

Prezentarea tezei

Tema tezei de doctorat o reprezintă generarea recomandărilor medicale computerizate prin dezvoltarea unor metode de achiziție a datelor pacienților și a regulilor medicale. Scopul final este mărirea gradului de interoperabilitate a sistemelor de generare a recomandărilor medicale.

Analiza domeniului a relevat existența unei multitudini de abordărilor pentru generarea de recomandări medicale (*Asbru*, *PROforma*, *Glif*, *Arden*, *Protege*, *Degel*), însă toate aceste soluții prezintă un grad scăzut de interoperabilitate, puține din acestea implementând standardele existente în domeniul informaticii medicale. Un alt dezavantaj al abordărilor existente în literatură îl reprezintă lipsa de metode pentru descoperirea de reguli medicale care să poată fi folosite mai apoi de mai multe sisteme de asistare a deciziei, regulile medicale computerizate bazându-se strict pe formalizarea ghidurilor medicale narative.

Astfel, pe baza unei analize a domeniului au fost definite următoarele **obiective**:

- A. dezvoltarea unei arhitecturi *CDS* care permite integrarea de surse de date multiple utilizând standardele *HL7 (CDA, vMR)*;
- B. descrierea și dezvoltarea de metode care permit reprezentarea standardului *vMR* (standard special creat pentru reprezentarea datelor pentru sistemele *CDS*) cu ajutorul tehnologiei *Topic Maps*;
- C. descrierea și dezvoltarea de metode pentru generarea de noi reguli medicale cu ajutorul tehnologiei *Data Mining* și reprezentarea acestora cu ajutorul standardului *Arden Syntax*;
- D. creșterea interoperabilității *CDS* prin achiziționarea de reguli medicale într-un format acceptat ca standard din diferite formalisme care nu sunt recunoscute ca standarde;
- E. evaluare soluțiilor propuse.

Actualitatea temei de cercetare abordate constă în faptul că domeniul îngrijirii sănătății este foarte important și afectează întreaga populație și de aceea orice demers de îmbunătățire a actului de îngrijire este bine venit. Astfel, sistemele de asistare a deciziei medicale (*Clinical Decision Support, CDS*) devin o parte importantă și indispensabilă îngrijirii medicale, datorită potențialului acestora de a îmbunătăți calitatea actului medical și de a reduce costurile. Acestea sunt din ce în ce mai folosite în diverse arii și tot mai multe instituții au început să le dezvolte, sistemele de generare a recomandărilor medicale reprezentând o tematică de actualitate care oferă multe direcții de cercetare. Pe de altă parte, din studiul bibliografiei rezultă că diverse colective de cercetători din toată lumea abordează în prezent acest domeniu, cu unele din aceste colective autorul având strânse relații de colaborare.

Conținutul tezei

Capitolul 1 - În acest capitol sunt prezentate câteva dintre cele mai actuale provocări ce apar în dezvoltarea sistemelor de asistare a deciziilor clinice (*CDS-Clinical Decision Support*). Este prezentat contextul medical referitor la calitatea asistenței medicale în ceea ce privește *sistemele de asistare a deciziilor clinice*. Sunt identificate diferite probleme cu privire la trei arii ale acestor sisteme:

- achiziționarea datelor specifice pacientului;
- reprezentare cunoștințelor medicale;
- descoperirea de noi reguli.

Sunt de asemenea puse în discuție mai multe soluții cu scopul de a rezolva problemele existente, având ca rezultat îmbunătățirea sistemelor de asistare a deciziilor clinice.

Pentru a genera recomandări medicale sunt necesare reguli medicale și date ale pacientului. Implementarea unui sistem *CDS* (dezvoltat într-o anumită unitate medicală) în unități medicale diferite întâmpină mari probleme legate de achiziționarea datelor pacientului deoarece, de obicei, soluțiile *CDS* sunt dezvoltate în jurul unei anumite structuri de date. Pentru a rezolva această problemă s-a stabilit ca direcție de cercetare utilizarea standardelor *HL7* pentru achiziția datelor pacienților.

În ceea ce privește regulile medicale, o problema majoră este utilizarea de formalisme multiple pentru reprezentarea regulilor medicale, ceea ce face ca folosirea de reguli medicale în sisteme *CDS* diferite să fie practic imposibilă. Pentru a rezolva aceasta problemă s-a stabilit ca direcție de cercetare reprezentarea cu ajutorul unui formalism acceptat ca standard (*Arden Syntax*) a regulilor medicale reprezentate în alte formalisme (ex. formalismul *ASTI*).

