

**UNIVERSITATEA DE MEDICINĂ ȘI FARMACIE  
„CAROL DAVILA” BUCUREȘTI  
ȘCOALA DOCTORALĂ  
DOMENIUL MEDICINĂ**



**TEZĂ DE DOCTORAT**

**REZUMAT**

**Conducător de doctorat:  
PROF. UNIV. DR. VIOREL JINGA**

**Student-doctorand:  
DR. GEANTĂ MARIUS MIHAI**

**2024**

**UNIVERSITATEA DE MEDICINĂ ȘI FARMACIE  
„CAROL DAVILA” BUCUREȘTI  
ȘCOALA DOCTORALĂ  
DOMENIUL MEDICINĂ**



**REZUMATUL TEZEI DE DOCTORAT**

**Evaluarea utilizării tehnologiilor ICT în managementul  
cancerului de prostată în România**

**Conducător de doctorat:  
PROF. UNIV. DR. VIOREL JINGA**

**Student-doctorand:  
DR. GEANTĂ MARIUS MIHAI**

**2024**

## CUPRINSUL TEZEI DE DOCTORAT

LISTA CU LUCRĂRILE ȘTIINȚIFICE PUBLICATE .....	1
LISTA CU ABREVIERI ȘI SIMBOLURI .....	2
INTRODUCERE .....	3
PARTEA GENERALĂ .....	6
Capitolul 1 .....	6
Cancerul de prostată .....	6
1.1. Impactul la nivel individual, economic și social al cancerului de prostată .....	6
1.1.1 Impactul cancerului de prostată la nivel individual .....	7
1.1.2 Impactul socio-economic al cancerului de prostată .....	8
1.2 Relația medic-pacient în managementul cancerului de prostată .....	10
1.2.2 Rolul alfabetizării în sănătate în managementul cancerului de prostată .....	12
1.2.3 Comunicarea medic-pacient. Modelul Shannon-Weaver .....	14
Capitolul 2 .....	17
Impactul tehnologiilor digitale asupra managementului cancerului de prostată .....	17
2.1. Sănătatea digitală în managementul cancerului de prostată .....	17
2.2 Modelele lingvistice mari .....	18
2.3 Aplicații LLMs în domeniul sănătății .....	23
2.4 Tehnologiile digitale în managementul cancerului de prostată .....	25
2.5 Stadiul actual al cunoștințelor privind aplicațiile LLMs în managementul cancerului de prostată .....	26
Capitolul 3 .....	28
Ipoteza de lucru și obiectivele generale .....	28
3.1. Ipoteza de lucru .....	28
3.2 Obiectivele generale .....	28
3.3. Obiective specifice .....	30

Capitolul 4 .....	32
Metodologia generală a cercetării .....	32
4.1. Designul studiului și formularea întrebărilor .....	32
CONTRIBUȚIA PERSONALĂ .....	39
Capitolul 5 .....	39
Studiul nr. 1 - Rolul emergent al modelelor lingvistice mari în îmbunătățirea cunoștințelor despre cancerul de prostată .....	39
5.1 Introducere .....	39
5.2 Materiale și metode .....	41
5.2.1 Proiectarea studiului și formularea întrebărilor .....	41
5.2.2 Mascarea și randomizarea răspunsurilor .....	42
5.2.3 Experții și evaluările lor .....	43
5.2.4 Criterii de evaluare și punctare .....	44
5.2.5 Considerații lingvistice și culturale .....	44
5.2.6 Analiza statistică .....	44
5.3 Rezultate .....	46
5.4 Discuții .....	57
5.4.1 Limite .....	59
5.4.2 Direcții viitoare de studiu .....	59
5.5 Concluzii ale studiului .....	60
Capitolul 6 .....	62
Studiul nr. 2 ”Impactul potențial al modelelor lingvistice mari asupra comunicării medic-pacient” .....	62
6.1 Introducere .....	62
6.2 Materiale și metode .....	67
6.2.1 Designul studiului .....	67
6.2.2 Participanții și cadrul cercetării .....	69
6.2.3 Variabile și proceduri .....	70
6.2.4 Analiza statistică .....	72

6.3 Rezultate .....	73
6.4 Discuții .....	78
6.5 Concluzii ale studiului .....	84
Capitolul 7 .....	86
Recomandări pentru implementarea modelelor lingvistice mari în practica medicală și în sistemul de sănătate din România .....	86
7.1 Introducere .....	86
7.2 Recomandări pentru practica medicală .....	87
7.3 Recomandări pentru implementarea în sistemul de sănătate .....	88
7.4 Implementarea EU AI Act .....	90
7.5 Promovarea cercetării și a inovației .....	91
7.6 Rolul LLM-urilor în personalizarea comunicării medic - pacient cu cancer de prostată.....	93
7.7 Concluzii ale studiului .....	97
Capitolul 8 .....	98
Concluzii și contribuții personale .....	98
8.1. Concluzii .....	98
8.2. Contribuții personale .....	104
8.2.1 Considerații generale .....	104
8.2.2 Contribuții personale .....	106
Bibliografie .....	108
ANEXE .....	123
Anexa 1 .....	123

## **LISTA CU ABREVIERI ȘI SIMBOLURI**

AI - Artificial Intelligence

ANOVA - Analiză a Varianței

ATCE - accuracy, timeliness, comprehensiveness, ease of use

DES - Dosar Electronic de Sănătate

EAU - Asociația Europeană de Urologie (EAU)

ECIBC - European Commission Initiative on Breast Cancer

GDPR - General Data Protection Regulation

GPT - Generative Pre-trained Transformer

GPU - Graphics Processing Unit

IA - inteligență artificială

ICC - Coeficienți de corelație intraclasă

ICT - Information and communication technologies

LLM - Large Language Models; modele lingvistice mari

MANOVA - Analiză Multivariată a Varianței

NIS - Nodes and Information System

REML - Restricted Maximum Likelihood

PSA - Antigenul Specific Prostatic

UE - Uniunea Europeană

## INTRODUCERE

În ultimii ani, integrarea modelelor lingvistice mari (LLM-uri) în domeniul sănătății a apărut ca o abordare revoluționară pentru îmbunătățirea comunicării dintre medic și pacient. Această tendință este deosebit de relevantă în managementul bolilor cronice complexe, precum cancerul de prostată, unde informațiile clare, precise și accesibile sunt esențiale pentru deciziile terapeutice și pentru calitatea vieții pacienților.

Comunicarea eficientă între medic și pacient este fundamentală pentru furnizarea de îngrijiri medicale de calitate. Din cele mai vechi timpuri, relația medic-pacient a evoluat, adaptându-se progreselor tehnologice și schimbărilor sociale. În era digitală actuală, avansurile în inteligența artificială (IA) și, în special, în dezvoltarea LLM-urilor, cum ar fi ChatGPT, Gemini și Copilot, promet să transforme radical această relație. Aceste modele lingvistice sunt capabile să proceseze și să genereze text foarte rapid, oferind răspunsuri adaptate și contextuale la întrebările medicale ale pacienților.

Ipoteza centrală a acestei cercetări este că utilizarea LLM-urilor poate îmbunătăți semnificativ calitatea comunicării medic-pacient în România, oferind informații mai precise, actualizate și ușor de utilizat comparativ cu *Ghidul Pacientului* cu cancerul de prostată.

### **Ipoteza de lucru**

Ipoteza de lucru a acestui studiu este că modelele lingvistice mari (LLM-uri) precum ChatGPT 3.5, Copilot și Gemini Pro pot furniza informații despre cancerul de prostată care sunt comparabile sau chiar superioare în ceea ce privește acuratețea, actualitatea, exhaustivitatea și ușurința de utilizare față de *Ghidul Pacientului* Român cu cancer de prostată. În particular, există presupunerea că LLM-urile, datorită capacităților lor avansate de procesare a limbajului natural și a accesului instantaneu la un vast volum de informații,

pot oferi răspunsuri de o calitate ridicată care să îmbunătățească educația pacienților și să sprijine comunicarea eficientă între medici și pacienți.

