Optimizarea interacțiunii om-computer prin euristici de uzabilitate specifice domeniului, destinate evaluării platformelor software pentru ingineria și managementul cerințelor

Teză de doctorat – Rezumat

pentru obținerea titlului științific de doctor la
Universitatea Politehnica Timișoara
în domeniul de doctorat
Electronică, Telecomunicații și Tehnologii Informaționale

Autor: ing. ROTARU Oana Alexandra
Conducător științific: Prof.univ.dr.ing. Radu VASIU

luna:06 anul:2026

1. Motivație

Tehnologia Informației (IT) a transformat fundamental industria auto, determinând trecerea de la sisteme pur mecanice (ME) la sisteme extrem de complexe, bazate pe software (SW) și din ce în ce mai autonome. Această transformare a avut un impact semnificativ asupra practicilor de inginerie, în special prin adoptarea platformelor software și a sistemelor informatice care susțin procesele de dezvoltare [1].

Pe măsură ce vehiculele moderne se bazează tot mai mult pe software, complexitatea și diversitatea cerințelor — acoperind siguranța, funcționalitatea și conformitatea cu reglementările — au crescut considerabil. Această evoluție a condus la necesitatea unor cadre structurate pentru ciclul de viață al software-ului, în special a instrumentelor de tip Application Lifecycle Management (ALM), cu un accent puternic pe Ingineria și Managementul Cerințelor (Requirements Engineering and Management – REM) [2].

Ingineria și Managementul Cerințelor reprezintă o disciplină fundamentală în cadrul ingineriei sistemelor și ingineriei software, fiind susținută de standarde internaționale precum ISO/IEC/IEEE 29148 [3] și strâns legată de standarde mai ample privind ciclul de viață (ISO/IEC/IEEE 15288 [4] și 12207 [5]). În domeniul auto, activitățile ce tin de ingineria si managementul cerintelor sunt esențiale datorită dezvoltării sistemelor critice din punct de vedere al siguranței. Pentru a sprijini aceste activități, sunt utilizate pe scară largă platforme software specializate precum Polarion [6], Jama Connect [7] și IBM DOORS Next [8] (DNG).

Aceste platforme sunt utilizate de roluri diverse, cu atributii și obiective variate. În consecință, eficiența lor depinde nu doar de capabilitățile funcționale, ci și de modul în care utilizatorii pot interacționa eficient cu acestea. Acest aspect plasează cercetarea în domeniul Interacțiunii Om–Calculator (Human-Computer Interaction – HCI), care se concentrează pe proiectarea și evaluarea sistemelor centrate pe utilizator. Concepte-cheie ale HCI includ experiența utilizatorului (User Experience – UX), uzabilitatea și designul interfeței, așa cum sunt definite de standarde precum ISO 9241 [9].

Deși metodele generale de evaluare a uzabilității — precum evaluarea euristică bazată pe principiile lui Nielsen [10] — oferă o bază solidă, caracterul lor generic poate limita eficiența în abordarea provocărilor specifice domeniului întâlnite în platformele REM complexe.

Această teză explorează intersecția dintre Ingineria si Managementul Cerintelor (REM) și

Interactiunea Om-Computer (HCI), având ca scop echilibrarea funcționalității inginerești avansate cu interacțiunea eficientă a utilizatorului. Lucrarea propune metode de evaluare a uzabilității specifice domeniului, adaptate platformelor software utilizate în sectorul auto. În plus, investighează potențialul modelelor generative de inteligență artificială (Large Language Models – LLMs) ca instrumente bazate pe inteligență artificială (AI) pentru realizarea evaluărilor euristice (HE -Heuristic Evaluation) [11].

Cercetarea din cadrul tezei își propune să răspundă la următoarele întrebări principale:

1. Cum se reflectă considerentele Interactiunii Om-Computer în evaluarea platformelor dedicate Ingineriei și Managementului Cerințelor?
2. Sunt euristicile generice eficiente în identificarea problemelor de uzabilitate specifice domeniului, asociate fluxurilor de lucru ale platformelor dedicate REM folosite în industria auto?
3. În comparație cu euristicile generice, cât de eficiente sunt euristicile specifice domeniului, adaptate platformelor dedicate REM, în identificarea problemelor de uzabilitate?
4. Cu ce acuratețe pot LLM-urile să realizeze evaluări euristice ale platformelor dedicate REM, în ceea ce privește identificarea problemelor, severitatea acestora și acoperirea generală?

Pentru a răspunde acestor întrebări, teza analizează mai întâi stadiul actual al domeniului HCI în relatie cu platformele dedicate REM, apoi propune un set de euristici specifice domeniului, concepute pentru evaluarea uzabilității platformelor utilizate în contextul industriei auto.

Teza este structurată în șapte capitole, după cum urmează:

Capitolul 1 „Motivație" prezintă rațiunea alegerii temei de cercetare și este structurat în trei secțiuni: o prezentare generală a subiectului, în care sunt explicate domeniile interconectate, lista articolelor publicate în cadrul tezei și prezentarea structurii lucrării.

Capitolul 2 „Fundamentele Ingineriei Cerințelor în Industria Auto" oferă baza teoretică a cercetării, concentrată pe Ingineria Auto (AE) și Ingineria Cerințelor (RE). Scopul principal al acestui capitol este de a prezenta domeniul RE utilizat în industria auto, inclusiv procesele, instrumentele, precum și funcționalitățile și capabilitățile acestora. De asemenea, capitolul abordează aplicarea inteligenței artificiale în domeniul RE, cu accent pe LLM-uri și rolul ingineriei prompturilor în modelarea rezultatelor. Sunt prezentate conceptele principale ale domeniului și motivația investigării experienței utilizatorului în instrumentele REM.

Capitolul 3 „Interacțiunea Om–Calculator pentru Instrumentele de Inginerie și Management al Cerințelor" prezintă o analiză a literaturii privind uzabilitatea și UX în contextul platformelor dedicate REM. Capitolul începe cu o bază teoretică asupra evoluției UX, urmată de o discuție privind diferențele dintre conceptele de uzabilitate și UX, și continuă cu o analiză a metodelor de evaluare a uzabilității, cu accent pe evaluarea euristică. De asemenea, sunt prezentate lucrări relevante privind euristicile specifice domeniului și aplicabilitatea UX pe diferite platforme digitale, în instrumentele REM, precum și utilizarea LLM-urilor în evaluarea uzabilității.

Capitolul 4 „Propunerea de euristici de uzabilitate specifice REM pentru platformele dedicate REM din domeniul auto" cuprinde propunerea de euristici specifice instrumentelor REM, obținute prin aplicarea metodologiei Prometheus [12], destinată dezvoltării euristicilor specifice domeniului. Etapele metodologiei sunt prezentate teoretic și aplicate în studiul de față, ultima etapă fiind detaliată în capitolul următor, ca parte a validării euristicilor.

Capitolul 5 „Validarea euristicilor de uzabilitate specifice Ingineriei și Managementului Cerințelor" prezintă validarea propunerii de euristici REM. Pentru validare, este aplicată etapa finală a metodologiei Prometheus, fiind realizate două evaluări euristice: una utilizând

euristicile lui Nielsen, ca set de control, și alta utilizând euristicile specifice REM, ca set de domeniu. Sunt descrise mediul și procedura de evaluare, iar rezultatele sunt structurate astfel: sinteza tuturor problemelor identificate (prin ambele seturi de euristici), sugestiile evaluatorilor, compararea transversală a problemelor rezultate, indicatorii de calitate calculați și, în final, interpretarea rezultatelor.

Capitolul 6 „Analiză exploratorie: validarea euristicilor prin intermediul LLM-urilor" reprezintă un capitol exploratoriu axat pe evaluarea euristică utilizând LLM-uri. Sunt prezentate motivația acestei abordări, metodologia de evaluare aplicată, precum și rezultatele și discuțiile aferente. Scopul principal este de a explora evaluările euristice bazate pe LLM și de a compara rezultatele cu cele obținute prin evaluări realizate de experți umani.

Capitolul 7 „Contribuții și concluzii" este capitolul final al tezei, concentrat pe prezentarea contribuțiilor și a concluziilor. Acesta include un rezumat al contribuțiilor, o discuție succintă privind contribuțiile teoretice la literatura de specialitate, implicațiile pentru industrie și, în final, direcțiile de cercetare viitoare.