Utilizarea ghidurilor medicale nu poate acoperi toate situațiile posibile astfel încât s-a stabilit o a treia direcție de cercetare și anume descoperirea de reguli medicale și reprezentarea acestora cu ajutorul standardului *Arden Syntax*, pentru a permite utilizarea acestor reguli în cât mai multe sisteme *CDS* posibile.

Capitolul 2 - Este prezentată o analiză a câtorva din cele mai importate soluții existente în domeniu: *Glif*, *PROforma*, *Asbru* și *Gello*. Aceste abordări sunt prezentate din punct de vedere al modelelor concepționale, reprezentarea computerizată, achiziționarea datelor pacientului și vizualizarea recomandărilor. Soluțiile existente sunt apoi comparate pe baza a 5 criterii:

- acceptarea ca intrări a datelor pacientului reprezentate cu ajutorul unor standarde;
- reprezentarea recomandărilor medicale generate (ieșirile sistemului) cu ajutorul unor standarde;
- utilizarea ca formalism de reprezentarea a tehnologiei "task network";
- utilizarea ca formalism de reprezentarea a tehnologiei "Rule-Based";
- implementarea unor modalități de descoperire a regulilor medicale.

Pe baza acestei analize s-a observat că cel mai des utilizat standard pentru reprezentarea recomandărilor medicale este *Arden Syntax*, iar ca standard de reprezentare a datelor pacientului se utilizează standardul *HL7 CDA*. Dintre soluțiile analizate, Egadss "open source" utilizează cele mai multe standarde.

De asemenea, s-a realizat o analiză a evoluției în timp a soluțiilor existente. În Figura 1 este prezentată evoluția în timp a acestor sisteme.

Sunt prezentate câteva din cele mai cunoscute standarde pentru reprezentarea datelor pacienților: *HL7 CDA* (Clinical Document Architecture), *HL7 vMR* (virtual Medical Record), *HL7 CCD* (Continuity of Care Record).

Este evidențiată partea de vizualizare a recomandărilor medicale (la momentul creării și al execuției) prezentându-se diferite metode și software-uri care să permită vizualizarea ghidurilor și protocoalelor medicale la diverse etape: concepere, testare, execuție. Este dezvoltată o metodă originală de evaluare a calității interfeței cu utilizatorul, bazată pe aplicarea de chestionare personalului medical referitoare la diferite caracteristici ale interfețelor *CDS*. Au rezultat 5 caracteristici pe care ar trebui să le aibă interfețele sistemelor *CDS*:

- posibilitatea de a vizualiza pe aceeași interfață a recomandările medicale precum și a datelor referitoare la pacient;
- avertismente cu privire la tratament;
- afișarea de alarme;

- variante multiple de recomandări;
- posibilitatea de introducere a unor păreri vizavi de recomandări.

