



**Universitatea Tehnică a Moldovei**

# **APLICAȚIE PENTRU ANALIZA ȘI PROCESAREA DATELOR SEMI-STRUCTURATE**

**Student:**

**Cervac Petru**

**Conducător:**

**conf. univ. dr. Sudacevschi Viorica**

**Chișinău - 2020**

## ADNOTARE

**La teza de masterat cu tema „Aplicație pentru analiza și procesare datelor semi-structurate”  
a st. gr. CRI-191M, Cervac Petru**

**Cuvinte cheie:** Date semi-structurate, aplicație desktop, raport, laborator de testare, C#, WPF

Teza de masterat are drept scop dezvoltarea unei aplicații desktop pentru analiza și procesarea datelor semi-structurate generate de către laboratorul unui producător de conectoare electrice. Teza expune arhitectura aplicației și descrie structura detaliată a părților componente ale aplicației.

**Tehnologiile utilizate sunt:** limbajul de programare C# pentru dezvoltarea aplicației, framework-ul pentru dezvoltarea aplicațiilor grafice framework-ul Windows Presentation Foundation pentru dezvoltarea interfeței grafice de utilizator, mediul de dezvoltare integrat Visual Studio 2019 pentru dezvoltarea aplicației, biblioteca DevExpress.Reporting pentru generarea de rapoarte.

**Memoriul explicativ** conține introducere, 3 capitole, concluzii, bibliografie cu 17 titluri, dintre care 56 pagini text de bază, 22 figuri și 4 anexe.

**Capitolul I** descrie domeniile datelor semi-structurate și literaturii gri, precum și prezintă provocările întâlnite

**Capitolul II** descrie totalitatea de instrumente necesare pentru implementarea proiectului precum limbajele de programare și mediile de dezvoltare utilizate.

**Capitolul III** descrie în detalii arhitectura aplicației dezvoltate și modulele din componența acesteia

## ANNOTATION

**For the thesis „Application for analysis and processing of semi structured data”**

**Keywords:** semi-structured data, Desktop application, report, C#, WPF

The master’s thesis aims to develop a desktop application for the analysis and processing of semi-structured data generated but the laboratory of a manufacturer of electrical connectors. The thesis exposes the application’s architecture and describes the structure of the component parts of the application in details.

**The technologies used are:** C# programming language for the development of the application, Window Presentation Foundation framework for graphical user interface development, DevExpress.Reporting library for report generation.

**Explanatory memo** contains an introduction, 3 chapters, conclusions, 56 pages of basic text, 22 figures and 4 annexes.

**Chapter I** describes the areas of semi-structured data and gray literature, as well as the challenges encountered.

**Chapter II** describes all the tools used to implement the project such as used programming languages and development environments.

**Chapter III** describes in detail the application’s architecture and its modules.

## Cuprins

LISTA FIGURILOR .....	10
INTRODUCERE .....	11
CAPITOLUL I CADRUL TEORETIC .....	12
1.1 Era informațională .....	12
1.1.1 Date, informații, cunoștințe și înțelepciune .....	12
1.1.2 Tipuri de date .....	13
1.2 Literatura Gri .....	14
1.2.1 Regulamente .....	15
1.2.2 Provocări .....	16
1.3 Studiu de caz .....	16
1.3.1 Provocări .....	17
1.4 Soluții .....	17
1.4.1 Ignorarea problemelor .....	17
1.4.2 Expert .....	17
1.4.3 Instruirea personalului .....	18
1.5 Automatizarea procesului de generare de rapoarte .....	18
1.6 Automatizarea generării de rapoarte în compania X .....	19
1.7 Analiza soluțiilor existente .....	20
1.7.1 Windwar Reporting .....	20
1.7.2 Telerik Reporting .....	20
1.1.1 Cristal Reports .....	20
1.7.3 Microsoft Word .....	20
1.7.4 LaTeX .....	21
1.7.5 DITA .....	21
CAPITOLUL II INSTRUMENTELE UTILIZATE .....	22
2.1 C# .....	22