2. Fundamentele Ingineriei Cerințelor în Industria Auto

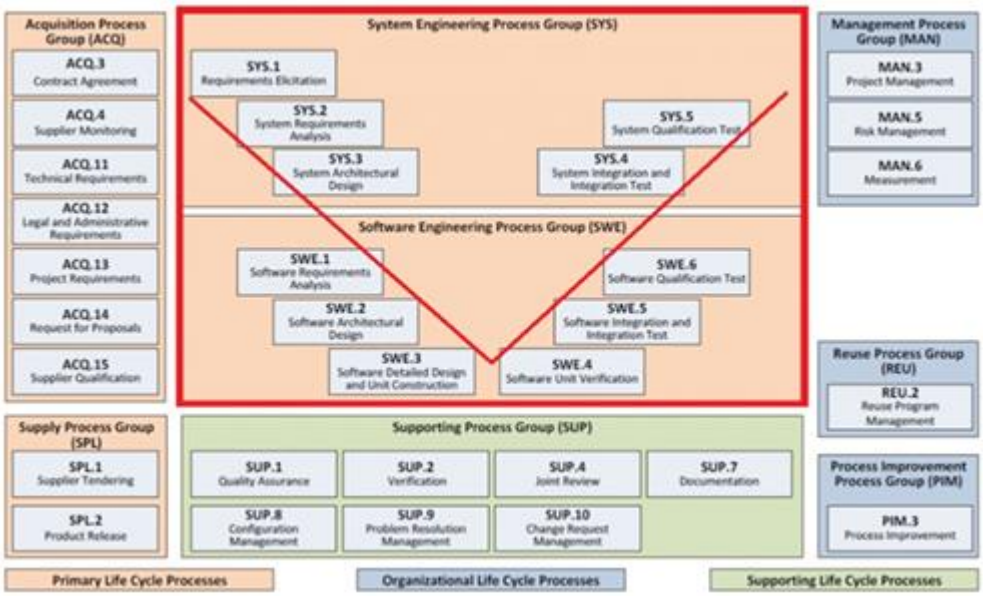


Figura 1. Modelul de referință al procesului Automotive SPICE [13]

Dezvoltarea modernă a sistemelor auto se bazează pe procese bine definite, menite să asigure siguranța, calitatea și conformitatea, fiind susținute de modelul V [14], care stabilește o relație directă între definirea cerințelor și activitățile corespunzătoare de verificare și validare. Abordările de dezvoltare pot include metodologii tradiționale de tip Waterfall [15], care oferă structură și trasabilitate, metodologii Agile, care permit flexibilitate și livrare iterativă, sau modele hibride care combină aceste abordări pentru a răspunde constrângerilor hardware și dinamicii software [16].

În acest context, Ingineria Cerințelor (Requirements Engineering – RE) reprezintă o disciplină fundamentală, responsabilă de elicitarea, analiza, specificarea, validarea și gestionarea cerințelor sistemului pe întreg ciclul de viață [17]. Așa cum este ilustrat în Figura 1, RE asigură alinierea între nevoile părților interesate și implementarea tehnică,

bazându-se pe principii precum claritatea, completitudinea, consistența, trasabilitatea, fezabilitatea și verificabilitatea.

În domeniul auto, Ingineria Cerintelor (RE) este esențială pentru gestionarea sistemelor critice din punct de vedere al siguranței, susținerea conformității cu reglementările și tratarea atât a cerințelor funcționale, cât și a celor nefuncționale în medii complexe. Activitățile RE se extind pe întreg ciclul de dezvoltare, de la definirea conceptului până la proiectarea sistemului, implementare, testare și suport post-producție, fiind strâns aliniate cu structura modelului V [14].

Trasabilitatea joacă un rol fundamental în corelarea cerințelor între diferite niveluri de abstractizare și artefacte ale ciclului de viață, asigurând acoperire, transparență și conformitate, în timp ce managementul schimbărilor oferă fluxuri de lucru structurate pentru gestionarea evoluției cerințelor prin analiza impactului, aprobare și verificare [18].

RE în domeniul auto este puternic reglementată de standarde precum ISO 26262 pentru siguranța funcțională a sistemelor auto [19], ASPICE pentru evaluarea capabilității proceselor [20], ISO/IEC/IEEE 29148 pentru practicile de RE [3] și ISO/SAE 21434 pentru cerințele de securitate cibernetică [21].

Pentru a susține aceste activități, sunt utilizate instrumente specializate de tip REM, care permit redactarea structurată, organizarea ierarhică, trasabilitatea bidirecțională, controlul versiunilor, automatizarea fluxurilor de lucru, colaborarea și integrarea cu alte sisteme de inginerie [22]. Exemple de astfel de instrumente includ IBM DOORS [23], DNG [8], Siemens Polarion [6], Jama Connect [7] și ReqView [24], fiecare oferind niveluri diferite de scalabilitate, colaborare și capacități de integrare. Tabelul 1 prezintă o analiză a unor instrumente REM populare din perspectiva funcționalităților oferite.

Table 1 Analiză comparativa a funcționalităților platformelor destinate REM.

Funcționalitate	**DOORS**	**DNG**	**Siemens Polarion**	**Jama Connect**	**ReqView**
Structurare și organizare ierarhică	Da	Da	Da	Da	Da
Trasabilitate	Da	Da	Da	Da	Da
Versionare și baseline-uri	Da	Da	Da	Da	Da
Managementul configurației	Parțial	Da	Da	Parțial	Parțial
Colaborare și revizuire	Nu	Da	Da	Da	Nu
Integrare cu platforme de testare și MBSE	Parțial	Da	Da	Parțial	Nu
Suport pentru import/export ReqIF	Da	Da	Da	Da	Da
Interfață web	Nu	Da	Da	Da	Parțial

Selecția platformelor de tip REM depinde de complexitatea proiectului, distribuția echipei, maturitatea proceselor, necesitățile de integrare, cerințele de reglementare și considerentele legate de costuri. În ultimii ani, inteligența artificială (AI), în special modelele de tip LLM [25], a apărut ca o tehnologie de suport în domeniul Ingineriei Cerințelor, facilitând activități precum elicitarea cerințelor, clasificarea acestora, evaluarea calității, susținerea trasabilității și analiza impactului schimbărilor. În pofida acestor oportunități, integrarea AI rămâne limitată de provocări precum acuratețea, specificitatea domeniului și riscul generării unor rezultate nesigure.

Ingineria prompturilor (prompt engineering) joacă un rol esențial în valorificarea eficientă a LLM-urilor, implicând proiectarea structurată a intrărilor prin instrucțiuni clare, informații contextuale și constrângeri bine definite, în vederea îmbunătățirii calității rezultatelor. Diferite strategii de formulare a prompturilor, inclusiv abordările zero-shot, few-shot și chain-of-thought, permit interacțiuni mai controlate și orientate către obiective cu sistemele AI [26].

În ansamblu, capitolul evidențiază faptul că Ingineria Cerințelor reprezintă un pilon fundamental al dezvoltării în industria auto, susținut de procese bine structurate, standarde riguroase și instrumente specializate, în timp ce tehnologiile emergente din domeniul inteligenței artificiale introduc noi oportunități pentru creșterea eficienței și a capacităților de analiză.

3. Interacțiunea Om-Computer pentru platformele destinate REM

Eficiența platformelor destinate REM depinde nu doar de funcționalități, ci și de uzabilitate și experiența utilizatorului (UX), care influențează direct eficiența, acuratețea și satisfacția utilizatorului în medii auto complexe [27].

Interacțiunea Om–Computer (Human-Computer Interaction – HCI) oferă fundamentul teoretic pentru înțelegerea UX și a uzabilității, punând accent pe proiectarea centrată pe utilizator și pe metode structurate de evaluare, precum testarea empirică, tehnicile de inspecție și evaluarea euristică. Toate regulile HCI sunt cuprinse în standardul ISO 9241 [9].

Deși euristicile generice oferă o bază solidă, acestea nu dispun de specificitate contextuală pentru sisteme complexe precum instrumentele REM, care implică un grad ridicat de încărcare cognitivă, structuri de trasabilitate, versionare și fluxuri de lucru multidisciplinare [28].

Aceste limitări evidențiază necesitatea unor euristici specifice domeniului, validate prin metodologii structurate, pentru a surprinde mai adecvat problemele de uzabilitate legate de fluxurile de lucru.

Cercetările recente și aplicațiile din diverse platforme digitale demonstrează importanța unor abordări UX adaptate și a combinării metodelor de evaluare, în timp ce inteligența artificială și LLM-urile introduc noi oportunități pentru evaluarea uzabilității, ca instrumente complementare, având în vedere limitările acestora [11].

Studiul [29] subliniază necesitatea extinderii cercetării privind uzabilitatea instrumentelor REM, cu accent pe colaborare, trasabilitate și vizualizare, pentru o mai bună înțelegere a problemelor recurente. O altă cercetare anterioară, care abordează această lacună, analizează evaluarea uzabilității funcționalităților de colaborare ale instrumentelor REM [30]. Studiul prezintă evaluarea uzabilității funcțiilor de colaborare din DNG, într-un departament din industria auto, utilizând metode precum testarea cu utilizatori, focus grupuri și chestionarul SUS. O altă cercetare relevantă abordează analiza și compararea nevoilor specifice rolurilor Agile pentru interfețele platformelor de tip REM, fiind un studiu bazat pe chestionar [31].

În ansamblu, evaluarea eficientă a uzabilității platformelor destinate REM necesită o combinație între euristici specifice domeniului, metodologii structurate și suport oferit de tehnologii AI emergente, pentru a face față complexității sistemelor auto moderne.

4. Propunerea de euristici de uzabilitate specifice REM pentru platformele de tip REM din domeniul auto

Utilizarea metodelor generice de evaluare euristică este adesea insuficientă pentru domenii complexe și specializate, caracterizate prin fluxuri de lucru specifice, terminologie proprie și constrângeri particulare, deoarece aceste euristici oferă doar îndrumări generale și nu reușesc să surprindă problemele profunde de uzabilitate. Deși cadre precum euristicile lui Nielsen [32] sau cele ale lui Gerhardt-Powals [33] oferă o bază valoroasă, acestea necesită adaptare pentru a aborda eficient complexitățile interacțiunii specifice domeniului.