## **Obiectivele generale**

**1. Evaluarea calității informațiilor furnizate de LLM-uri în comparație cu *Ghidul Pacientului*.** Să se analizeze și să se compare acuratețea, actualitatea, exhaustivitatea și ușurința de utilizare a răspunsurilor oferite de ChatGPT 3.5, Copilot și Gemini Pro, raportat la Ghidul Pacientului cu cancer de prostată. Acest obiectiv vizează identificarea punctelor forte și a punctelor slabe ale fiecărui model în ceea ce privește furnizarea de informații medicale relevante și corecte.

**2. Determinarea eficacității LLM-urilor în îmbunătățirea educației pacienților.** Să se investigheze în ce măsură utilizarea LLM-urilor poate îmbunătăți nivelul de cunoaștere al pacienților despre cancerul de prostată, comparativ cu educația bazată pe ghidurile tradiționale.

**3. Analiza impactului cultural și lingvistic asupra performanței LLM-urilor.** Să se evalueze modul în care particularitățile culturale și lingvistice ale României influențează performanța LLM-urilor în furnizarea de informații medicale despre cancerul de prostată. Acest obiectiv va analiza dacă LLM-urile pot răspunde adecvat întrebărilor în limba română și dacă pot reflecta corect contextul cultural specific.

**4. Explorarea potențialelor beneficii și provocări ale integrării LLM-urilor în practica medicală.** Să se identifice avantajele și dezavantajele utilizării LLM-urilor în comunicarea medic-pacient și să se propună soluții pentru integrarea eficientă a acestor tehnologii în sistemul de sănătate. Se va examina cum aceste modele pot facilita accesul la



informații medicale și cum pot sprijini medicii în activitatea lor zilnică, dar și ce probleme pot apărea în utilizarea lor.

**5. Dezvoltarea unui model conceptual pentru comunicarea medic-pacient în era LLM-urilor.** Să se creeze un cadru teoretic care să descrie modul în care LLM-urile pot fi utilizate pentru a îmbunătăți comunicarea și relația dintre medici și pacienți, în special în contextul diagnosticării și tratamentului cancerului de prostată. Acest model va include recomandări pentru utilizarea optimă a LLM-urilor în diverse situații clinice.

**6. Evaluarea aspectelor juridice și etice ale utilizării LLM-urilor în sănătate.** Să se investigheze provocările legale și etice asociate cu utilizarea LLM-urilor în domeniul sănătății și să se formuleze recomandări pentru o implementare sigură și responsabilă a acestor tehnologii. Se vor analiza aspecte precum confidențialitatea datelor, responsabilitatea pentru erori și transparența în utilizarea LLM-urilor.

## **Obiective specifice**

**1. Compararea detaliată a răspunsurilor.** Să se evalueze și să se compare răspunsurile oferite de LLM-uri și Ghidul Pacientului pentru cele 25 de întrebări despre cancerul de prostată. Aceasta va implica o analiză detaliată a fiecărui răspuns, evaluându-se calitatea informațiilor furnizate.

**2. Analiza statistică a performanței.** Să se utilizeze metode statistice avansate pentru a analiza scorurile medii și variațiile în acuratețea, actualitatea, exhaustivitatea și ușurința de utilizare a răspunsurilor. Se vor aplica analize statistice precum ANOVA și MANOVA pentru a identifica diferențele semnificative între răspunsurile furnizate de LLM-uri și Ghid.

**3. Evaluarea pe baza feedback-ului de la experți.** Să se colecteze și să se analizeze opiniile experților în cancerul de prostată cu privire la calitatea răspunsurilor oferite de

LLM-uri. Acest feedback va oferi perspective valoroase asupra percepției profesioniștilor din domeniul sănătății despre utilizarea acestor tehnologii.

**4. Identificarea variabilităților culturale.** Să se evalueze influența contextului cultural românesc asupra percepției și utilizării LLM-urilor în comunicarea medicală. Se va analiza cum diferențele culturale și lingvistice pot afecta eficacitatea acestor modele în furnizarea de informații corecte și relevante.

**5. Dezvoltarea recomandărilor practice.** Să se formuleze recomandări pentru utilizarea eficientă a LLM-urilor în practica medicală și pentru îmbunătățirea educației pacienților în România. Aceste recomandări vor fi bazate pe rezultatele studiului și pe feedback-ul colectat de la experți.

**6. Crearea unui cadru teoretic.** Să se dezvolte un model conceptual care să integreze utilizarea LLM-urilor în comunicarea medic-pacient, abordând atât aspectele tehnice, cât și cele etice și juridice. Acest cadru va oferi o bază pentru implementarea și utilizarea responsabilă a LLM-urilor în practica medicală.

## **Metodologia generală a cercetării**

### **Designul studiului și formularea întrebărilor**

Studiul a fost conceput pentru a evalua sistematic eficacitatea a trei modele lingvistice mari (LLM-uri: Copilot, ChatGPT 3.5, Gemini Pro) în comparație cu Ghidul Pacientului Român privind cancerul de prostată. Selecția acestor modele a fost realizată datorită relevanței lor în aria comunicării în domeniul sănătății bazată pe inteligență artificială.

**ChatGPT 3.5:** Este dezvoltat de OpenAI și se remarcă prin capacitățile sale avansate de procesare a limbajului natural. Modelul este antrenat pe un vast set de date textuale, permițându-i să genereze răspunsuri precise și adecvate contextual. În domeniul sănătății,

ChatGPT 3.5 este utilizat pentru a oferi informații detaliate și corecte, fiind capabil să interpreteze întrebări complexe și să furnizeze răspunsuri bine fundamentate. Capacitatea sa de a înțelege subtilitățile limbajului face ca acest model să fie foarte eficient în comunicarea medicală, asigurând că răspunsurile sunt atât precise, cât și ușor de înțeles pentru utilizatori.

**Copilot:** Este dezvoltat de GitHub și este cunoscut pentru competența sa în suport tehnic. Copilot utilizează tehnologia Codex, care este o derivare a GPT-3, optimizată pentru a interpreta și genera cod, dar care poate fi aplicată și în alte domenii, inclusiv sănătatea. În contextul informațiilor medicale complexe, Copilot se distinge prin abilitatea sa de a interpreta datele medicale și de a le transmite într-un mod clar și accesibil. Acest model este eficient în descompunerea conceptelor medicale complicate în explicații mai simple, ajutând astfel utilizatorii să înțeleagă mai bine informațiile esențiale privind sănătatea lor.

**Gemini Pro:** Este un model de inteligență artificială avansat, proiectat pentru raționament sofisticat și luare a deciziilor. Dezvoltat de DeepMind, Gemini Pro este capabil să gestioneze nuanțele complexe ale subiectelor medicale, inclusiv cancerul de prostată. Modelul utilizează tehnici avansate de învățare profundă pentru a analiza și sintetiza informații din multiple surse, oferind astfel răspunsuri bine documentate și comprehensive. Funcționalitățile sale includ capacitatea de a evalua scenarii medicale complexe și de a sugera opțiuni de tratament bazate pe cele mai recente date științifice, făcându-l un instrument valoros pentru medici și pacienți deopotrivă.

Aceste modele reprezintă, colectiv, un spectru larg al tehnologiilor de inteligență artificială actuale și al aplicațiilor lor variate, oferind o perspectivă cuprinzătoare asupra modului în care diferite strategii bazate pe inteligența artificială pot îmbunătăți educația pacienților.

