bazează pe criterii de validitate și fiabilitate a datelor, asigurând un cadru robust pentru analiza informațiilor fiscale.

Capitolul **Arhitectura și specificațiile de proiectare ale platformei de colectare și analiză a datelor** explică în detaliu tipurile de informații gestionate și tehnologiile utilizate, inclusiv arhitectura modulară care facilitează colectarea, transmiterea, preprocesarea și integrarea datelor. Platforma este proiectată să gestioneze eficient fluxurile mari de date și să asigure validarea și prezentarea clară a informațiilor relevante.

În capitolul **Implementarea platformei**, se discută modul în care au fost dezvoltate și implementate modulele platformei, inclusiv cele de primire, preprocesare, integrare, analiză și publicare a rezultatelor. Fiecare modul este detaliat în termeni de funcționalitate și contribuții la fluxul general de analiză a datelor fiscale.

**Concluziile** sintetizează principalele realizări ale tezei, evidențiind contribuțiile originale, care includ dezvoltarea unei arhitecturi inovatoare pentru colectarea și analiza datelor fiscale și demonstrarea impactului Big Data în îmbunătățirea conformității fiscale și optimizarea veniturilor. De asemenea, sunt identificate lacunele de cercetare și se propun direcții viitoare pentru explorarea avansată a tehnologiilor Big Data în context fiscal. Lista lucrărilor originale reflectă angajamentul autorului în diseminarea rezultatelor cercetării și avansarea cunoștințelor în acest domeniu emergent.