Literatura existentă din domenii precum televiziunea interactivă [34], sistemele medicale [35] și mediile virtuale de învățare [36] demonstrează că euristicile adaptate oferă perspective mai relevante și mai precise asupra uzabilității, în special atunci când sunt dezvoltate utilizând metodologii structurate, cum ar fi Prometheus. Cercetările anterioare realizate în cadrul acestei teze confirmă faptul că euristicile generice sunt eficiente în identificarea problemelor de bază de uzabilitate, în timp ce problemele mai profunde, asociate fluxurilor de lucru specifice domeniului, rămân insuficient abordate, ceea ce evidențiază necesitatea unor cadre de evaluare personalizate.

În sisteme complexe, precum platformele REM din domeniul dezvoltării de sisteme auto, această limitare devine și mai evidentă, din cauza existenței unor activități specializate de inginerie a cerințelor, inclusiv trasabilitatea, validarea, managementul schimbărilor și fluxurile de lucru colaborative [22].

Aceste platforme implică multiple roluri de utilizator, cu responsabilități distincte, ceea ce crește complexitatea interacțiunii și amplifică riscul apariției problemelor de uzabilitate în diferite contexte de utilizare.

Studiile anterioare privind uzabilitatea platformelor de tip REM din domeniul auto, precum DOORS și DNG, evidențiază probleme recurente legate de navigare, complexitatea interfeței, mecanismele de trasabilitate, accesibilitatea funcțiilor de versionare și facilitățile de colaborare [30], [29]. Evaluarea empirică a acestor instrumente arată, de asemenea, că testarea uzabilității, deși eficientă, este consumatoare de resurse, din cauza necesității unor medii reale, a implicării mai multor utilizatori și a utilizării unor scenarii extinse.

Aceste provocări motivează dezvoltarea unor euristici specifice domeniului, ca o abordare mai eficientă și scalabilă pentru evaluarea uzabilității, capabilă să identifice probleme structurale profunde, reducând în același timp dependența de testarea extinsă cu utilizatori.

Pentru a susține acest obiectiv, metodologia Prometheus [37] este adoptată ca un cadru structurat pentru dezvoltarea euristicilor, oferind un proces sistematic, multi-etapizat, care include analiza domeniului, selecția euristicilor, evaluarea specificității, normalizarea, prioritizarea, descrierea, validarea și rafinarea acestora.

În cadrul acestei cercetări, euristicile specifice REM rezultate sunt aplicate într-o evaluare euristică comparativă asupra unui instrument industrial (DNG), implicând evaluatori experți și fiind comparate cu euristicile standard, pentru a evalua eficiența acestora în identificarea problemelor de uzabilitate relevante pentru domeniu.

De asemenea, euristicile propuse — atât cele specifice REM, cât și cele generice — sunt evaluate utilizând un model de tip LLM, ChatGPT-5 [38], pentru a explora capacitatea acestuia de a realiza evaluări structurate de uzabilitate și de a oferi o validare secundară a cadrului propus.

Teza contribuie printr-un cadru nou de evaluare a uzabilității, adaptat platformelor REM din domeniul dezvoltării sistemelor auto, care permite identificarea problemelor de uzabilitate de natură structurală și legate de fluxurile de lucru. Metodologia Prometheus [37] a fost aplicată pentru obținerea setului de euristici specifice REM, iar etapele parcurse sunt prezentate în Figura 2.

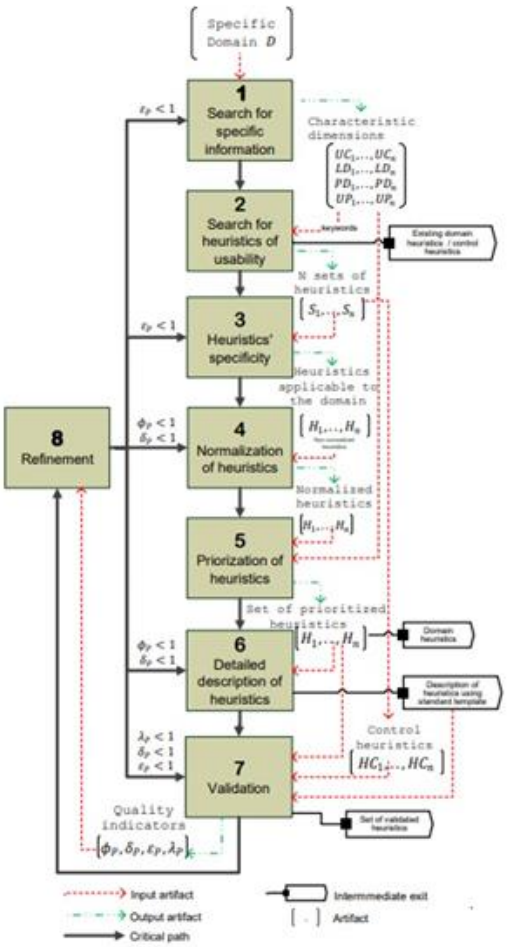


Figura 2. Etapele metodologiei Prometheus [37]

După aplicarea tuturor etapelor, euristicile de uzabilitate specifice REM sunt prezentate după cum urmează:

HREM1. Feedback și starea sistemului

Descriere: Instrumentul REM ar trebui să ofere utilizatorilor feedback clar și contextual și să îi mențină informați cu privire la operațiunile sistemului. Acest lucru poate include schimbări vizibile de stare, mesaje de eroare/succes, indicatori de progres pentru activități precum corelarea artefactelor sau specificarea cerințelor.

HREM2. Consistența interfeței și a terminologiei

Descriere: Interfața instrumentului REM trebuie să fie consistentă în ceea ce privește elementele vizuale, mesajele și terminologia. Termenii utilizați trebuie să fie specifici domeniului REM și adaptați contextual scenariilor de utilizare.

HREM3. Prevenirea erorilor

Descriere: Interfața instrumentului REM trebuie să prevină apariția erorilor în timpul activităților uzuale, precum corelarea artefactelor, editarea, specificarea sau ștergerea.

HREM4. Flexibilitate

Descriere: Interfața instrumentului REM ar trebui să permită utilizatorilor realizarea acțiunilor în moduri multiple, în funcție de nivelul lor de expertiză, preferințe sau nevoi. Interfața ar trebui să ofere scurtături sau vizualizări auxiliare pentru a îmbunătăți flexibilitatea utilizării.

HREM5. Controlul utilizatorului și reversibilitatea

Descriere: În interacțiunea cu instrumentele REM, utilizatorii trebuie să simtă că dețin controlul asupra acțiunilor lor și să poată reveni cu ușurință asupra acestora. Aceste acțiuni pot include modificări ale cerințelor, ștergeri sau actualizări ale legăturilor dintre artefacte.

HREM6. Familiaritatea cu domeniul

Descriere: Interfața instrumentului REM trebuie să susțină utilizarea unui limbaj specific domeniului și fluxuri de lucru de tip RE, așteptate de utilizatorii din domeniu.

HREM7. Recuperarea din erori / diagnosticabilitate

Descriere: Instrumentele REM trebuie să permită recuperarea acțiunilor în cazul unor erori ale utilizatorului. Utilizatorii trebuie să poată înțelege problema, să identifice sursa acesteia (de exemplu, o legătură incorectă) și să remedieze situația cu ușurință, acolo unde este posibil.

HREM8. Reducerea încărcării memoriei

Descriere: Interfața instrumentelor REM trebuie să minimizeze necesitatea memorării informațiilor între diferite ecrane sau interacțiuni. Termenii, funcțiile, widget-urile și stările artefactelor trebuie să fie vizibile și ușor de reținut din interacțiunile anterioare.

HREM9. Secțiune Help și documentație

Descriere: Instrumentele REM trebuie să ofere acces facil la ajutor și documentație. Aceasta include instrucțiuni generale de utilizare, precum și ghiduri specifice domeniului REM pentru acțiuni mai complexe, precum stabilirea trasabilității sau gestionarea versiunilor.

HREM10. Suport pentru colaborare și comunicare

Descriere: Instrumentele REM trebuie să susțină mecanisme de colaborare și comunicare între părțile implicate în activități de RE. Acestea includ funcționalități precum comentarii, etichetarea utilizatorilor, notificări și vizibilitatea modificărilor realizate de alți utilizatori.

HREM11. Design estetic și minimalist

Descriere: Instrumentele REM trebuie să ofere interfețe clare, în care informația este prezentată doar în măsura necesară pentru sarcinile curente, utilizând elemente vizuale și pictograme relevante pentru fluxurile de lucru RE.

HREM12. Suport pentru revizuire și validare

Descriere: Interfețele instrumentelor REM trebuie să ofere suport clar structurat pentru fluxurile de revizuire și validare. Acestea trebuie să distingă vizual etapele de revizuire, să ofere mecanisme clare de aprobare sau respingere și să permită accesul la date relevante de validare (de exemplu, status, comentarii). Aceste funcționalități trebuie să sprijine vizual sesiunile de validare colaborativă.

HREM13. Accesibilitatea versiunilor și a baseline-urilor

Descriere: Interfața instrumentelor REM trebuie să ofere acces facil și intuitiv la artefactele versionate și la baseline-uri. Utilizatorii trebuie să poată compara și interpreta diferite versiuni prin structuri vizuale clare și fluxuri ghidate.

HREM14. Urmărirea istoricului artefactelor

Descriere: Instrumentele REM trebuie să permită utilizatorilor să urmărească cu ușurință istoricul modificărilor pentru fiecare artefact, printr-o structurare clară a informațiilor privind ce s-a modificat, când, de ce și de către cine. Utilizatorii trebuie să aibă acces la cronologii vizuale, filtre și sinteze.