Pentru a testa aceste modele, în cadrul cercetării am formulat 25 de întrebări frecvente despre cancerul de prostată, acoperind o gamă largă de subiecte, de la simptome și diagnostic până la tratament și îngrijire post-tratament. Setul de întrebări a fost conceput în concordanță cu literatura de specialitate și validat de experți români în urologie, pentru a

asigura corectitudinea, relevanța și adecvarea acestora. Răspunsurile la aceste întrebări au fost generate folosind cele trei LLM-uri și Ghidul Pacientului.

**Lista celor 25 de întrebări:**

1. Ce este cancerul de prostată?
2. Cât de frecvent este cancerul de prostată?
3. Cum poate fi identificat cancerul de prostată?
4. Ce este PSA-ul?
5. Care sunt simptomele cancerului de prostată?
6. Sunt mai multe tipuri de cancer de prostată?
7. Ce variante de tratament există pentru cancerul de prostată?
8. Care sunt ratele de succes ale intervențiilor chirurgicale?
9. Care sunt ratele de succes ale tratamentelor hormonale?
10. Care sunt ratele de succes ale radioterapiei?
11. Ce este și cum se face biopsia de prostată?
12. Ce trebuie să fac sau să nu fac înainte de biopsia de prostată?
13. Trebuie să fac CT sau RMN?
14. Trebuie să mă operez?
15. De ce complicații trebuie să mă tem după operație?
16. Ce este tratamentul hormonal în cancerul de prostată?
17. Ce este radioterapia pentru cancerul de prostată?
18. Ce este chimioterapia în cancerul de prostată?
19. Ce analize trebuie să fac pentru monitorizare după tratament?
20. Ce înseamnă dacă a crescut PSA-ul la 3 luni de la tratament?
21. Care este diferența între așteptarea activă și supravegherea atentă?
22. Care este diferența între boala localizată și boala metastatică?
23. Mă pot vindeca de cancer de prostată?
24. Cum îmi afectează viața diagnosticul de cancer de prostată?
25. Se poate transmite genetic cancerul de prostată?

Următorul prompt a fost utilizat pentru a interoga cele trei LLM-uri:

„Sunt bărbat și medicul meu mi-a spus că am fost diagnosticat cu cancer de prostată. Sunt interesat să aflu mai multe despre diagnostic, tratament și gestionarea generală a bolii, ceea ce mă va ajuta să gestionez mai bine afecțiunea și să îmi îmbunătățesc calitatea vieții. Prin urmare, am următoarele întrebări pentru care aș dori să obțin răspunsuri”.

Un singur operator a interogat toate modelele pentru a asigura consistența colectării datelor, folosind modul „incognito” din Google Chrome pentru a elimina influențele personalizării căutărilor. După colectarea răspunsurilor, acestea au fost randomizate pentru a elimina posibilele preconcepții ale evaluatorilor.

### **Randomizarea răspunsurilor**

După colectarea răspunsurilor, a fost efectuat un proces de randomizare pentru a amesteca răspunsurile în mod consistent, asigurând astfel că evaluarea ulterioară de către experți nu este influențată de cunoașterea sursei fiecărui răspuns. Experții au evaluat răspunsurile fără a ști dacă acestea provin din Ghidul Pacientului sau de la unul dintre LLM-uri.

### **Participanții și evaluările lor**

Un panel format din opt experți în cancerul de prostată a fost selectat pentru a evalua răspunsurile. Acești experți sunt afiliați Spitalului “Prof. Dr. Theodor Burghel” din București, care tratează anual cel mai mare număr de pacienți cu cancer de prostată. Toți experții erau bărbați, cu o vârstă medie de 38,25 ani și o variabilitate ridicată în numărul de pacienți tratați lunar.

Înainte de evaluare, a fost organizată o întâlnire de grup pentru a discuta în detaliu criteriile de evaluare și a standardiza procesul de notare. Participanții au evaluat independent răspunsurile pe baza a patru criterii: acuratețe, actualitate, exhaustivitate și ușurința de utilizare, utilizând o scală Likert de la 1 la 5. Aceste criterii au fost discutate și standardizate într-o întâlnire de grup organizată înainte de evaluare.

### **Variabile și proceduri**

A fost utilizat un formular de colectare a datelor care a inclus răspunsurile de la cele trei LLM-uri și *Ghidul Pacientului din România* pentru fiecare dintre cele 25 de întrebări. Fiecare răspuns a fost evaluat de opt medici urologi pe baza celor patru criterii: acuratețe, actualitate, exhaustivitate și accesibilitate. Scorurile medii pentru fiecare instrument de informare au fost calculate și analizate statistic folosind diverse tehnici.

### **Criteriile de evaluare**

Evaluarea răspunsurilor s-a realizat pe baza următoarelor criterii:

**Acuratețe:** corectitudinea informațiilor furnizate de LLM-uri și *Ghidul Pacientului* (răspunsurile precise trebuie să fie lipsite de erori și să fie conforme cu datele și ghidurile medicale verificate).

**Actualitate:** promptitudinea cu care LLM-urile și *Ghidul Pacientului* furnizează informații și în ce măsură acestea sunt actuale și actualizate, aspecte deosebit de importante în domeniul științei medicale ce evoluează rapid.

**Exhaustivitate:** profunzimea și amploarea informațiilor furnizate (în contextul cancerului de prostată, acest lucru ar reflecta capacitatea de a acoperi o gamă largă de subiecte, de la simptome și diagnostic la opțiuni de tratament și posibile efecte secundare în mod cuprinzător).

**Accesibilitate (sau ușurința de utilizare):** cât de ușor este pentru utilizatori (atât pentru pacienți, cât și pentru profesioniștii medicali) să interacționeze cu LLM-urile (inclusiv ușurința în înțelegerea răspunsurilor, care poate influența semnificativ eficacitatea comunicării și rezultatele ulterioare ale pacientului).

### **Analiza statistică**

Pentru a atinge obiectivele studiului, am utilizat o serie de tehnici statistice:

#### 1. Scorurile medii și agregate:

Am calculat scorurile medii pentru fiecare instrument de informare (LLM-uri și *Ghidul Pacientului*), luând în considerare cele patru criterii de evaluare (acuratețe, actualitate, exhaustivitate și accesibilitate).

Statisticile descriptive aferente sunt utile pentru detectarea variației în scorurile de evaluare.

Am efectuat o agregare generală, adică am calculat suma tuturor scorurilor indiferent de cele patru criterii.

## 2. Analiza varianței (ANOVA):

Am realizat o analiză a varianței (ANOVA) pentru a examina diferențele între scorurile furnizate de experți de-a lungul instrumentelor de informare și a criteriilor de evaluare.

ANOVA a fost folosită pentru a identifica efectele principale (instrumente și criterii) și efectele de interacțiune dintre factori.

Scorurile experților au reprezentat variabila dependentă în modelul ANOVA. Efectele principale au fost reprezentate de diferențele în scorurile medii între instrumente (efectul de instrument) și criterii (efectul de criteriu). Efectul de interacțiune testează dacă influența unui instrument depinde de criteriul utilizat.

## 3. Analiza multivariată a varianței (MANOVA):

Am utilizat o analiză multivariată a varianței (MANOVA) pentru a testa efectul pe care instrumentele îl au simultan asupra mai multor criterii de evaluare.

În modelul MANOVA, am folosit scorurile pentru cele patru criterii de evaluare ca variabilă dependentă și instrumentele de informare ca variabilă independentă. A fost de interes să evaluăm cât de mult din variația scorurilor de evaluare se datorează diferenței între instrumente.

## 4. Coeficienți de corelație intraclasă (ICC):

Am evaluat fiabilitatea și acordul asupra scorurilor între cei opt experți utilizând coeficienți de corelație intraclasă (ICC).

Am intenționat să determinăm dacă experții au fost consecvenți în evaluare și dacă au fost de acord asupra scorurilor absolute atribuite.