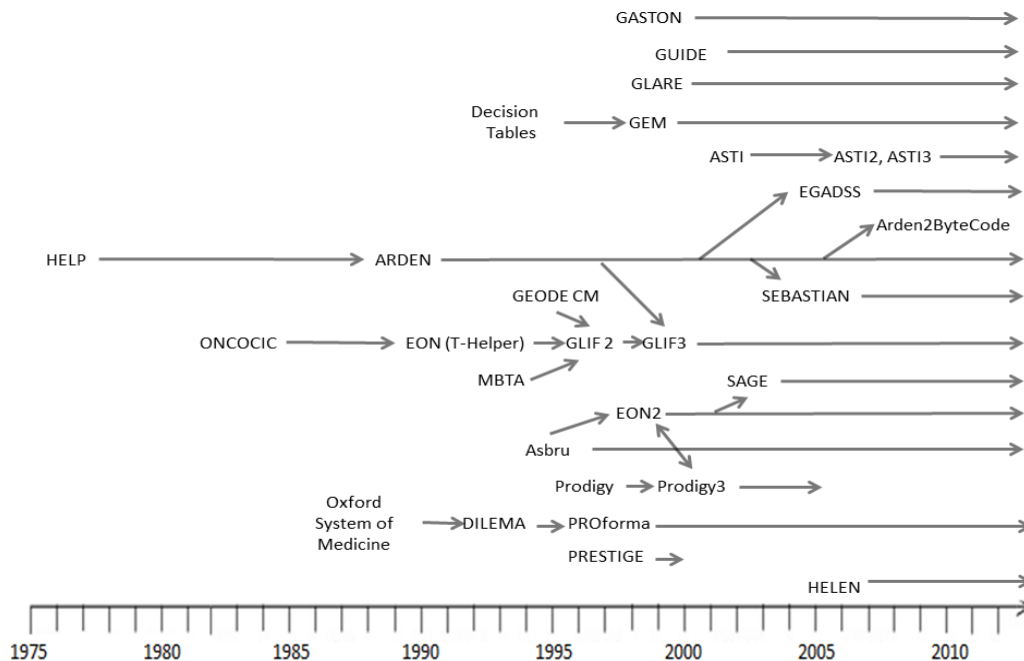


Figura 1. Istoria evoluției diferitelor abordări în domeniu

Capitolul 3 - Conține prezentarea unei arhitecturi CDS (Figura 2), care este dezvoltată pentru a permite accesul la date și cunoștințe medicale complexe. Pentru dezvoltarea acestei arhitecturi au fost luate în considerare mai multe standarde: VMR HL7, HL7 CDA, Arden Syntax. Arhitectura are ca principale componente:

- *Data Manager* - utilizat pentru a conecta toate celelalte componente. De asemenea au fost dezvoltate servicii care concatenează date din mai multe documente HL7 CDA și date în format vMR pentru a le putea transmite mai departe la partea de inferență (*Egads*).
- *Egads* - este o soluție „open source” pentru suport decizional clinic, care are ca intrări mesaje HL7 CDA sub formă de fișiere XML, reprezentând datele pacientului, și fișiere Arden Syntax, conținând reguli medicale. Pe baza acestora se generează recomandări medicale pe care le transpune în format HL7 CDA L2. A fost dezvoltat un serviciu care permite conectarea *Egads* la *Data Manager*, putându-se astfel accesa mai multe surse de date. Tot prin acest serviciu se transmit către *Data Manager* și recomandările generate.
- *HL7 CDA Component* - extrage date dintr-o bază de date locală și le prezintă ca mesaj HL7 CDA în format XML apoi le transmite la *Data Manager*.
- *TM- vMR* - este o bază de cunoștințe în care datele sunt reprezentate în format vMR cu ajutorul tehnologiei *Topic Maps*. Au fost dezvoltate servicii care permit extragerea datelor pacientului din *TM-vMR* și transmite aceste date către *Data Manager*.
- interfață pentru reprezentarea recomandărilor medicală și a datelor pacienților.

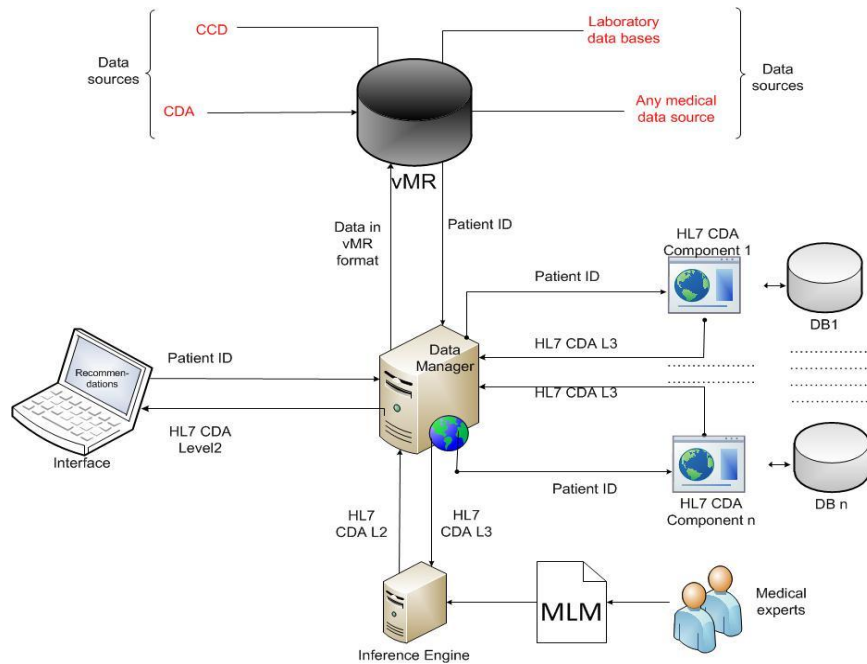


Figura 2. Arhitectura sistemului

Soluția a fost evaluată și validată pentru managementul diabetului din Spitalul Clinic Județean de Urgență Timișoara utilizându-se 30 de seturi de date. Seturile de date referitoare la un pacient conțin valori pentru: glicemie, rezervă alcalină, pH, sodiu, potasiu, greutate, uree, corpi cetonici. Aceste date au fost apoi reprezentate în format *HL7 CDA* și apoi în format *vMR*, pentru a testa validitatea *HL7 CDA Component* și *TM-vMR*.