HREM15. Suport pentru navigare și filtrare

Descriere: Interfețele instrumentelor REM trebuie să ofere mecanisme eficiente de navigare și opțiuni de filtrare, pentru a ajuta utilizatorii să opereze asupra artefactelor relevante. Navigarea în cadrul modulelor și artefactelor REM trebuie să respecte logica trasabilității din RE.

Euristicile propuse mai sus, HREM1–HREM15, constituie un cadru adaptat pentru evaluarea uzabilității instrumentelor REM utilizate în industria auto. Setul de euristici este conceput pentru a acoperi aspecte esențiale ale uzabilității, punând accent pe activitățile, funcționalitățile și fluxurile de lucru caracteristice domeniului RE și instrumentelor REM.

Euristicile HREM1–HREM11 sunt derivate din setul normalizat de euristici HN1–HN11 și adaptate domeniului REM, în timp ce HREM12–HREM15 reprezintă euristici noi, care nu erau acoperite de setul normalizat.

Împreună, aceste euristici acoperă provocările de uzabilitate întâlnite în interacțiunea cu instrumentele REM, reflectând complexitatea acestora, colaborarea între multiple roluri și caracterul puternic reglementat al sectorului auto.

În conformitate cu metodologia Prometheus, setul de euristici propus va fi validat în etapa următoare, prezentată în capitolul următor al tezei. Eficiența euristicilor specifice REM va fi evaluată prin comparație cu un set de control, reprezentat de euristicile lui Nielsen. Metodologia Prometheus susține această etapă de validare prin indicatori cantitativi, utilizați ulterior pentru rafinare, dacă este necesar. În plus, vor fi obținute și perspective calitative, menite să completeze evaluarea.

5. Validarea euristicilor de uzabilitate specifice REM propuse

Etapele 7 și 8 ale metodologiei Prometheus se concentrează pe validarea și rafinarea iterativă a euristicilor specifice domeniului REM propuse. Aceste etape sunt esențiale pentru a demonstra relevanța, utilitatea și fidelitatea setului de euristici adaptat atunci când este utilizat în evaluări de uzabilitate în cadrul domeniului.

Etapa 7 are ca obiectiv compararea euristicilor specifice REM propuse cu un set de control, reprezentat de euristicile lui Nielsen, pentru a evalua eficacitatea acestora în identificarea problemelor de uzabilitate. Astfel, validarea este realizată prin metoda evaluării euristice (HE), utilizând atât setul de euristici Nielsen, cât și setul de euristici specifice REM, aplicate asupra unei platforme reprezentative pentru domeniu, DNG.

Evaluarea euristică a fost organizată utilizând un spațiu de test dedicat, configurat în DNG, cu licență corporativă validă și cu aprobarea unui departament al unei companii din industria auto, în cadrul căruia instrumentul este utilizat. Spațiul de test dedicat a fost creat special pentru evaluările euristice, fără a interfera cu alte module de proiect sau configurații utilizate în cadrul companiei.

În cadrul unei configurații locale denumite „Test”, creată pentru evaluare, au fost generate trei module de cerințe. Acestea sunt denumite: „Heuristic Evaluation test document #1”, „Heuristic Evaluation test document #2”, „Heuristic Evaluation test document #3”. Fiecare modul conținea mai multe cerințe, denumite utilizând identificatori structurați, precum 1.1, 2.1, 3.2. Capturile aferente spațiului de test sunt prezentate în Figura 4.

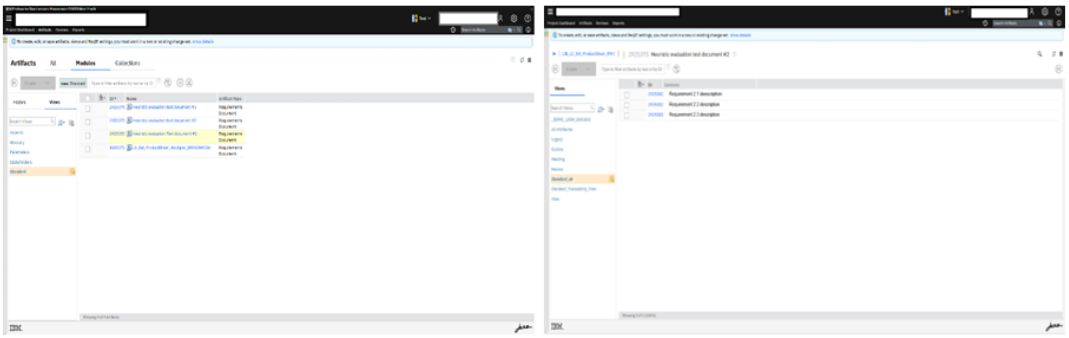


Figura 4. Spațiul de test dedicat din DNG

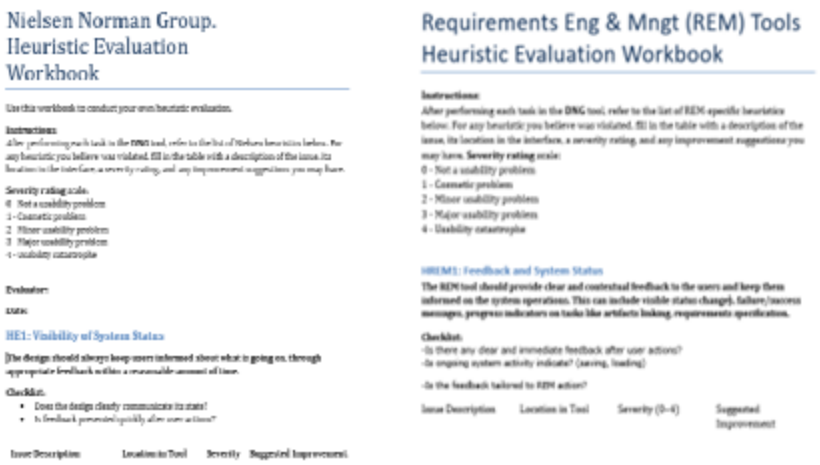


Figura 5. Ghidul de evaluare euristică : stânga-euristici de control; dreapta – euristici REM

Zece evaluatori au fost recrutați pentru realizarea evaluării euristice (HE). Cinci evaluatori au analizat platforma de tip REM utilizând setul de euristici specifice domeniului, iar ceilalți cinci au constituit grupul de control, evaluând instrumentul utilizând setul de euristici de referință. Evaluatorii au fost selectați pe baza experienței relevante în domeniul uzabilității și al evaluării interfețelor. Din punct de vedere al pregătirii, toți evaluatorii au avut cunoștințe solide privind principiile de uzabilitate și HCI.

Pentru a asigura desfășurarea eficientă a evaluării euristice, au fost elaborate și utilizate următoarele materiale: ghidul de evaluare (evaluation workbooks) și un document introductiv privind Ingineria Cerințelor (RE), ale cărui extrase sunt prezentate în Figura 5.

Un set de 6 sarcini structurate a fost dezvoltat pentru sesiunile de evaluare euristică (HE), cu scopul de a activa în mod complet atât euristicile de control, cât și pe cele specifice domeniului. Aceste sarcini au fost concepute pentru a simula interacțiuni realiste ale utilizatorilor în cadrul unui instrument REM — în acest caz, DNG — urmărind să asigure executarea principalelor activități și fluxuri de lucru specifice RE. Sarcinile utilizate în cadrul evaluării euristice sunt prezentate în Tabelul 2.

Pentru ambele sesiuni de evaluare, partea de introducere, care a inclus explicarea configurației și o etapă de explorare liberă a instrumentului, a avut o durată mediană de 15 minute. Evaluarea bazată pe euristicile lui Nielsen a avut o durată medie de aproximativ 75 de minute, în timp ce evaluarea bazată pe euristicile specifice REM a durat în medie aproximativ 105 minute, datorită numărului mai mare de euristici și descrierilor mai detaliate ale acestora.

În urma sesiunilor de evaluare euristică realizate utilizând setul de euristici de control, au fost identificate 30 de probleme unice de uzabilitate, după eliminarea duplicatelor. În urma sesiunilor realizate utilizând setul de euristici specifice REM propuse, au fost identificate 46 de probleme unice de uzabilitate, după eliminarea duplicatelor. Problemele identificate au acoperit toate euristicile — zece în cazul setului de control și cincisprezece în cazul setului specific domeniului — și au fost evaluate din punct de vedere al severității de către evaluatori.

Fiecărei probleme i-a fost atribuit un nivel de severitate cuprins între 0 și 4 (pe o scală Likert cu 5 niveluri), în conformitate cu practicile evaluării euristice. Distribuția severității pentru fiecare sesiune de evaluare este prezentată în Figura 6, iar distribuția severității pe fiecare euristică este ilustrată în Figura 7.

Table 2 Detaliile sarcinilor pentru sesiunile de evaluare euristică.