2.2 NET Framework .....	22
2.3 Windows Presentation Framework(WPF) .....	23
2.3.1 Extensible Application Markup Language .....	23
2.3.2 Model View ViewModel .....	23
2.4 DevExpress Reporting .....	24
2.5 Domain Driven Design .....	25
CAPITOLUL III REZULTATELE CERCETĂRII .....	27
3.1 Arhitectura aplicației.....	29
3.1.1 Arhitectura de tip ceapă .....	30
3.1.2 Administrarea dependențelor dintre module.....	31
3.2 Modelarea domeniului de lucru .....	32
3.2.1 Rezultatele obținute de la W 434 .....	35
3.2.2 Rezultatele obținute de la VNA .....	36
3.2.3 Validarea măsurărilor.....	37
3.3 Subsistemele aplicației.....	40
3.3.1 Subsistemul de stocare.....	40
3.3.2 Interogarea datelor din depozit .....	41
3.3.3 Convertirea fișierelor în obiecte de domeniu.....	42
3.3.4 Subsistemul de raportare.....	43
3.4 Interfața grafică.....	43
3.4.1 Fereastra de încărcare a fișierelor .....	44
3.4.2 Fereastra de configurare a raportului .....	46
3.4.3 Fereastra de vizualizare a raportului generat .....	46
3.4.4 MVVM.....	47
3.4.5 Înregistrarea resurselor.....	49
3.4.6 Subsistemul de statusuri.....	49
CONCLUZII .....	51
BIBLIOGRAFIE .....	53
ANEXA 1 .....	55
ANEXA 2.....	58
ANEXA 3 .....	59
ANEXA 4.....	60

## INTRODUCERE

Începând cu mijlocul secolul 20, omenirea a experimentat o dezvoltare spectaculoasă a tehnologiilor informaționale. Această perioadă este numită epoca digitală. În această perioadă resursele cheie au devenit datele. Datele stau la baza informațiilor. Informațiile permit obținerea de cunoștințe. Până în 2018, s-a estimat că omenirea a produs peste 18 Zetabyte-ti de date, cu un trend pozitiv. Cantitatea enormă de date reprezintă o oportunitate ce merită a fi exploatată. O piedică în valorificarea datelor ce ne înconjoară reprezintă modalitățile de prelucrare a acestora.

În cazul în care datele au o structură clar definită acestea sunt numite date structurate. Calculatorul poate lucra ușor cu acest tip de date. Algoritmii pentru procesarea datelor structurate dezvoltati în ultimii 60 de ani oferă o performanță impresionantă. Din nefericire, majoritatea datelor ce ne înconjoară nu sunt date structurate. Ele sunt date nestructurate, care nu au o structură clar definită. Lucrul cu datele nestructurate este îngreunat. Datele nestructurate sunt mai greu de păstrat și prelucrat. Acestea nu pot fi păstrate într-o bază de date. Domeniul de prelucrare al datelor nestructurate este în continuă dezvoltare. Dezvoltarea unor metode pentru procesarea și analiza a datelor devine tot mai stringentă.

Printre cei mai mari beneficiari ai cantității enorme de date ce ne înconjoară sunt companiile. În realitățile actuale, succesul companiei depinde de capacitatea ei să valorifice datele ce îi stau la dispoziție. Ignorarea acestui lucru reprezintă un risc mult prea mare pentru viitorul acesteia.

Scopul proiectului constă în elaborarea unei aplicații desktop pentru analiza și procesarea datelor semi-structurate generate de către o un producător de conectoare electrice. Aplicația dezvoltată va permite valorificarea datelor obținute în urma testării conectorilor produse. Aplicația va permite colectarea datelor generate, extragerea informațiilor utile din acestea și afișarea informației sub formă de raport.