Regulile medicale pentru managementul diabetului au fost reprezentate cu ajutorul formalismului *Arden Syntax* (Figura 3).

```

if nac < 149 and glicemia>50 then // NaCl / glucoza
NaClp:= "9 la mie";
  if h=1 or h=0 then NaCl:=1000; endif;
  if h>1 and h<6 then NaCl:= 500;endif;
  if h>5 then NaCl:= 250; endif;
endif;

if nac < 149 and glicemia<50 then // NaCl / glucoza
glucozap:="5 la mie";
  if h=1 or h=0 then glucoza:="1000, 5 la mie"; endif;
  if h>1 and h<6 then glucoza:= 500;endif;
  if h>5 then glucoza:= 250; endif;
endif;

if nac > 150 and glicemia<250 then // NaCl / glucoza
glucozap:="5";
  if h=1 or h=0 then glucoza:="1000, 5 la mie"; endif;
  if h>1 and h<6 then glucoza:= "500, 5 la mie";endif;
  if h>5 then glucoza:= 250; endif;
endif;

if nac > 150 and glicemia>249 then // NaCl / glucoza
NaClp:= "4.5 la mie";
  if h=1 or h=0 then NaCl:=1000; endif;
  if h>1 and h<6 then NaCl:= 500;endif;
  if h>5 then NaCl:= 250; endif;
endif;

```

Figura 3. Reprezentarea regulilor medicale în format Arden Syntax

Astfel, având datele pacientului și regulile medicale reprezentate, introducându-se în interfață ID-ul unui pacient, s-au generat recomandări medicale (Figura 4) care au fost confirmate ca valide de către personalul medical.



Figura 4. Recomandări medicale pe baza managementului diabetului.

Capitolul 4 - Prezintă o soluție pentru situațiile în care ghidurile medicale narative computerizate nu pot oferi recomandări medicale, fiind necesare reguli medicale (care nu sunt bazate pe dovezi), în scopul de a oferi un sfat medical în locul recomandării. Această soluție se bazează pe utilizarea tehnologiei *data mining* pentru extragerea regulilor medicale din baze de date.

Pentru a spori utilitatea acestor noi reguli, este dezvoltat un convertor care permite reprezentarea regulilor în limbajul *Arden Syntax*. Acest convertor traduce regulile medicale obținute cu ajutorul tehnicilor *data mining* în formatul *Arden Syntax*, care este recunoscut ca un formalism standard. Acest modul este numit *DB2AS*.

DB2AS are ca intrare regulile extrase cu ajutorul *data mining*, reprezentate în format *XML*. Ieșirile modulului *DB2AS* sunt fișiere *Arden Syntax*.

Etapele pentru generarea de *MLM*-uri (fișiere *Arden Syntax*) sunt:

- extragerea și crearea de categorii (*variable*, *value*, *operator*, *conclusion_variable*, *conclusion_value*) pe baza regulilor extrase (Figura 5). Aceste categorii de date sunt utilizate pentru generarea diferitelor părți ale unui *MLM*;
- generarea părților *Library* și *Maintenance*;
- generarea părții *Knowledge*.

```
IF background=U AND para=(-inf-2,6] AND gestation_weeks=(34.3-39.4] AND episio=0 THEN apgar=(8-inf]
```