ID		Task details
1	Denumire sarcini	**Crearea și specificarea unei cerințe noi**
	Precondiții	Este necesar un set de modificări (change set) deschis;
	Pași	Se creează o cerință nouă în „Heuristic evaluation document #1”, denumită „Requirement 1.4”;
		Se setează statusul la „In work” și se adaugă un comentariu.
2	Denumire sarcini	**Corelarea artefactelor și explorarea trasabilității**
	Precondiții	Este necesar un set de modificări deschis
	Pași	În „Heuristic evaluation document #2”, „Requirement 2.1” derivă din „Requirement 1.2” (din documentul #1);
		Se creează legătura între cerințe;
		Se adaugă un comentariu la „Requirement 2.1”;
		Se deschide exploratorul de legături pentru vizualizare.
3	Denumire sarcini	**Editarea unui artefact**
	Precondiții	Este necesar un set de modificări deschis;
	Pași	Se modifică „Requirement 3.1” din „Heuristic evaluation document #3”;
		Se finalizează setul de modificări;
		Se urmărește istoricul modificărilor.
4	Denumire sarcini	**Căutarea și filtrarea cerințelor**
	Precondiții	Este necesar un set de modificări deschis;
	Pași	Se filtrează cerințele din „Heuristic evaluation document #1” după status „Released”;
		Se creează o vizualizare personalizată cu artefactele filtrate.
5	Denumire sarcini	**Accesarea și compararea baseline-urilor**
	Precondiții	În afara unui set de modificări;
	Pași	Se creează un baseline pentru „Heuristic evaluation document #2”;
		Se compară baseline-ul creat cu cel anterior.
6	Denumire sarcini	**Revizuirea și verificarea pentru validare**
	Precondiții	În afara unui set de modificări;
	Pași	Se creează o revizuire pentru „Requirement 1.3” din „Heuristic evaluation document #1”;
		Se finalizează revizuirea;
		Se verifică artefactele de validare asociate.

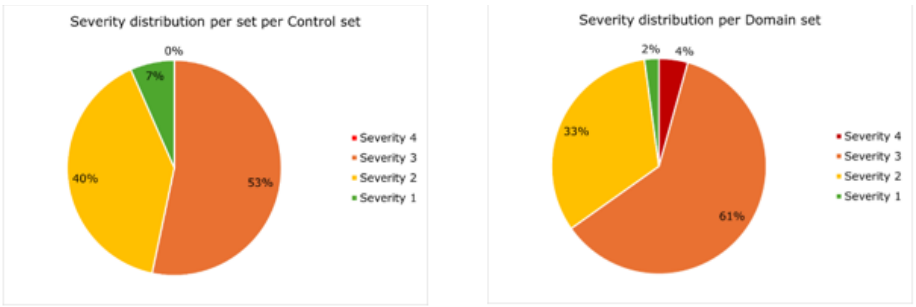


Figura 6. Distribuția severității problemelor de uzabilitate pentru setul de control și setul specific REM

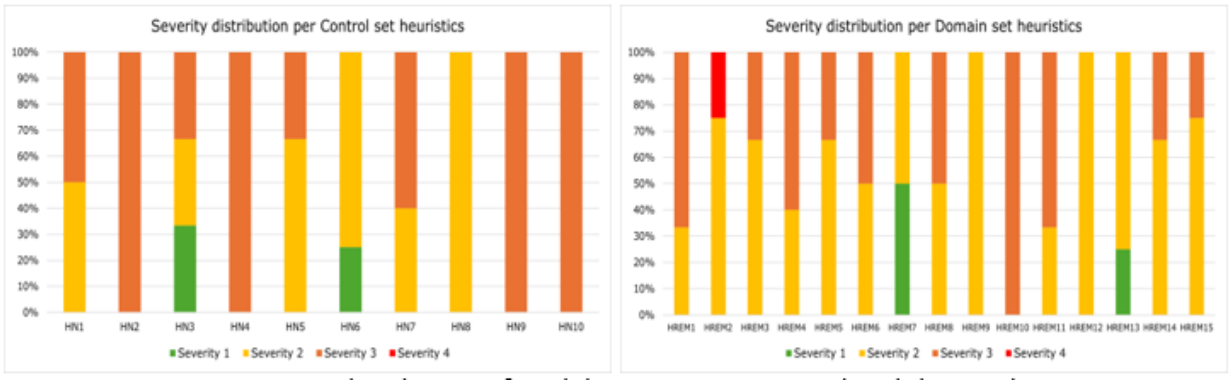


Figura 7. Distribuția severității problemelor de uzabilitate pentru setul de control și setul specific REM

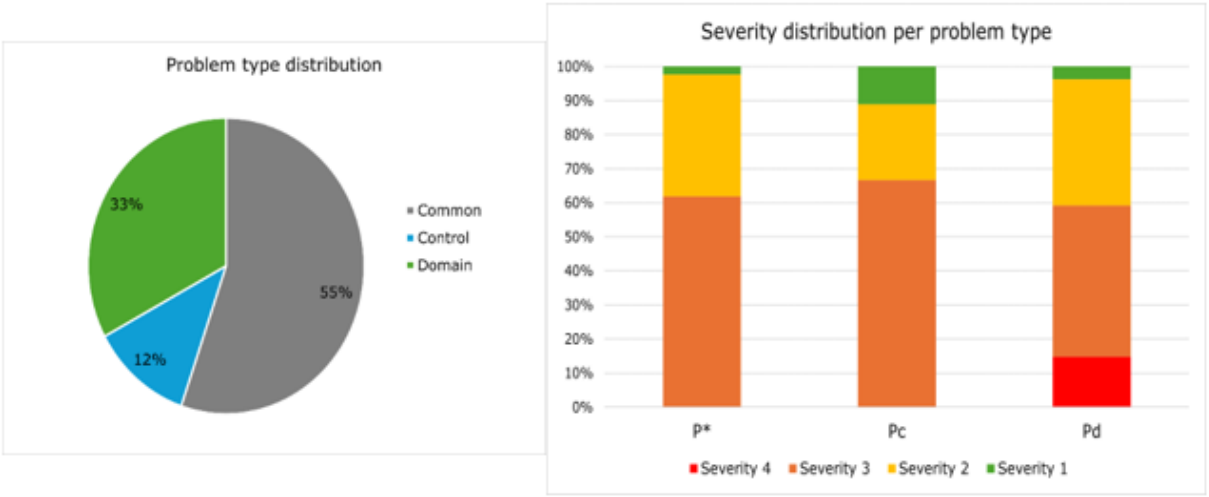


Figura 8. Distribuția tipurilor de probleme identificate în evaluarea euristică și a nivelurilor de severitate corespunzătoare fiecărui tip de problemă

Din totalul problemelor identificate, 42 de probleme au fost comune (21 din fiecare set), fiind detectate utilizând ambele seturi de euristici, notate ca $P*$. De asemenea, 9 probleme au fost identificate exclusiv de grupul de control, utilizând setul de euristici Nielsen, notate ca P_C, în timp ce 25 de probleme au fost identificate în mod unic de euristicile specifice domeniului, notate ca P_D.

Fie $\mathbb{P}$, definit în Ecuația 5.3, mulțimea totală a problemelor identificate prin aplicarea atât a euristicilor de control (P_{C*}), definite în Ecuația 5.2, cât și a euristicilor specifice domeniului (P_{D*}), definite în Ecuația 5.1.

$$P_D^* = P_D \cup P^* = 51 \quad (5.1)$$

$$P_C^* = P_C \cup P^* = 67 \quad (5.2)$$

$$\mathbb{P} = P_D \cup P_C \cup P^* = P_D^* \cup P_C^* = 71 \quad (5.3)$$

Pentru a completa analiza comparativă, rezultatele evaluărilor euristice au fost analizate utilizând indicatori de calitate, după cum urmează:

Rata problemelor unice (Φ_P), care cuantifică problemele identificate exclusiv de setul de euristici specifice domeniului în comparație cu setul de control. Acest indicator este definit în Ecuația 5.4 și evidențiază care set de euristici a identificat un număr mai mare de probleme unice. Dacă $\Phi_P > 1$, aceasta indică faptul că euristicile specifice domeniului au relevat mai multe probleme.

$$\Phi_P = P_D/P_C = P_D^*/P_C^* = 2.78 \quad (5.4)$$

Rata de dispersie a problemelor (δ_P), care măsoară cât de uniform sunt distribuite problemele între euristici, cu scopul de a evita supraîncărcarea sau redundanța acestora, fiind definită în Ecuația 5.5. Distribuția problemelor în ambele grupuri de euristici este analizată prin valorile δ_C și δ_D, care reprezintă abaterea standard a distribuției problemelor pentru euristicile de control, respectiv pentru cele specifice domeniului.

Abaterea standard pentru un set de probleme este definită în Ecuația 5.6, unde *n* reprezintă numărul total de euristici, x_i numărul de probleme identificate utilizând euristica *i*, iar μ_k media numărului de probleme per euristică.

Dacă $\delta_P > 1$, distribuția problemelor în cazul euristicilor specifice domeniului este mai uniformă decât în cazul setului de control.

$$\delta_P = \delta_C/\delta_D = 1.08 \quad (5.5)$$

$$\delta_k = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \mu_k)^2} \quad (5.6)$$

Rata severității problemelor (λ_P), care evaluează dacă euristicile specifice domeniului conduc, în medie, la identificarea unor probleme mai severe. Indicatorul este definit în Ecuația 5.7, unde λ_D reprezintă severitatea medie a problemelor identificate utilizând setul de euristici specifice domeniului, iar λ_C reprezintă severitatea medie a problemelor identificate utilizând setul de control.

În Ecuația 5.8 este calculată severitatea medie, unde S_{k_i} reprezintă scorul de severitate al problemei *i* din setul *k*, iar *n* reprezintă numărul total de probleme luate în considerare (pentru setul respectiv de euristici).