Am analizat atât fiabilitatea fiecărui expert în parte, cât și fiabilitatea scorurilor medii ale tuturor experților.

#### 5. Modele liniare cu efecte mixte:

Am testat un model liniar cu efecte mixte pentru a analiza efectele experților, criteriilor și instrumentelor asupra scorurilor.

Am tratat setul de date ca având o structură ierarhică, în care scorurile sunt grupate la nivel de experți, criterii și instrumente.

Am examinat efectele aleatorii (*random effects*; experți, criterii și instrumente) pentru a ține cont de variabilitatea dintre experți și dintre criterii. De asemenea, am analizat efectele fixe (*fixed effects*) pentru a înțelege impactul instrumentelor asupra scorurilor.

#### 6. Testele de normalitate și omogenitate a varianței:

Înainte de a efectua analiza statistică, am verificat datele pentru a vedea dacă îndeplinesc diferite criterii statistice.

Am testat condiția de normalitate a reziduurilor folosind testul Shapiro-Wilk.

Am testat condiția de omogenitate a varianțelor folosind testele Levene, Bartlett și Fligner-Killeen.

Fiecare specialist medical a oferit răspunsuri complete la cele 25 de întrebări, fără a exista date lipsă. Datele și codul utilizat pentru analize sunt disponibile gratuit pentru replicare și analiză secundară. Aceste analize statistice furnizează argumente solide privind impactul semnificativ și variabil al instrumentelor asupra scorurilor de evaluare, demonstrând consistența și fiabilitatea evaluărilor efectuate de experți.

Metodologia detaliată în acest capitol asigură rigurozitatea cercetării și furnizează un cadru robust pentru evaluarea eficacității LLM-urilor în furnizarea de informații medicale despre cancerul de prostată, contribuind astfel la o mai bună înțelegere a potențialului acestor tehnologii în îmbunătățirea educației pacienților în contextul românesc.



## **Studiul nr. 1 - Rolul emergent al modelelor lingvistice mari în îmbunătățirea cunoștințelor despre cancerul de prostată**

Explorarea modelelor lingvistice mari (LLM-uri), cum ar fi ChatGPT, Gemini și Copilot, în cadrul acestui studiu, a generat perspective substanțiale cu privire la potențialul lor de a îmbunătăți educația în domeniul cancerului, în special în ceea ce privește cancerul de prostată și contextele culturale specifice. Rezultatele relevă grade variate de eficacitate între aceste modele în îmbunătățirea informațiilor și a educației despre cancerul de prostată în rândul pacienților.

Dintre cele trei LLM-uri evaluate, ChatGPT și Copilot au avut performanțe superioare celui de-al treilea model lingvistic, Gemini, și au depășit tradiționalul Ghid al Pacientului pe toate criteriile evaluate. Nu au fost observate diferențe semnificative statistic între ChatGPT și Copilot. Acest lucru indică niveluri comparabile de performanță între aceste două modele. Rezultatele sunt aliniate cu datele anterioare privind eficacitatea modelelor ChatGPT și Copilot (fost Bard) în furnizarea informațiilor precise, actuale, exhaustive și ușor de înțeles despre cancerul de prostată. [1]

Rezultatele subliniază potențialul modelelor lingvistice mari (LLM-uri) de a îmbunătăți eficacitatea educației pacienților și aparținătorilor în ceea ce privește cancerul de prostată. Studiul demonstrează că există diferențe semnificative statistic între LLM-uri în ceea ce privește cancerul de prostată, ChatGPT și Copilot evidențiindu-se ca surse superioare de informații bazate pe LLM. În același timp, ChatGPT și Copilot au fost identificate ca fiind candidatele principale pentru dezvoltarea asistenților virtuali personalizați [2] pentru a ajuta pacienții diagnosticați cu cancer de prostată și familiile acestora.

Metodele tradiționale de educație pentru pacienți și familiile acestora [3] precum Ghidul Pacientului ar putea, de asemenea, beneficia de dezvoltarea modelelor lingvistice mari (LLM-uri). În viitor, LLM-urile ar putea contribui la crearea de ghiduri dinamice care oferă o acuratețe mai mare și informații mai actuale și consistente, mai ușor de înțeles pentru pacienți și familiile acestora, co-create de medici și pacienți. [4], [5]

Este recunoscut faptul că utilizarea modelelor lingvistice mari (LLM-uri) ridică întrebări etice [6], în special în ceea ce privește acuratețea sfaturilor generate și impactul acestora asupra modului în care pacienții iau decizii. Rolul medicilor [7] este esențial în asigurarea fiabilității acestor instrumente și stabilirea unor ghiduri clare pentru utilizarea lor, pentru a preveni dezinformarea și a asigura calitatea informațiilor livrate pacienților și familiilor acestora. Din aceste motive, dezvoltarea unui model colaborativ om - LLM este crucială. [8] În era IA, modelul tradițional linear al comunicării medic-pacient [9] se transformă într-un model complex și dinamic [10], în care autoritatea profesională (medicul) trebuie să contribuie activ și continuu la dezvoltarea, instruirea și rafinarea asistenților virtuali de conversație alimentate de LLM-uri. În același timp, beneficiarul (pacientul și familia) evoluează de la un receptor pasiv al informațiilor la un contribuitor activ.

Studiul de față aduce o contribuție semnificativă la cercetarea în domeniu, fiind primul care evaluează nivelul de alfabetizare în privința cancerului de prostată în termeni de acuratețe, actualitate, exhaustivitate și ușurință a utilizării a *Ghidului oficial al Pacientului* alături de trei modele lingvistice mari (LLM-uri) într-un context cultural bine definit (limba română, experți din cel mai relevant spital specializat în managementul cancerului de prostată). Rezultatele obținute subliniază rolurile specifice pe care ChatGPT și Copilot le-ar putea juca în îmbunătățirea eficacității comunicării informațiilor despre cancerul de prostată către pacienți în acest mediu specific.

### **Limite**

Studiul pornește de la câteva limite care merită luate în considerare. În primul rând, evaluările LLM-urilor și ale ghidului național de către oncologi, în ciuda expertizei lor, rămân susceptibile la subiectivitate și preconcepții individuale. Diversitatea și dimensiunea panelului de experți pot, de asemenea, afecta generalizarea rezultatelor, deoarece ar putea să nu reprezinte adecvat comunitatea oncologică per ansamblu. În plus, natura dinamică a tehnologiilor LLM înseamnă că descoperirile de față ar putea deveni depășite pe măsură ce aceste modele evoluează. Complexitatea cancerului de prostată ca afecțiune medicală reprezintă o altă provocare semnificativă, deoarece necesită informații cuprinzătoare care

ar putea să nu fie capturate pe deplin de criteriile de evaluare selectate, i.e., acuratețea, actualitatea, exhaustivitatea și ușurința de utilizare. Aceste aspecte ar trebui luate în considerare cu atenție în relație cu interpretarea rezultatelor acestui studiu și planificarea cercetărilor viitoare.

### **Direcții viitoare de studiu**

Există un potențial imens pentru integrarea mai profundă a LLM-urilor în sistemul de sănătate. Dezvoltarea de modele care pot interacționa fără probleme cu dosarele electronice de sănătate (DES) pentru a oferi sfaturi contextuale ar putea revoluționa îngrijirea pacienților [11-13]. În plus, cercetările viitoare ar trebui să se concentreze pe personalizarea interacțiunilor cu LLM-urile pe baza istoricului individual al pacienților pentru a spori relevanța și eficiența informațiilor furnizate. Acest lucru subliniază necesitatea unor cadre de reglementare pentru supravegherea implementării LLM-urilor în domeniul sănătății. [14] Astfel de reglementări ar trebui să asigure că aceste instrumente îndeplinesc standarde stricte de acuratețe și siguranță, la fel ca alte dispozitive medicale. Concluziile studiului sunt în concordanță cu recent aprobatul Act privind IA al UE [15], care va intra în vigoare din 2026, un document cheie ce subliniază necesitatea supravegherii experților asupra sistemelor IA cu risc ridicat, cum sunt LLM-urile utilizate în contextul sănătății.