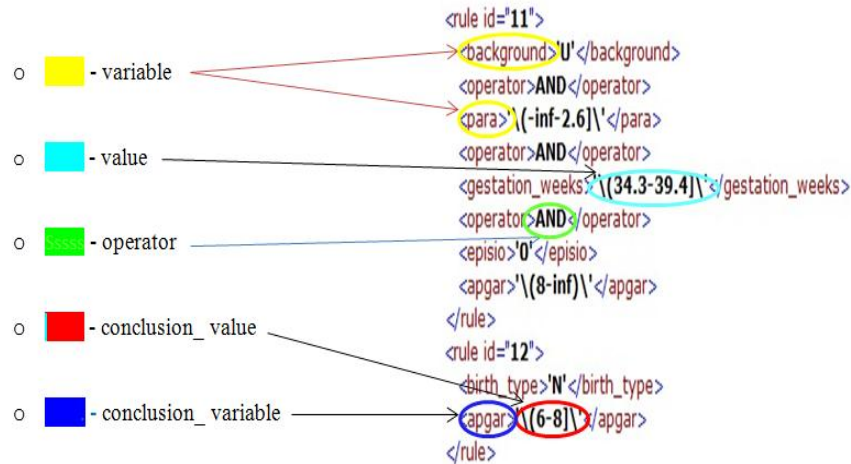


Figura 5. Creare de categorii (*variable*, *value*, *operator*, *conclusion_variable*, *conclusion_value*) pe baza regulilor extrase

Implementarea DB2AS se face folosind platforma *.NET* cu *C#*.

Acest sistem a fost testat folosind datele privind nou-născuții de la *Clinica de Obstetrică - Ginecologie Bega, Timișoara*. Pe baza acestor date și a sistemului propus s-au putut genera sfaturi medicale cu privire la: scorul Apgar, riscul la naștere și cezariană. Pentru a testa noile reguli s-au elaborat scenarii pe baza datelor din baza de date a nou-născuților de la *Clinica de Obstetrică - Ginecologie Bega, Timișoara*. Sfaturile medicale generate au fost conforme cu așteptările din aceste scenarii. În Figura 6 se poate observa modul în care se ajunge de la reguli extrase cu ajutorul *data mining* la reguli în format *Arden Syntax* apoi la sfaturi medicale.

```
IF luna=1 AND varsta='(31.8-38.4]' AND mediul=U AND G=9 OR G=10 AND P=2 AND prezentatie=CEF AND ia=8 AND sex=F AND tip_nastere=N AND videx=0 AND epizio=1 AND emp=0 THEN risc=2
IF mediul=U AND G=1 AND P=1 AND prezentatie=CEF AND ia=10 AND sex=M AND tip_nastere=C AND videx=0 AND epizio=0 AND emp=0 THEN risc=1
IF mediul=U AND G=1 AND P=1 AND prezentatie=CEF AND ia=10 AND tip_nastere=N AND videx=0 AND epizio=1 AND emp=0 THEN risc=0
```



```
IF (mounth=1) AND (age>31.8 AND age<38.4 or age=38.4) AND (background=U) AND (G=9) OR (G=10) AND (P=2) AND (presentation=CEF) AND (ia=8) AND (sex=F) AND (birth_type=N) AND (videx=0) AND (epizio=1) AND (emp=0) THEN solutie:="risc=2";
Conclude true; endif;
IF (background=U) AND (G=1) AND (P=1) AND (presentation=CEF) AND (ia=10) AND (sex=M) AND (birth_type=C) AND (videx=0) AND (epizio=0) AND (emp=0) THEN solutie:="risc=1";
Conclude true; endif;
IF (background=U) AND (G=1) AND (P=1) AND (presentation=CEF) AND (ia=10) AND (birth_type=N) AND (videx=0) AND (epizio=1) AND (emp=0) THEN solutie:="risc=0";
```



Medical Recommendations	Patient Data
The medical advice is: risk=0	- Age: 18 - background: R - gesta: 2 - para: 2 - weeks: 40

Figura 6. De la reguli obținute cu ajutorul *data mining* la sfaturi medicale

Capitolul 5 - Conține o analiză referitoare la traducerea regulilor medicale reprezentate cu ajutorul formalismului *ASTI* în formalismul *Arden Syntax* precum și obținerea datelor necesare (ale pacientului), dintr-o bază de date care implementează standardul *vMR*, pentru a permite inferența. Pe baza acestei analize s-a observat faptul că cele mai mari probleme în translatarea dintr-un formalism în altul o reprezintă construcțiile de tip obiect din *ASTI* și modul în care aceste construcții sunt utilizate în interiorul regulilor. Formalismul *Arden Syntax* permite definirea de obiecte începând cu versiunea 2.5 (versiunea curentă este 2.8), însă pentru aceste obiecte nu pot fi definite funcții. Astfel sunt prezentate trei soluții pentru reprezentarea regulilor medicale din formalismul *ASTI* în formalismul *Arden Syntax*:

- Utilizarea obiectelor *Arden Syntax* pentru reprezentarea conceptelor de medicament și tratament din modelul *ASTI*.
- Utilizarea de fișiere multiple *Arden Syntax*, fiecare din acest tip de fișier tratând diferite tipuri de construcții *ASTI* (date pacient, medicament, tratament) și modul în care aceste sunt utilizate în interiorul regulilor, fișierele *Arden Syntax* fiind interconectate.
- Utilizarea de funcții externe pentru compararea conceptelor de tratament și medicație.