O valoare $\lambda_P > 1$ indică faptul că, în medie, setul de euristici specifice domeniului a condus la identificarea unor probleme mai severe decât setul de control.

$$\lambda_P = \lambda_D/\lambda_C = 1.06 \quad (5.7)$$

$$\lambda_k = \frac{1}{n}\sum_{i=1}^{n} S_{k_i} \quad (5.8)$$

Rata specificității problemelor (ε_P), care măsoară gradul de specificitate a problemelor identificate pentru domeniu, printr-un scor ponderat bazat pe FSI pentru fiecare euristică. Indicatorul este calculat utilizând Ecuația 5.9, unde ε_D reprezintă specificitatea medie a euristicilor specifice domeniului, iar ε_C reprezintă specificitatea medie a setului de control.

În Ecuația 5.10 este detaliată ε_C, care reprezintă media simplă a specificității pentru fiecare problemă. În Ecuația 5.11, ε_D corespunde sumei ponderate a numărului de probleme din fiecare euristică, înmulțit cu scorul FSI asociat acesteia.

Dacă ε_P> 1, atunci problemele identificate prin euristicile specifice domeniului sunt, în medie, mai specifice decât cele identificate prin euristicile de control.

$$\varepsilon_P = \varepsilon_D/\varepsilon_C = 2.58 \quad (5.9)$$

$$\varepsilon_D = \frac{\sum_{H_i}(|P \in H_i^D| * FSI_i)}{|P_D^*|} = 0.77 \quad (5.10)$$

$$\varepsilon_P = \frac{\sum_{i=1}^{n} \varepsilon(P_i)}{|P_C^*|} = 2.58 \quad (5.11)$$

Analiza indicatorilor de calitate furnizată de metodologia Prometheus în etapele 7 și 8 evidențiază relevanța utilizării euristicilor specifice domeniului REM, în comparație cu euristicile generice, precum cele ale lui Nielsen. Indicatori precum **Rata problemelor unice** (Ecuația 5.4), cu o valoare de Φ_P = 2,78, și **Rata specificității problemelor** (Ecuația 5.9), cu o valoare de ε_P = 2,58, s-au dovedit a fi cei mai relevanți. Valorile obținute ilustrează faptul că setul de euristici specifice domeniului a reușit să identifice un număr mai mare de probleme unice și mai bine adaptate domeniului decât setul de control.

Rata severității problemelor (Ecuația 5.7), cu o valoare de λ_P= 1,06, și **Rata de dispersie a problemelor** (Ecuația 5.8), cu o valoare de δ_P= 1,08, sugerează că diferențele dintre cele două seturi sunt relativ reduse în ceea ce privește severitatea medie a problemelor și distribuția acestora între euristici.

Analiza indicatorilor de calitate confirmă faptul că euristicile specifice domeniului propuse sunt eficiente nu doar în identificarea majorității problemelor generale de uzabilitate, ci și în detectarea unor probleme de uzabilitate cu impact ridicat, dependente de context și strâns legate de fluxurile de lucru din RE, oferind astfel o evaluare euristică mai țintită.

Deși euristicile lui Nielsen rămân eficiente pentru identificarea problemelor fundamentale de uzabilitate, setul propus de euristici specifice REM oferă perspective mai profunde și mai aplicabile în medii REM complexe. O concluzie posibilă este că o abordare complementară, care combină ambele cadre în evaluarea instrumentelor REM, ar asigura atât identificarea problemelor generale, cât și captarea provocărilor specifice domeniului.

Având în vedere criteriile oferite de metodologia Prometheus pentru analiza indicatorilor de calitate, nu sunt necesare rafinări ale setului actual de euristici specifice REM. Deși ar putea fi realizate ajustări minore în ceea ce privește echilibrul între severitate și dispersia problemelor, acestea nu sunt considerate critice, motiv pentru care nu se vor efectua rafinări în cadrul prezentei teze.

6. Analiză exploratorie: validarea euristicilor prin intermediul LLM-urilor

Studiul exploratoriu a fost conceput pentru a evalua capacitatea unui model generativ de inteligență artificială (LLM) — ChatGPT 5 — de a realiza evaluări euristice ale unui instrument REM, utilizând atât euristicile generice (Nielsen), cât și euristicile specifice REM propuse.

În acest context, scenariul de evaluare este descris după cum urmează: au fost realizate două sesiuni de evaluare separate, fiecare corespunzând unui set de euristici, aplicate asupra acelorași șase sarcini structurate utilizate anterior în evaluările realizate de

evaluatori umani. LLM-ul a fost ghidat printr-o abordare structurată de tip prompt, care combină definirea rolului, raționamentul de tip chain-of-thought și formatarea predefinită a rezultatelor, pentru a asigura obținerea unor rezultate consistente și comparabile [26].

Fiecare sarcină a fost evaluată independent, utilizând prompturi care includeau rolul evaluatorului, obiectivele, pașii de raționament, setul de euristici, descrierea sarcinii, contextul spațiului de lucru și capturi de ecran relevante ale interfeței utilizator. Mediul de evaluare a reprodus fidel configurația testelor cu utilizatori umani, utilizând același instrument REM (DNG), aceleași sarcini și scenarii de interacțiune comparabile, cu limitarea că evaluarea s-a bazat pe capturi de ecran statice, și nu pe interacțiuni reale.
Fiecare sarcină a fost evaluată individual, utilizând un șablon standardizat de prompt, formulat după cum urmează:
„**Rol:** Presupune că ești un expert în uzabilitate care evaluează un instrument de Inginerie și Management al Cerințelor din domeniul auto (IBM DOORS Next Generation), prin metoda evaluării euristice.
Obiectiv: Identificarea problemelor de uzabilitate pe baza setului de euristici furnizat, evaluarea severității acestora pe o scală de la 1 la 4 și propunerea unor îmbunătățiri.
Raționament: Înainte de a furniza răspunsurile, analizează interfața și parcurge interacțiunea pas cu pas:

1. Observă fiecare captură de ecran pentru a înțelege contextul complet.
2. Identifică elementele interfeței utilizator, fluxurile potențiale și terminologia utilizată.
3. Compară observațiile realizate cu setul de euristici furnizat pentru a identifica posibile încălcări ale acestora.
4. Pentru fiecare problemă identificată, analizează motivul pentru care reprezintă o problemă și impactul asupra uzabilității.
5. Atribuie un nivel de severitate problemei identificate.
6. Propune o îmbunătățire care să abordeze problema identificată.

Set de euristici: Nielsen / REM-specific.
Descrierea sarcinii: Sarcina 1 … Sarcina 6.
Contextul spațiului de lucru: A fost creat un spațiu de lucru de tip „Test” în instrumentul REM pentru această evaluare euristică. Spațiul conține trei module de cerințe, fiecare incluzând un număr redus de artefacte.
Capturi de ecran: sunt incluse capturi relevante pentru sarcina curentă.
Formatul rezultatului: Euristica încălcată | Descrierea problemei | Severitate (1–4) | Îmbunătățire propusă”

Studiul a fost realizat în condiții controlate, utilizând câte o singură rulare pentru fiecare sarcină, fără regenerarea răspunsurilor, reflectând astfel un mod realist de utilizare a LLM-ului și asigurând consistența rezultatelor pentru analiză.

Evaluarea euristică realizată de LLM utilizând euristicile lui Nielsen a identificat un total de 37 de probleme de uzabilitate, în timp ce evaluarea bazată pe euristicile specifice REM a identificat un total de 30 de probleme de uzabilitate. Distribuția problemelor raportate este prezentată în Figura 9 și include: probleme comune identificate și de evaluatori umani, probleme noi, care nu au fost identificate de evaluatori umani, precum și probleme potențiale sau nevalidate.

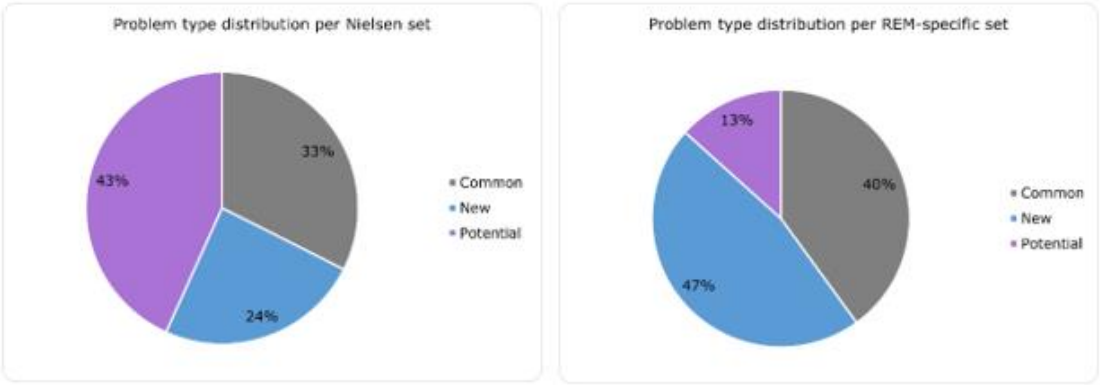


Figura 9. Distribuția tipurilor de probleme identificate în evaluarea euristică bazată pe LLM

După cum se observă în distribuția problemelor prezentată în Figura 9, procentul mai redus, de 13%, al problemelor nevalidate sau potențiale, comparativ cu evaluarea bazată pe euristicile lui Nielsen, poate sugera că euristicile specifice domeniului oferă un cadru mai contextual pentru evaluările bazate pe LLM, reducând generarea de rezultate ipotetice (halucinații).