Aplicația dezvoltată va fi utilizată de laboratorul unui producător de conectoare electrice pentru valorificarea datelor obținute în urma testării componentelor. Datele semi structurate sunt utilizate pentru a trage concluzii cu privire la calitatea conectorilor produse de către companie. Această informație are un impact direct asupra afacerii companiei deoarece rapoartele vor fi utilizate de către consumatori pentru a lua decizia dacă de procurat componentele oferite de companie sau nu.

În prezent procesul de convertire a datelor inițiale în informații utile are loc manual. Aplicația are ca scop automatizarea procesului.

## BIBLIOGRAFIE

- [1] Economist, „Regulating the internet giants,” The Economist Newspaper Limited, 06 05 2017. [Interactiv]. Available: <https://www.economist.com/leaders/2017/05/06/the-worlds-most-valuable-resource-is-no-longer-oil-but-data>. [Accesat 29 11 2020].
- [2] P. Buneman, „Semistructured data,” în *Proceedings of the sixteenth ACM SIGACT-SIGMOD-SIGART symposium on Principles of database systems*, Tucson Arizona USA, 1997.
- [3] S. Abiteboul, „Querying Semi-Structured Data,” în *International Conference on Database Theory*, Delphi, Grecia, 1997.
- [4] J. Feather și P. Struges, „International Encyclopedia of Information and Library Science,” 2nd ed., New York, Routledge, 2013, p. 724.
- [5] C. P. Auger, „Use of Reports Literature,” London, Butterworth, 1975.
- [6] National Gray Literature Collection, „U.S. Interagency Gray Literature Working Group Definition 1995,” National Gray Literature Collection, 1997. [Interactiv]. Available: <http://allcatsrgrey.org.uk/wp/knowledgebase/u-s-interagency-gray-literature-working-group-definition-1995/>.
- [7] National Gray Literature Collection, „Luxembourg Definition 1997,” National Gray Literature Collection, 1997. [Interactiv]. Available: <http://allcatsrgrey.org.uk/wp/knowledgebase/luxembourg-definition-1997/>.
- [8] National Gray Literature Collection, „New York Definition 2004,” National Gray Literature Collection, 2004. [Interactiv]. Available: <http://allcatsrgrey.org.uk/wp/knowledgebase/new-york-definition-2004/>.
- [9] J. Albahari și B. Albahari, *C# 6.0 in a Nutshell*, 6th ed., Sebastopol, California: O'reilly, 2015, p. 1114.
- [10] J. Gossman, „Introduction to Model/View/ViewModel pattern for building WPF apps,” Microsoft Corporation, 08 10 2005. [Interactiv]. Available: <https://docs.microsoft.com/de-de/archive/blogs/johngossman/introduction-to-modelviewviewmodel-pattern-for-building-wpf-apps>.
- [11] M. Fowler, „Presentation Model,” 19 07 2004. [Interactiv]. Available: <https://martinfowler.com/eaaDev/PresentationModel.html>. [Accesat 15 11 2020].

- [12] Developer Express Inc, „Reporting,” Developer Express Inc, 26 10 2020. [Interactiv]. Available: <https://docs.devexpress.com/XtraReports/2162/reporting>.
- [13] E. Evans, Domain Driven Design: Tackling Complexity in the Heart of Software, Boston: Addison Wesley, 2003, p. 529.
- [14] J. Palermo , „Onion Architecture,” Clear Measure Inc., 29 07 2008. [Interactiv]. Available: <https://jeffreypalermo.com/2008/07/the-onion-architecture-part-1/>. [Accesat 15 09 2020].
- [15] M. Seemann, Dependency Injection Principles, Practices, and Patterns, 1st ed., New York: Manning Publications, 2019, p. 552.
- [16] G. Fairbanks, Just Enough Software Architecture, Boulder, Colorado: MARSHALL & BRAINERD, 2010, p. 360.
- [17] M. Fowler, Patterns of Enterprise Application Architecture, 1st ed., Boston: Addison-Wesley Professional, 2002, p. 560 .
- [18] D. Reinsel , J. Gantz și J. Ryding, „The Digitization of the World,” International Data Corporation, Framingham, USA, 2018.