În acest fel, în locul achiziționării de reguli medicale în format *Arden Syntax* din ghidurile medicale narrative, acestea pot fi traduse direct din formalismul *ASTI*, reducând personalul medical necesar.

Pentru a genera recomandări medicale bazate pe regulile medicale reprezentate cu ajutorul formalismului *ASTI*, s-a analizat posibilitatea utilizării unei surse de date care implementează modelul de date *vMR* (standard de reprezentare a datelor pacienților pentru *CDS*). Această analiză a fost făcută cu scopul de a permite utilizarea unor softuri care implementează modelul *ASTI*, în cât mai multe unități medicale (*HL7 vMR* fiind singurul standard dedicat sistemelor *CDS*). Analiza a relevat că modelul de date actual *vMR* nu permite reprezentarea tuturor datelor pacienților care sunt necesare pentru inferență. Acesta este motivul pentru care am sugerat extinderea modelului de date *vMR* prin adăugarea de noi clase (Figura 7) cu privire la atributele ale tratamentului care nu au putut fi găsite în modelul de date *vMR*.

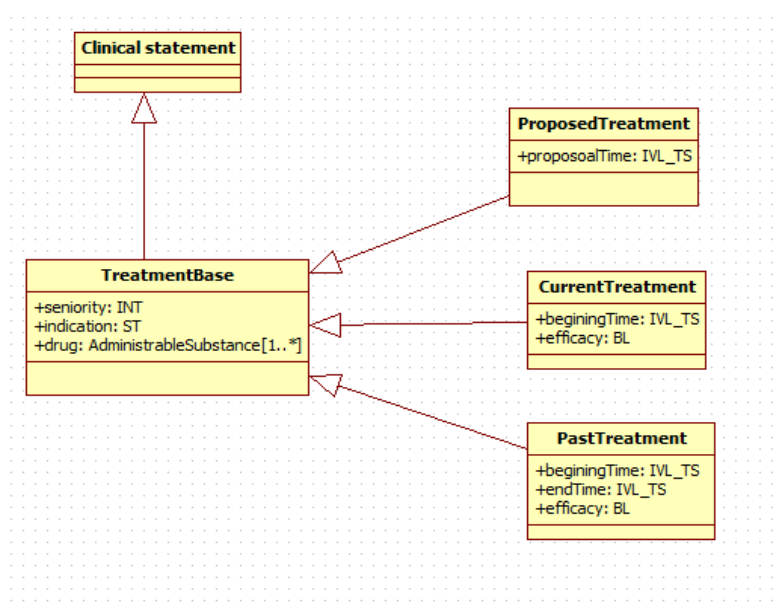


Figura 7. Clasele tratament în *vMR*

Pentru testarea metodelor propuse s-au tradus 20 de reguli *ASTI* în format *Arden Syntax* iar noile clase din *vMR* au fost reprezentate cu ajutorul *TM-vMR* (din Capitolul 3). Utilizând sistemul propus în Capitolul 3, precum și noile reguli traduse și noile clase *vMR*, s-au generat recomandări medicale. Pentru aceleași seturi de date ale pacientului, recomandările medicale obținute în acest mod au fost identice cu cele generate cu ajutorul *ASTI*, validând astfel aceste metode.

Capitolul 6 - Recapitulează discuțiile și concluziile cercetării, subliniază contribuțiile aduse de autor și precizează direcțiile de dezvoltare propuse pentru a îmbunătăți sistemele clinice de suport a deciziei clinice.

Contribuțiile originale revendicate de autor în concluzii sunt următoarele:

1. Contribuții privind analiza și prezentarea diverselor aspecte ale domeniului:

- 1.1. prezentarea critică și compararea diferitelor soluții cu concluzii utile pentru propriile cercetări;
- 1.2. analiza evoluției diferitelor abordări începând cu anul 1975 până în anul 2012.
- 1.3. analiza regulamentului *HIPAA* (Health Insurance Portability and Accountability Act) pentru a implementa o comunicare sigură între diferitele componente ale sistemului.