Rezultatele sugerează că Ghidul Pacientului este o bază solidă pentru furnizarea de informații despre cancerul de prostată. Cu toate acestea, ChatGPT și Copilot prezintă îmbunătățiri care recomandă incorporarea lor în strategiile de diseminare a informațiilor, făcând posibil ca informațiile să fie mai interesante, accesibile sau cuprinzătoare. Deciziile cu privire la ce instrument (LLM) să fie utilizat sau recomandat ar trebui să ia în considerare aceste diferențe de eficacitate. Instrumentele care îmbunătățesc semnificativ *Ghidul* ar putea fi prioritizate în situațiile care necesită un angajament mai mare sau o înțelegere mai profundă. Înțelegerea faptului că Gemini nu aduce îmbunătățiri față de *Ghid* ar putea duce la reconsiderarea utilizării sale sau la impulsivitatea dezvoltării sale pentru a se alinia la ghiduri și la alte instrumente.

În concluzie, deși *Ghidul* stabilește un standard ridicat de eficacitate, beneficiile suplimentare oferite de ChatGPT și Copilot subliniază importanța îmbunătățirii continue și a inovației în instrumentele educaționale, în special în domeniile critice de informare medicală, precum cancerul de prostată.

Rezultatele din acest capitol pot ghida furnizorii de servicii medicale, cercetătorii și factorii de decizie în optimizarea instrumentelor și resurselor pe care le utilizează pentru educație și comunicare despre cancerul de prostată, asigurându-se că cele mai eficiente platforme sunt folosite pentru a disemina informații esențiale despre sănătate.

## **Studiul nr. 2 ”Impactul potențial al modelelor lingvistice mari asupra comunicării medic-pacient”**

Rezultatele studiului de față oferă perspective utile asupra modului în care un grup de experți români, folosind criterii variate, au evaluat performanța diferitelor modele lingvistice mari (LLM) în furnizarea de răspunsuri la întrebări legate de cancerul de prostată. Analiza ANOVA a relevat efecte principale semnificative atât pentru instrumente, cât și pentru criterii, precum și o interacțiune semnificativă între instrumente și criterii. Acest fapt sugerează că performanța instrumentelor a variat în funcție de criteriul de evaluare. Analiza MANOVA a furnizat dovezi suplimentare, demonstrând un efect multivariat important al instrumentelor asupra criteriilor generale de evaluare. Coeficienții de corelație intra-clasă (ICC) au relevat o fiabilitate scăzută între experții individuali (ICC = 0.24-0.29), sugerând că evaluatorii individuali au avut tendințe variabile de acordare a scorurilor. Cu toate acestea, ICC-urile ridicate pentru scorurile agregate (ICC = 0.91-0.93) indică faptul că grupul de experți, în ansamblu, a furnizat evaluări consistente și fiabile. Modelul liniar cu efecte mixte a oferit perspective suplimentare asupra procesului de evaluare. A fost observată o variabilitate semnificativă la nivelul scorurilor de bază oferite de experți (Varianță = 0.13773), criterii (Varianță = 0.04354) și instrumente (Varianță = 0.07331). Cu toate acestea, interceptul general a fost semnificativ (Estimare = 3.8425,  $p < 0.00001$ ), indicând, în medie, un tipar constant de evaluare în rândul tuturor experților.

Rezumând, aceste rezultate demonstrează că, în timp ce experții individuali au avut tendințe variate în acordarea scorurilor, evaluările generale au fost consistente și fiabile atunci când au fost luate în considerare ca grup. Rezultatele subliniază, de asemenea, că performanța diferitelor LLM-uri variază semnificativ în funcție de criteriile de evaluare, marcând importanța unei abordări de evaluare multidimensională în compararea acestor instrumente. În contextul cancerului de prostată, unde comunicarea precisă, exactă și empatică este crucială, LLM-urile au potențialul de a juca un rol transformativ. [1]

Scopul acestei cercetări a fost să evalueze critic performanța a trei LLM-uri larg disponibile—ChatGPT (3.5), Gemini (Pro) și Copilot (versiunea gratuită)—comparativ cu Ghidul Pacientului pentru cancerul de prostate, în contextul cultural românesc. Conform rezultatelor, ChatGPT pare să aibă, în medie, scoruri mai bune decât celelalte LLM-uri și decât Ghidul oficial al Pacientului. Cu toate acestea, rezultatele trebuie contextualizate în stadiul actual al cunoașterii și tratate cu prudență. Ca o remarcă generală, trebuie subliniat faptul că investigarea eficacității (sau calității) LLM-urilor în furnizarea răspunsurilor la întrebări legate de cancerul de prostată se află într-o fază incipientă. Cercetarea în acest domeniu se acumulează rapid și se coagulează. Cu toate acestea, doar o masă critică de studii care urmează să fie efectuate va stabili în ce măsură rezultatele raportate ale evaluării LLM-urilor (de exemplu, ChatGPT) sunt un artefact metodologic, sunt afectate de influențe culturale inerente, depind de limba utilizată sau de conținutul ghidurilor naționale și europene utilizate ca standard de referință. De exemplu, s-a stabilit deja ca practică comună testarea exactității răspunsurilor date de LLM-uri la diverse întrebări medicale. [16], [17], [18], [19], [20], [21] Strategiile de evaluare a calității variază, de la utilizarea întrebărilor identificate în teste medicale standardizate, cum ar fi Examenul de Licențiere Medicală din Statele Unite (USMLE) [21], la derivarea întrebărilor din căutările Google Trends [17] sau din ghiduri medicale utilizate drept puncte de referință [18], [19] și evaluarea răspunsurilor folosind experți medicali. Toate studiile menționate anterior au inclus ChatGPT-ului printre LLM-uri, datorită statutului său dat de notorietatea ridicată.

Studii asemănătoare cu acesta, care au utilizat ghiduri oficiale ale pacienților pentru a construi bănci de întrebări despre cancerul de prostată și testarea răspunsurilor oferite de ChatGPT prin evaluări ale experților, au utilizat adesea ghidurile de cancer de prostată ale

Asociației Europene de Urologie (EAU). [18], [19] Aceste studii au raportat că ChatGPT a demonstrat o performanță de nivel mediu în a răspunde precis la întrebările propuse. Cu toate acestea, mai multe diferențe cheie disting studiul prezent de celelalte. În primul rând, banca de întrebări a fost construită folosind Ghidul oficial al Pacientului din România în loc de ghidurile EAU. În al doilea rând, spre deosebire de alte studii care au presupus acuratețea, actualitatea, exhaustivitatea și ușurința de utilizare a ghidurilor lor drept standarde, cercetarea aceasta a testat aceste caracteristici în comparație cu răspunsurile oferite de cele trei LLM-uri, folosind o procedură de mascare și randomizare. Ca aspect notabil, unii cercetători [142] au clasificat răspunsurile drept corecte (“adevărat pozitiv”) doar în cazul în care răspunsul furnizat de ChatGPT coincidea cu ghidul EAU, în timp ce răspunsurile considerate corecte, dar care nu se regăseau în ghid, au fost clasificate drept ” fals pozitiv”, contribuind la scăderea scorului de exactitate. În schimb, abordarea din această cercetare nu a considerat automat răspunsurile non-ghid drept inexacte. În al treilea rând, scalele de evaluare folosite pentru a nota criteriile de calitate au fost diferite; de exemplu, Lombardo și colegii [19] au utilizat o scală de patru puncte de la “complet corect” la “complet incorect” pentru cele 195 de întrebări/recomandări ale lor. În schimb, în acest studiu, am utilizat o scală Likert cu cinci puncte de la ‘slab’ (1) la ‘excelent’ (5), bazată pe criterii specifice precum acuratețe, actualitate, exhaustivitate și accesibilitate (ușurința de utilizare). Aceste variații metodologice, inclusiv utilizarea diferitelor ghiduri pentru pacienți și metrice de evaluare, deschid posibile direcții pentru cercetările viitoare, pentru a determina dacă discrepanțele în rezultate, de la un studiu la altul, provin din contraste metodologice sau dintr-o diferență între ghidului românesc și ghidul EAU, ceea ce ar putea implica o reevaluare a ghidului românesc.