În ceea ce privește severitatea, problemele identificate utilizând setul de euristici Nielsen au fost încadrate în nivelurile de severitate 2 și 3; 41% dintre probleme au fost percepute ca fiind majore, iar 59% ca fiind minore, fără ca vreo problemă să fie clasificată drept cosmetică sau critică. Distribuția severității la nivelul întregului set de euristici este prezentată în Figura 10.

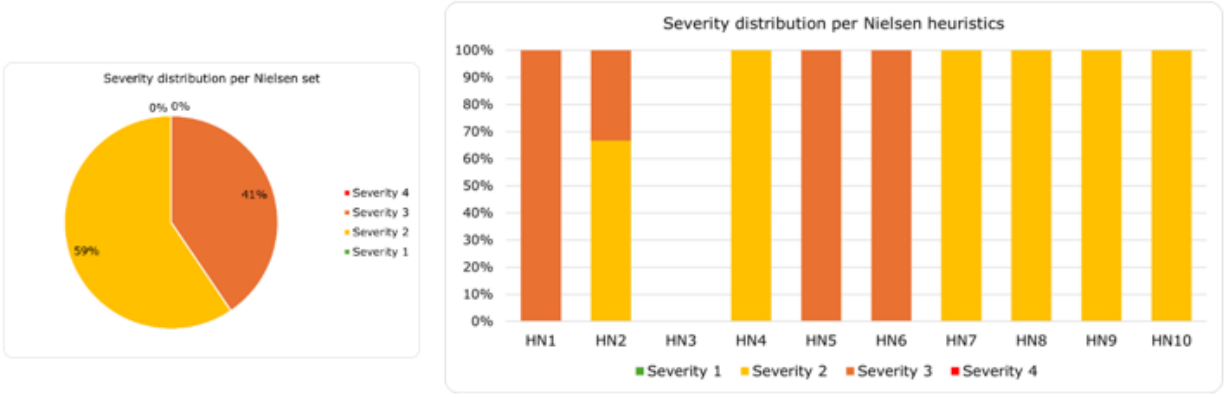


Figura 10. Distribuția severității pentru evaluarea euristică bazată pe LLM, utilizând euristicile lui Nielsen.

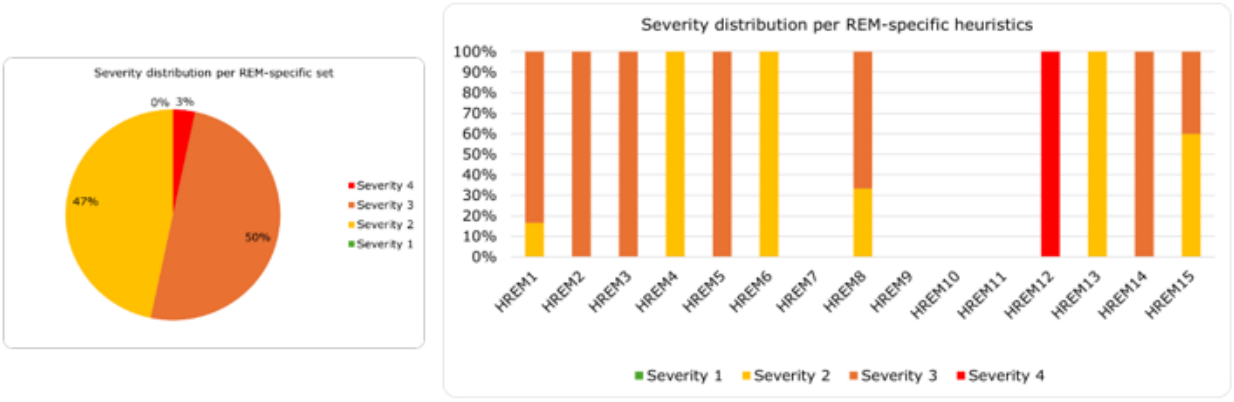


Figura 11. Distribuția severității problemelor în cadrul evaluării euristice bazate pe LLM, utilizând euristicile propuse specifice REM

În ceea ce privește severitatea, problemele identificate utilizând setul de euristici specifice REM au fost încadrate în niveluri de severitate de la 2 la 4. Din totalul de 30 de probleme, 50% au fost considerate majore, 47% minore și 3% critice. Distribuția severității pentru întregul set de euristici specifice REM este prezentată în Figura 11.

Acest studiu exploratoriu a evidențiat faptul că modelele generative de inteligență artificială precum ChatGPT-5, pot identifica un număr semnificativ de probleme utilizând ambele seturi de euristici, existând o suprapunere considerabilă cu problemele identificate de evaluatori umani.

Atunci când a fost ghidat de euristici generice, modelul a identificat în principal probleme de suprafață, legate de mecanismele de feedback, terminologie și claritatea interfeței. În schimb, atunci când au fost utilizate euristicile specifice REM, modelul a reușit să identifice probleme mai contextuale și legate de fluxurile de lucru, ceea ce indică faptul că ghidarea contextuală îmbunătățește fiabilitatea modelului și relevanța problemelor identificate.

Evaluările euristice bazate pe LLM prezintă însă limitări, în special din cauza lipsei interacțiunii reale cu sistemul, ceea ce poate conduce la generarea de rezultate ipotetice sau la apariția fenomenului de „halucinație" al modelului.

În general, LLM-urile pot reprezenta un instrument auxiliar puternic în evaluările euristice, fiind capabile să genereze perspective suplimentare și sugestii de îmbunătățire. Totuși, având în vedere limitările acestora, LLM-urile nu pot înlocui expertiza umană, evaluarea bazată pe interacțiune directă rămânând esențială.

7. Concluzii și Contribuții

Cercetarea pentru această teză a fost realizată în cadrul Departamentului de Comunicații, Facultatea de Electronică, Telecomunicații și Tehnologii Informaționale, Universitatea Politehnica Timișoara. Aceasta contribuie la domeniul doctoral Tehnologia Informației, situându-se la intersecția dintre Ingineria Software (SW), cu accent pe Ingineria Cerințelor (RE), și Interacțiunea Om–Calculator (HCI) [17], [39]. Teza abordează evaluarea uzabilității platformelor software REM și a mediilor complexe de inginerie software necesare pentru susținerea activităților de RE standardizate.

Lucrarea extinde principiile generale ale HCI la cazul platformelor software REM. În mod specific, teza poate fi sintetizată ca o propunere de cadru pentru evaluarea euristică (HE), adaptată instrumentelor REM utilizate în domeniul auto. Setul de euristici specifice REM

propus a fost validat, demonstrând o bază mai orientată și contextualizată a principiilor de uzabilitate pentru evaluarea platformelor software specializate.

În ultimii ani, tema euristicilor specifice domeniului a atras o atenție crescândă în domeniul HCI. Deși regulile lui Ben Shneiderman au fost publicate încă din 1986 [40], iar euristicile lui Nielsen în 1994 [32], dezvoltarea cadrelor euristice rămâne un subiect de interes actual. În ultimul deceniu, au fost propuse mai multe metodologii pentru dezvoltarea sistematică a euristicilor specifice domeniului, printre cele mai cunoscute fiind metodologia Prometheus (Jimenez et al., 2017 [37]) și metodologia propusă de Rusu et al. (2011 [41]). Acestea au fost aplicate cu succes în domenii precum medii virtuale de învățare [36], aplicații de realitate virtuală [42], Grid Computing [43] sau televiziune digitală interactivă [34].

În ultimii ani, s-a observat un interes crescut pentru evaluările euristice bazate pe LLM-uri. Studiile arată că LLM-urile pot acționa ca evaluatori euristici, având capacitatea de a funcționa ca asistenți cognitivi și de a evalua uzabilitatea pe baza unor euristici definite [26]. Totuși, acestea sunt considerate instrumente auxiliare, utile pentru susținerea cercetării calitative și a evaluărilor din etapele incipiente.

Euristicile specifice domeniului propuse în această teză au fost validate prin două sesiuni de evaluare euristică: una utilizând setul generic (Nielsen) și una utilizând setul propus pentru REM. Au fost implicați zece evaluatori, iar analiza s-a bazat pe experiența acestora cu instrumentul DNG. Analiza indicatorilor de calitate obținuți a demonstrat că euristicile specifice domeniului sunt mai eficiente atât în identificarea problemelor generale de uzabilitate (suprapuse cu cele identificate prin euristici generice), cât și în identificarea problemelor contextuale, legate de fluxuri de lucru specializate din RE, esențiale în domeniul auto.

În plus, a fost explorată utilizarea instrumentelor AI în domeniul HCI. După validarea euristicilor, a fost realizată o analiză exploratorie utilizând LLM-uri ca potențiali evaluatori euristici [44]. Două evaluări euristice au fost realizate utilizând ChatGPT-5, modelul fiind ghidat printr-un șablon standard, aceleași sarcini și capturi de ecran statice. Rezultatele au evidențiat atât potențialul LLM-urilor, cât și limitările acestora. LLM-urile s-au dovedit utile ca instrumente complementare, dar nu pot înlocui evaluarea umană.