2. Contribuții teoretice:

- 2.1. dezvoltarea unei metode care permite traducerea regulilor medicale obținute prin *data mining* în fișiere specifice *Arden Syntax*; precum și crearea de categorii pe baza elementelor regulilor *data mining*: *variable*, *value*, *operator*, *conclusion_variable*, *conclusion_value*. Această clasificare a variabilelor a fost necesară pentru a genera secțiuni diferite ale *MLM*-urilor;
- 2.2. dezvoltare a trei metode (bazate pe: obiecte *Arden Syntax*, fișiere *Arden Syntax* multiple, funcții externe) care permit traducerea diferitelor construcții specifice *ASTI* (ex: obiecte reprezentând medicația) în *Arden Syntax*;
- 2.3. adăugarea de noi elemente la modelul de date *vMR*.

3. Contribuții privind dezvoltarea de soluții software originale:

- 3.1. propunerea, implementarea, testarea și validarea unei arhitecturi de sistem care permite integrarea unor surse de date multiple pe baza standardelor *HL7*.
- 3.2. dezvoltarea unei interfețe care pune în aplicare anumite caracteristici bine apreciate de către medici (specifice pentru sistemele de suport a deciziei clinice), pe baza unei analize anterioare a soluțiilor din domeniu;
- 3.3. implementarea în *C#* a modulelor software care traduc regulile obținute cu tehnologia *data mining* în reguli în format *Arden Syntax*.

4. Contribuții privind validarea și evaluarea soluțiilor propuse, bazate pe studii de caz reale:

- 4.1. dezvoltarea unei metode de evaluare a calității interfeței cu utilizatorul, bazată pe diferite caracteristici care sunt bine apreciate de către personalul medical;
- 4.2. dezvoltarea unor metode de validare a utilizării mai multor documente *HL7 CDA* și date în format *vMR*, utilizând 30 de seturi de date ale pacienților și reguli medicale cu privire la gestionarea diabetului zaharat utilizate în *Spitalul Județean Clinic de Urgență Timișoara*;

- 4.3. metode de validare a regulilor medicale reprezentate cu ajutorul *Arden Syntax* și extrase cu ajutorul *data mining*, pentru seturile de date reale pentru nou-născuți, de la *Clinica de Obstetrică și Ginecologie Bega, Timișoara*;
- 4.4. metode de validare a regulilor medicale în format *Arden Syntax* traduse din format *ASTI*, utilizând 30 seturile de date ale pacienților de la *Spitalul Județean Clinic de Urgență Timișoara*.

Direcții de dezvoltare:

O prima direcție de dezvoltare o reprezintă creșterea interoperabilității sistemelor de asistență a deciziei medicale prin adăugarea de noi standarde la soluția prezentată. Câteva din aceste standarde ar putea fi: *CCD (Continuity of Care Document)*, *Gello* (standard pentru reprezentarea regulilor medicale).

Folosirea de noi metode pentru descoperirea de reguli medicale sau utilizarea altor algoritmi *data mining*. De asemenea sistemul dezvoltat pentru obținerea de reguli medicale va fi utilizat pentru extragerea unor reguli mai relevante din alte baze de date. Un pas important este acceptarea rezultatului acestor reguli ca fiind o recomandare medicală și nu un sfat medical, prin validări multiple (a regulilor) de către specialiști în medicină.

O altă direcție de cercetare o reprezintă analizarea altor formalisme (care nu sunt un standard) de reprezentare a regulilor medicale pentru traducerea acestora într-un formalism acceptat ca standard, permițând utilizarea acestor reguli în cât mai multe sisteme de suport a deciziei medicale (care implementează standardul). Pe lângă aceasta este necesară dezvoltarea unor module care să permită traducerea automată a acestor formalisme.

În ceea ce privește dezvoltarea și utilizarea standardului *vMR*, este necesară analiza tipurilor de date (ale pacientului) necesare altor soluții decât cele studiate în dezvoltarea acestui standard. Pe baza acestor analize se poate verifica dacă standardul permite reprezentarea tuturor datelor necesare sau e nevoie de adăugarea unor noi elemente standardului.