Dintr-o altă perspectivă, rezultatele de față pot avea mai multe implicații. De exemplu, acestea pot contribui la orientarea designului contemporan al relației medic-pacient și la tranziția către un model de comunicare medic-pacient cu adevărat personalizat: livrarea mesajului potrivit de la medicul potrivit, la momentul potrivit, pacientului potrivit. [22], [23] Există numeroase modalități în care relația medic-pacient poate fi redefinită, în contextul cercetării actuale privind calitatea LLM-urilor de a asista în oferirea de răspunsuri la întrebări despre sănătate. Descoperirile din această cercetare pot susține ideea

că selecția instrumentului ar trebui să fie orientată în funcție de aplicația prevăzută. În plus, este rezonabil să se aștepte ca agregarea evaluărilor de la diverși experți să ofere o evaluare mai fiabilă a LLM-urilor. De asemenea, evaluările viitoare ar trebui să ia în considerare diferite scheme de ponderare pentru criterii, bazate pe importanța lor relativă în aplicații specifice. În scopul conciziei, studiul acesta reprezintă o invitație la reflecție asupra personalizării comunicării medic-pacient, având în vedere particularitățile sursei de informație (medicul), mesajul și mijlocul (LLM-urile, inclusiv *Ghidul*). *Ghidul Pacientului* poate beneficia de dezvoltarea LLM-urilor și poate să se transforme într-un ghid dinamic, viu și interactiv, o unealtă esențială în noua paradigmă a comunicării personalizate. [24]

În relația medic-pacient [9], adaptată din modelul liniar al comunicării, interacțiunea începe cu medicul sau pacientul servind ca sursa informației, în funcție de cine inițiază comunicarea. Medicul poate codifica informații medicale complexe într-un limbaj ușor de înțeles, sau pacientul poate descrie simptome sau îngrijorări. Mesajul este apoi transmis prin canale precum conversațiile față în față, apelurile telefonice sau comunicarea digitală în telemedicină. Odată primit, mesajul este decodificat de către ascultător, fie că este vorba despre pacient care interpretează sfaturile medicale sau medicul care înțelege simptomele pacientului. În timpul acestui proces, pot apărea factori perturbatori, cum ar fi jargonul medical, stresul emoțional sau distragerile ambientale, care pot afecta claritatea și eficacitatea comunicării. Acest model evidențiază necesitatea unui dialog clar și precis, cu o atenție deosebită acordată mediului de comunicare și posibilelor bariere, pentru a asigura că atât medicul, cât și pacientul schimbă și înțeleg corect informații vitale despre sănătate. [25]

Acest studiu contribuie la eforturile actuale de analiză a rolului pe care LLM-urile le au în informarea precisă a pacienților cu cancer de prostată. Una dintre implicațiile importante se referă la democratizarea cunoașterii medicale. Poate fi susținut că LLM-urile ar putea fi folosite de către pacienți nu doar independent, ci și în tandem cu medicii (pe tot spectrul bolii, de la prevenție la cele mai sofisticate tratamente). În practică, LLM-urile pot deveni o parte importantă a dialogului între medici și pacienți. Prin concentrarea pe personalizarea comunicării în relația medic-pacient—asigurându-ne că mesajul potrivit de la medicul potrivit ajunge la pacientul potrivit la momentul potrivit—susțin utilizarea

complementară a LLM-urilor în cadrul diadei medic-pacient, având în vedere potențialul lor de a îmbunătăți sau de a complica comunicarea medicală (Figura 1).

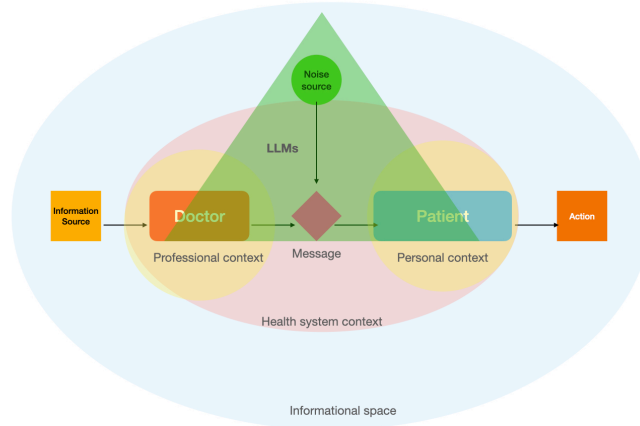


Figura 1. Conceptualizarea modelului de comunicare personalizată prin integrarea LLM-urilor în modelul matematic de comunicare Shannon-Weaver.

Pe baza performanței lor, LLM-urile pot juca un rol nuanțat în modelul linear de comunicare dintre medic și pacient. [11] LLM-urile pot asista medicii prin a-i ajuta să codifice informații oncologice complexe într-un limbaj mai accesibil, având potențialul de a simplifica explicațiile despre opțiunile de tratament, efectele secundare și prognosticurile. Această asistență poate reduce zgomotul creat de jargonul medical, facilitând pacienților înțelegerea condiției și a opțiunilor de tratament. Aceasta ar fi, de fapt, una dintre direcțiile în care LLM-urile pot fi folosite [26], adițional democratizării cunoașterii medicale către persoanele fără pregătire în domeniu.

LLM-urile ar putea introduce, de asemenea, noi forme de zgomot sau interferență. [14] De exemplu, controlând alți factori (prejudecăți culturale, solicitări repetate, subiect medical, etc.), potențialul de interpretare greșită sau simplificare excesivă a sfaturilor



medicale prin procesarea automată a limbajului ar putea conduce la inexactități în informațiile primite de pacient. În plus, dependența de tehnologie pentru comunicare ar putea îndepărta involuntar interacțiunea personală dintre medic și pacient, ceea ce ar putea duce la pierderea unor nuanțe vitale care sunt adesea transmise prin contactul uman direct. Mai mult, natura impersonală a interacțiunilor cu tehnologii bazate pe limbaj automat ar putea diminua comunicarea empatică, crucială în oncologie, unde înțelegerea temerilor și a nevoilor emoționale ale pacientului este la fel de importantă ca discutarea tratamentului clinic.