Pe baza rezultatelor obținute, au fost formulate răspunsuri la întrebările de cercetare:

1. **Cum se reflectă considerentele HCI în evaluarea instrumentelor REM?**
   Considerentele HCI sunt relevante pentru instrumentele REM, în special în ceea ce privește uzabilitatea, navigarea, feedbackul, prevenirea erorilor și încărcarea cognitivă, dar includ și aspecte specifice precum colaborarea și vizualizarea trasabilității.
2. **Sunt euristicile generice eficiente pentru probleme specifice domeniului?**
   Acestea sunt eficiente pentru probleme generale de uzabilitate, dar limitate în identificarea problemelor contextuale.
3. **Cât de eficiente sunt euristicile propuse, specifice REM?**
   Acestea identifică atât problemele generale, cât și probleme mai profunde, asociate fluxurilor de lucru.
4. **Cât de eficient este un LLM în evaluări euristice?**
   LLM-urile pot identifica probleme relevante, dar prezintă limitări și necesită validare umană.

Din perspectivă practică, concluziile tezei au implicații importante pentru industria auto. Instrumentele REM sunt componente critice ale infrastructurii digitale utilizate în dezvoltarea sistemelor critice din punct de vedere al siguranței. Deficiențele de uzabilitate pot genera riscuri inginerești, cum ar fi pierderea trasabilității sau interpretarea incorectă a cerințelor.

Din perspectivă academică, teza contribuie prin integrarea HCI în contextul sistemelor software de inginerie auto și evidențierea necesității alinierii principiilor UX cu fluxurile inginerești și standardele relevante.

În concluzie, lucrarea conectează domeniile Ingineriei Software și HCI, analizând impactul uzabilității și al AI asupra evaluării sistemelor complexe. Rezultatele evidențiază necesitatea utilizării cadrelor de evaluare specifice domeniului în contextul evoluției industriei auto.

Contribuții teoretice:

1. Cadru teoretic pentru RE în industria auto
2. Clasificarea funcțională a instrumentelor REM
3. Integrarea HCI în sistemele REM
4. Introducerea euristicilor specifice domeniului auto
5. Cercetare asupra utilizării LLM-urilor în evaluarea uzabilității

Contribuții practice:

1. Propunerea euristicilor REM-specifice
2. Aplicarea metodologiei Prometheus
3. Cadru comparativ de evaluare euristică
4. Validare prin indicatori cantitativi și calitativi
5. Studiu exploratoriu LLM

Bibliografie

[1] M. Vogel et al., “Metrics in automotive software development: A systematic literature review,” *J. Softw. Evol. Process*, vol. 33, no. 2, p. e2296, 2021, doi: 10.1002/smr.2296.
[2] E. Tüzün, B. Tekinerdogan, Y. Macit, and K. İnce, “Adopting integrated application lifecycle management within a large-scale software company: An action research approach,” *J. Syst. Softw.*, vol. 149, pp. 63–82, Mar. 2019, doi: 10.1016/j.jss.2018.11.021.
[3] ISO/IEC/IEEE 29148:2018 - Systems and software engineering — Life cycle processes — Requirements engineering, 2, 2018. Accessed: Aug. 23, 2025. [Online]. Available: https://www.iso.org/standard/72089.html
[4] ISO/IEC/IEEE 15288:2023(en), Systems and software engineering — System life cycle processes. Accessed: May 05, 2026. [Online]. Available: https://www.iso.org/obp/ui/en/#iso:std:iso-iec-ieee:15288:ed-2:v1:en
[5] ISO/IEC/IEEE 12207:2026(en), Systems and software engineering — Software life cycle processes. Accessed: May 05, 2026. [Online]. Available: https://www.iso.org/obp/ui/en/#iso:std:iso-iec-ieee:12207:ed-2:v1:en
[6] Siemens Product Lifecycle Management Software Inc., “Requirements Management - Polarion REQUIREMENTS.” Accessed: Sep. 03, 2025. [Online]. Available: https://polarion.plm.automation.siemens.com/products/polarion-requirements
[7] Jama Software, “Jama Connect.” Accessed: May 13, 2025. [Online]. Available: https://www.jamasoftware.com/platform/jama-connect/
[8] S. SAS, “IBM Engineering Requirements Management DOORS Next.” Accessed: May 20, 2025. [Online]. Available: https://www.sodiuswillert.com/en/ibm-elm/ibm-engineering-requirements-management-doors-next
[9] ISO 9241-11:2018, Ergonomics of human-system interaction — Part 11: Usability: Definitions and concepts. Accessed: Aug. 21, 2025. [Online]. Available: https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-2:v1:en
[10] J. Nielsen, “Usability 101: Introduction to Usability,” Nielsen Norman Group. Accessed: Oct. 08, 2025. [Online]. Available: https://www.nngroup.com/articles/usability-101-introduction-to-usability/
[11] M. Aubin Le Quéré et al., “LLMs as Research Tools: Applications and Evaluations in HCI Data Work,” in *CHI EA ’24*, New York: ACM, May 2024, pp. 1–7. doi: 10.1145/3613905.3636301.
[12] C. Jimenez, H. Allende Cid, and I. Figueroa, “PROMETHEUS: Procedural Methodology for Developing Usability Heuristics,” *IEEE Lat. Am. Trans.*, vol. 15, no. 3, pp. 541–549, Mar. 2017, doi: 10.1109/TLA.2017.7867606.
[13] J. C. Bebel et al., “Managing complexity in an autonomous vehicle,” *IEEE Aerosp. Electron. Syst. Mag.*, vol. 23, no. 3, pp. 3–13, 2008.
[14] “V-Model - an overview | ScienceDirect Topics.” Accessed: Aug. 21, 2025. [Online]. Available: https://www.sciencedirect.com/topics/engineering/v-model
[15] S. Pargaonkar, “A Comprehensive Research Analysis of SDLC Agile & Waterfall Model,” *Int. J. Sci. Res. Publ.*, vol. 13, no. 8, pp. 120–124, Aug. 2023.

[16] K. Berger, “Adoption of Agile methods in automotive software development.” Accessed: Jan. 19, 2025. [Online]. Available: https://essay.utwente.nl/81865/
[17] F. Flores et al., “Requirements Engineering: A Review of Processes and Techniques,” 2009, doi: 10.4018/978-1-59904-887-1.ch006.
[18] O. Gotel et al., “Traceability Fundamentals,” in *Software and Systems Traceability*, 2012, pp. 3–22. doi: 10.1007/978-1-4471-2239-5_1.
[19] ISO 26262-8:2018, Road vehicles — Functional safety — Part 8. Accessed: May 15, 2025. [Online]. Available: https://www.iso.org/standard/68386.html
[20] M. Knop, “An ASPICE Overview.” Accessed: May 04, 2025. [Online]. Available: https://www.suse.com/c/an-aspice-overview/
[21] ISO/SAE 21434:2021, Road vehicles — Cybersecurity engineering. Accessed: May 19, 2025. [Online]. Available: https://www.iso.org/standard/70918.html
[22] J. M. Carrillo de Gea et al., “Requirements engineering tools: Capabilities, survey and assessment,” *Inf. Softw. Technol.*, vol. 54, no. 10, pp. 1142–1157, Oct. 2012.
[23] S. SAS, “IBM Engineering Requirements Management DOORS.” Accessed: May 20, 2025. [Online]. Available: https://www.sodiuswillert.com/en/ibm-elm/ibm-doors
[24] Eccam, “ReqView.” Accessed: May 13, 2025. [Online]. Available: https://www.reqview.com/
[25] M. A. K. Raiaan et al., “A Review on Large Language Models,” *IEEE Access*, vol. 12, pp. 26839–26874, 2024.
[26] P. Sahoo et al., “A Systematic Survey of Prompt Engineering,” arXiv, 2025. doi: 10.48550/arXiv.2402.07927.
[27] M. Körber et al., “User experience evaluation in an automotive context,” 2013.
[28] H. Gaspard-Boulinc and S. Conversy, “Usability Requirements for Requirement Engineering Tools,” 2014.
[29] O. Rotaru, S. Vert, and R. Vasiu, “Switching from DOORS to DNG,” Springer, 2024.
[30] O. Rotaru et al., “Usability Evaluation of Requirement Collaboration Features,” ICEIS 2025.
[31] O. Rotaru et al., “Analyzing Agile role-specific needs for REM tools,” to be published.
[32] J. Nielsen, “10 Usability Heuristics for UI Design.”
[33] J. Gerhardt-Powals, “Cognitive engineering principles,” *Int. J. HCI*, 1996.
[34] A. Solano et al., “Usability heuristics for interactive digital television,” 2011.
[35] J. Zhang et al., “Using usability heuristics to evaluate patient safety,” *J. Biomed. Inform.*, 2003.
[36] I. Figueroa et al., “Developing usability heuristics with PROMETHEUS,” *Comput. Stand. Interfaces*, 2019.
[37] C. Jimenez et al., “PROMETHEUS methodology,” arXiv, 2018.
[38] OpenAI, “GPT-5 System Card,” 2025.
[39] A. R. Mathew et al., “HCI: An overview,” IEEE, 2011.
[40] E. Wong, “Shneiderman’s Eight Golden Rules.”
[41] C. Rusu et al., “A Methodology to establish usability heuristics,” 2011.
[42] A. Mohammad and L. Pedersen, “Heuristics in VR learning,” *Informatics*, 2022.
[43] D. Quiñones et al., “Developing usability heuristics for Grid Computing,” Springer, 2016.
[44] G. Guerino et al., “Can GPT-4o Evaluate Usability Like Human Experts?” 2025.