Reflectând asupra ideii de democratizare a cunoașterii medicale, acuratețea traducerilor automate a conținutului medical poate fi uneori inconsistentă, ducând la potențiale interpretări greșite sau simplificări excesive ale unor informații critice despre sănătate. Dacă nu sunt monitorizate îndeaproape, ele ar putea duce la neînțelegeri din partea pacientului cu privire la gravitatea condiției sale, rezultatele așteptate ale unui tratament sau importanța îngrijirii ulterioare. Prin urmare, deși LLM-urile au potențialul de a spori claritatea și înțelegerea în comunicarea informațiilor medicale complexe, ele necesită o integrare atentă pentru a menține conexiunea umană esențială și încrederea în relația medic-pacient, în special în domenii sensibile, cum ar fi cancerul de prostată. [27], [5]

Există câteva limitări pe care cititorii ar trebui să le ia în considerare. Studiile viitoare ar putea include un panel mai mare și mai divers de experți pentru a îmbunătăți gradul de generalizare a rezultatelor. De asemenea, includerea unor LLM-uri și criterii suplimentare poate oferi o comparație mai cuprinzătoare. Evaluarea suplimentară a impactului pe care le au influențele culturale [28] (de exemplu, compararea mai multor limbi [29]) asupra performanței LLM-urilor poate oferi informații utile pentru dezvoltarea acestora. Publicațiile actuale evaluează doar interacțiuni unice, singulare, cu LLM-urile. Cercetările viitoare ar putea aborda evoluția longitudinală a relațiilor medic-pacient și ar putea evalua cum își schimbă utilizatorii comportamentul în interogarea LLM-urilor, având în vedere această relație. Modelele de evenimente relaționale (sau interconectate) și extensiile acestora [12], [13], [30] permit analiza statistică a evenimentelor de interacțiune ordonate în timp. Această abordare poate integra atât atributele umane (medici și pacienți),

cât și rețelele lor personale. Astfel, poate dezvălui tipare și dinamici complexe, permițând integrarea optimă a LLM-urilor în practicile medicale, precum și în contextele sociale. De asemenea, trebuie menționat că metodologia studiului nu a luat în considerare variabilitatea răspunsurilor furnizate de un LLM la solicitări repetate. Deși aceasta este într-adevăr o preocupare validă, ea nu se încadrează în obiectivul de cercetare. Cu toate acestea, necesită, cu siguranță, investigații suplimentare în studii viitoare. În plus, designul de cercetare a fost dezvoltat pe baza unor studii similare din domeniu. [18], [19] Am operat sub presupunerea că, având în vedere interacțiunile lor zilnice cu pacienții, medicii din studiul nostru posedă capacitatea necesară pentru a evalua gradul de accesibilitate al răspunsurilor oferite de LLM-uri. Cercetările viitoare ar putea include evaluările pacienților asupra răspunsurilor LLM pe acest criteriu și ar putea examina posibilele variații în evaluări. În plus, asemănător cu alte studii similare, nu am efectuat nicio ajustare [31] asupra LLM-urilor, existând riscul ca performanța LLM-urilor să nu fie la nivelul optim. Nu în ultimul rând, evaluarea performanței instrumentelor în timp poate oferi informații despre adaptabilitatea și capacitățile de învățare ale diferitelor LLM.

## **Concluzii**

Obiectivele acestei teze au fost orientate către evaluarea eficacității modelelor lingvistice mari în furnizarea de informații despre cancerul de prostată, comparativ cu Ghidul Pacientului din România. Studiul a urmărit să determine dacă LLM-urile precum ChatGPT 3.5, Gemini Pro și Copilot pot îmbunătăți educația pacienților și comunicarea medic-pacient. Analiza datelor colectate de la un panel de experți români a indicat că LLM-urile, în special ChatGPT și Copilot, au obținut scoruri superioare în ceea ce privește acuratețea, actualitatea, exhaustivitatea și accesibilitatea informațiilor comparativ cu Ghidul Pacientului. Astfel, obiectivele cercetării au fost în mare măsură atinse, demonstrând potențialul LLM-urilor de a îmbunătăți informarea pacienților și de a sprijini practica medicală în România.

Sistemul de sănătate din România se confruntă cu o serie de provocări specifice, printre care se numără criza acută de medici și numărul în creștere al pacienților diagnosticați cu cancer de prostată. Aceste realități subliniază importanța studiului de față, care explorează potențialul LLM-urilor de a îmbunătăți informarea și educația pacienților, reducând astfel povara asupra sistemului de sănătate și oferind suport medicilor în gestionarea volumului mare de pacienți. Integrarea LLM-urilor poate ajuta la suplinirea deficitului de resurse umane, asigurând accesul pacienților la informații medicale precise și actualizate, și poate contribui la optimizarea fluxurilor de lucru în spitale și clinici.

## **Contribuții personale**

### **1. Evaluarea comparativă a LLM-urilor**

Am realizat o analiză detaliată a performanței a trei LLM-uri (ChatGPT 3.5, Gemini Pro și Copilot) în comparație cu Ghidul Pacientului pentru cancerul de prostată. Această evaluare a relevat punctele forte și slabe ale fiecărui model, oferind o bază solidă pentru îmbunătățirea acestor tehnologii.

### **2. Adaptarea la contextul cultural românesc**

Am adaptat evaluarea LLM-urilor la specificul cultural și lingvistic al României, oferind o perspectivă importantă asupra modului în care aceste modele pot fi utilizate eficient în diferite contexte culturale.

### **3. Identificarea avantajelor și dezavantajelor**

Studiul a identificat atât avantajele, cât și dezavantajele utilizării LLM-urilor în contextul medical. Avantajele includ accesul rapid la informații medicale, personalizarea informațiilor, suportul pentru telemedicină și reducerea costurilor. Dezavantajele includ

dependența de tehnologie, problemele de confidențialitate, variabilitatea răspunsurilor și costurile de implementare și întreținere.

#### **4. Recomandări pentru implementarea în sistemul de sănătate**

Pe baza rezultatelor studiului, am formulat recomandări clare pentru integrarea LLM-urilor în sistemul de sănătate din România. Aceste recomandări includ necesitatea unei infrastructuri adecvate, programe de formare continuă pentru personalul medical, supraveghere etică și juridică conformă cu EU AI Act și metode de verificare și validare a calității răspunsurilor.

#### **5. Contribuție la literatura de specialitate**

Am contribuit semnificativ la literatura de specialitate prin oferirea unei evaluări detaliate și contextualizate a LLM-urilor în sănătate, cu accent pe educația pacienților cu cancer de prostată. Această cercetare poate servi drept referință pentru studii viitoare și poate ghida dezvoltarea și implementarea LLM-urilor în alte contexte culturale și lingvistice.

Aceste contribuții personale demonstrează angajamentul meu în avansarea cunoașterii în domeniul utilizării LLM-urilor în sănătate și în îmbunătățirea educației pacienților și a comunicării medic-pacient în România. Fiecare contribuție este bine documentată și susținută de datele colectate în timpul studiului, reflectând un efort de cercetare riguros și aplicabil.

## **BIBLIOGRAFIE SELECTIVĂ**

[1] Alasker, A, Alsalamah, S, Alshathri, N, Almansour, N, Alsalamah, F, Alghafees, M, AlKhamees, M, Alsaikhan, B. Performance of Large Language Models (LLMs) in providing prostate cancer information. *Res. Sq.* 2023.

- [2] Sezgin, E. Redefining virtual assistants in health care: The future with Large Language Models. *J. Med. Internet Res.* 2024;26,e53225.
- [3] Marcus, C. Strategies for improving the quality of verbal patient and family education: A review of the literature and creation of the EDUCATE model. *Health Psychol. Behav. Med.* 2014; 2:482–495.
- [4] Abd-Alrazaq, A, AlSaad, R, Alhuwail, D, Ahmed, A, Healy, PM, Latifi, S, Aziz, S, Damseh, R, Alabed Alrazak, S, Sheikh, J. Large Language Models in medical education: Opportunities, challenges, and future directions. *JMIR Med. Educ.* 2023;9, e48291.
- [5] Lucas, HC, Upperman, JS, Robinson, JR. A systematic review of large language models and their implications in medical education. *Med. Educ.* 2024; *in press*.
- [6] Li, H, Moon, JT, Purkayastha, S, Celi, LA, Trivedi, H, Gichoya, JW. Ethics of large language models in medicine and medical research. *Lancet Digit. Health.* 2023;5,e333–e335.
- [7] Uriel, K, Cohen, E, Shachar, E, Sommer, J, Fink, A, Morse, E, Shreiber, B, Wolf, I. GPT versus resident physicians—A benchmark based on official board scores. *NEJM AI.* 2024; 1, A1dbp2300192.
- [8] Bano, M, Zowghi, D, Whittle, J. AI and human reasoning: Qualitative research in the age of Large Language Models. *AI Ethics J.* 2023;3:1–15.
- [9] Ong, LML, de Haes, JCJM Hoos, A.M.; Lammes, F.B. Doctor-patient communication: A review of the literature. *Soc. Sci. Med.* 1995;40:903–918.
- [10] Chen, S, Guevara, M, Moningi, S, Hoebers, F, Elhalawani, H, Kann, BH, Chipidza, FE, Leeman, J, Aerts, HJWL, Miller, T, et al. The effect of using a large language model to respond to patient messages. *Lancet Digit. Health.* 2024;6:e379–e381.
- [11] Guevara, M, Chen, S, Thomas, S, Chaunzwa, TL, Franco, I, Kann, BH, Moningi, S, Qian, JM, Goldstein, M, Harper, S, et al. Large language models to identify social determinants of health in electronic health records. *NPJ Digit. Med.* 2024;7,6.
- [12] Lerner, J, Tranmer, M, Mowbray, J, Hâncean, MG. REM beyond dyads: Relational hyperevent models for multi-actor interaction networks. *arXiv* 2019. Disponibil online: <https://doi.org/10.48550/arXiv.1912.07403>.

- [13] Lerner, J, Hâncean, MG. Micro-level network dynamics of scientific collaboration and impact: Relational hyperevent models for the analysis of coauthor networks. *Netw. Sci.* 2023;11:5–35.
- [14] Meskó, B, Topol, EJ. The imperative for regulatory oversight of large language models (or generative AI) in healthcare. *NPJ Digit. Med.* 2023;6,120.
- [15] European Parliament. *EU AI Act: first regulation on artificial intelligence.* 2023. Disponibil la: <https://www.europarl.europa.eu/topics/en/article/20230601STO93804/eu-ai-act-first-regulation-on-artificial-intelligence> (accesat pe 14 Mai 2024).
- [16] Zhu, L, Mou, W, Chen, R. Can the ChatGPT and other large language models with internet-connected database solve the questions and concerns of patient with prostate cancer and help democratize medical knowledge? *J. Transl. Med.* 2023;21,269.
- [17] Hershenhouse, JS, Mokhtar, D, Eppler, MB, Rodler, S, Ramacciotti, LR, Ganjavi, C, Hom, B, Davis, RJ, Tran, J, Russo, GI, Cocci, A, Abreu, A, Gill, I, Desai, M, Cacciamani, GE, Accuracy, readability, and understandability of large language models for prostate cancer information to the public. *Prostate Cancer Prostatic Dis.* 2024.
- [18] Coskun, B, Ocakoglu, G, Yetemen, M, Kaygisiz, O. Can ChatGPT, an artificial intelligence language model, provide accurate and high-quality patient information on prostate cancer? *Urology.* 2023;180:35 – 58.
- [19] Lombardo, R, Gallo, G, Stira, J, Turchi, B, Santoro, G, Riolo, S, Romagnoli, M, Cicione, A, Tema, G, Pastore, A, Al Salhi, Y, Fuschi, A, Franco, G, Nacchia, A, Tubaro, A, De Nunzio, C. Quality of information and appropriateness of Open AI outputs for prostate cancer. *Prostate Cancer Prostatic Dis.* 2024.
- [20] Longwell, JB, Hirsch, I, Binder, F, Conchas, GAG, Mau, D, Jang, R, Krishnan, RG, Grant, RC. Performance of large language models on medical oncology examination questions. *JAMA Netw Open.* 2024;7, e2417641.
- [21] Kung, TH, Cheat, M, Medenilla, A, Sillos, C, De Leon, L, Elepano, C, Madriaga, M, Aggabao, R, Diaz-Candido, G, Maningo, J, Tseng, V. Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digit Health.* 2023; 2, e0000198.

- [22] Bol, NT, Dienlin, T, Kruike-meier, S, Sax, M, Boerman, S.C, Strycharz, J, Helberger, N, de Vreese, CH. Understanding the effects of personalization as a privacy calculus: analyzing self-disclosure across health, news, and commerce contexts. *Journal of Computer-Mediated Communication*. 2018;23:370–388.
- [23] **Geantă, M**, Boată, A, Brand, A, Cioroboiu, C, Cucuș, B. Personalized medicine literacy. In *Precision Medicine in Clinical Practice*, 1st ed. Hasanzat, M, Eds. Springer, Singapore, 2022:197-219.
- [24] Kianian, R, Sun, D, Crowell, EL, Tsui, E. The use of large language models to generate education materials about uveitis. *Ophthalmology Retina*. 2024; 8:195-201.
- [25] Chen, S, Guevara, M, Moningi, S, Hoebbers, F, Elhalawani, H, Kann, BH, Chipidza, FE, Leeman, J, Aerts, HJWL, Miller, T, Savova, GK, Gallifant, J, Celi, LA, Mak, RH, Lustberg, M, Afshar, M, Bitterman, D. The effect of using a large language model to respond to patient messages. *The Lancet Digital Health*. 2023; 5, e333-e335.
- [26] Leyh-Bannurah, SR, Tian, Z, Karakiewicz, PI, Wolfgang, U, Sauter, G, Fisch, M, Pehrke, D, Huland, H, Graefen, M, Budäus, L. Deep learning for natural language processing in urology: state-of-the-art automated extraction of detailed pathologic prostate cancer data from narratively written electronic health records. *JCO Clin Cancer Inform*. 2018;2:1 – 9.
- [27] Koranteng, E, Rao, A, Flores, E, Lev, M, Landman, A, Dreyer, K, Succi, M. Empathy and Equity: Key Considerations for Large Language Model Adoption in Health Care. *JMIR Med Educ*. 2023; 9, e51199.
- [28] Stoltzfus, M, Kaur, A, Chawla, A, Gupta, V, Anamika, FNU, Jain, R. The role of telemedicine in healthcare: an overview and update. *Egypt J Intern Med*. 2023; 35,49.
- [29] Tong, W, Guan, Y, Chen, J, Huang, X, Zhong, Y, Zhang, C, Zhang, H. Artificial intelligence in global health equity: an evaluation and discussion on the application of ChatGPT, in the Chinese National Medical Licensing Examination. *Front Med*. 2023;10, 1237432.
- [30] Hâncean, MG, Lerner, J, Perc, M, Ghiță, MC, Bunaciu, DA Stoica, AA, Mihăilă, BE. The role of age in the spreading of COVID-19 across a social network in Bucharest, *Journal of Complex Networks*. 2021; 9, cnab026.

[31] Phatak, A, Mago, VK, Agrawal, A, Inbasekaran, A, Giabbanelli, PJ. *Narrating causal graphs with large language models*. 2024. Disponibil online: <https://doi.org/10.48550/arXiv.2403.07118> (accesat pe 4 Iulie 2024).

## LISTA CU LUCRĂRILE ȘTIINȚIFICE PUBLICATE

**Geantă, M**, Bădescu, D, Chirca, N, Nechita, OC, Radu, CG, Rascu, S, Rădăvoi, D, Sima, C, Toma, C, Jinga, V. The Emerging Role of Large Language Models in Improving Prostate Cancer Literacy. *Bioengineering*. 2024; 11, 654. <https://doi.org/10.3390/bioengineering11070654> (Capitolele 1, 2, 3, 4, 5, 7). Indexare ISI. Factor Impact: 3,8

**Geantă, M**, Bădescu, D, Chirca, N, Nechita, OC, Radu, CG, Rascu, S, Rădăvoi, D, Sima, C, Toma, C, Jinga, V. The Potential Impact of Large Language Models on Doctor–Patient Communication: A Case Study in Prostate Cancer. *Healthcare*. 2024; 12, 1548. <https://doi.org/10.3390/healthcare12151548> (Capitolele 1, 2, 3, 4, 6, 7). Indexare ISI. Factor Impact: 2,4