

Tendinte actuale în evaluarea interfetelor om-calculator

Ec.dr. Costin PRIBEANU
Institutul de Cercetari în Informatica

Creșterea interesului pentru evaluarea interfetelor om-calculator a condus la o integrare mai bună a acestei activități în ciclul de viață a unui produs informatic. Acest articol prezintă principalele tendințe care se manifestă în metodologie, în organizarea evaluării la nivelul proiectului și în strategia de evaluare la nivel de firmă.

Cuvinte cheie: interacțiune om-calculator, evaluare, ingineria utilizabilității.

1. Introducere

Orientarea proiectării produselor informatice spre om a condus la o nouă abordare a proiectării interfeței cu utilizatorul, care pune în prim plan cerințele ergonomice și cognitive de utilizare ale acestora într-un context dat. Din acest motiv, criteriile de evaluare ale produselor informatice sunt cele de evaluare a unui instrument sau, mai general, a unui artefact, care trebuie să servească scopului pentru care a fost achiziționat.

Interesul pentru evaluarea interfetelor om-calculator a crescut considerabil în deceniul actual, odată cu cristalizarea conceptelor asociate interacțiunii om-calculator, referita cu acronimul HCI (Human-Computer Interaction), care constituie o reconsiderare, dintr-o perspectivă multidisciplinară, a conceptelor din disciplinele premergătoare: ingineria factorilor umani, ergonomia și ingineria interfetelor om-calculator.

Un termen asociat tot mai frecvent cu evaluarea produselor informatice este *utilizabilitatea*, definită de standardul ISO 9241-11 (1994) ca eficacitate, eficiență și satisfacție cu care un utilizator specificat realizează sarcini de lucru specificate. Din punctul de vedere al utilizatorului, ceea ce diferențiază două produse informatice care rezolvă, în principiu, aceleași probleme, este interfata om-calculator, adică modul în care acesta comunică cu calculatorul.

Această perspectivă conceptuală nouă a determinat o reconsiderare a evaluării interfetelor om-calculator, marcată prin apariția de noi strategii și abordări, care ar putea fi caracterizate astfel:

- crearea unui cadru participativ de dezvoltare și evaluare prin atragerea utilizatorilor ca participanți activi la acest proces;
- reconsiderarea resurselor umane și tehnice necesare la nivel de firmă pentru a răspunde cerințelor multidisciplinare ale procesului de interacțiune;
- noi metode și tehnici, care denotă o abordare realistă și pragmatică în raport cu bugetul unui proiect și disponibilitățile unei firme producătoare;
- posibilitatea evaluării în diferite etape ale ciclului de viață al unui produs, ca reflex al nevoii de dezvoltare iterativă.

Scopul acestui articol este de a face o trecere în revista a problematicei evaluării în general și a tendințelor care se manifestă în prezent în acest domeniu. După o succintă enumerare a obiectivelor evaluării și a principalelor strategii de evaluare se prezintă problematica asociată inspecțiilor de utilizabilitate, cu un accent pe evaluarea euristica.

Această alegere a fost determinată atât de interesul actual manifestat pentru aceste metode cât și de paleta largă a preocupărilor (criterii de evaluare, organizare la nivel de proiect și firmă) care face din expertiza de utilizabilitate abordarea cea mai reprezentativă pentru evoluția și tendințele actuale în evaluarea interfetelor om-calculator.

2. Abordări în evaluarea interfetelor om-calculator

După Dix et al. (1993) evaluarea are trei obiective de bază:

- evaluarea functionalitatii sistemului: utilizatorul trebuie sa poata efectua sarcina de lucru mai usor;
- evaluarea efectului pe care îl are interfata asupra utilizatorului: se are în vedere usurinta de a învăta, usurinta de a utiliza si atitudinea fata de sistem;
- identificarea unor probleme specifice legate de sistem: se au în vedere aspectele negative, care produc confuzii sau rezultate neasteptate.

Autorii identifica doua stiluri de evaluare: evaluarea în conditii de laborator si evaluarea în contextul utilizarii. Un aspect legat de evaluarea în contextul utilizarii îl constituie crearea unui cadru participativ de proiectare si evaluare.

Johnson (1992) face o structurare a problemelor evaluarii pe doua etape ale ciclului de realizare: specificare proiect si implementare, aratând ca la nivelul etapei de specificare problemele evaluarii sunt înca subiect de cercetare. Este importanta, întrucât ofera proiectantului cai de evaluare a proiectului înainte de alocarea resurselor pentru scrierea programelor si implementare. Unele metode de specificare pot ajuta în prevederea unui grad de utilizabilitate si usurinta în învătare.

Coutaz (1995) propune o abordare a evaluarii interfetelor om-calculator din perspectiva ingineriei software si sustine necesitatea de a explora intersectia acestui domeniu cu cel al interactiunii om-calculator. În acest sens, calitatea sistemelor software poate constitui un concept central.

Metodologia de evaluare este diversa: se pot distinge metode analitice cum sunt GOMS (Card, Moran & Newell, 1991) sau TAG (Payne & Green, 1985) si metode empirice (observare informala, chestionare, interviuri), consultanta oferita de specialisti în HCI si proiectarea participativa bazata pe instrumente de prototipizare rapida.

O clasificare mai cuprinzatoare a metodelor de evaluare, propusa de Whitefield et al (1991), este urmatoarea: metode analitice (GOMS, TAG); raportul unui expert; raportul utilizatorului (chestionare, interviuri); metode experimentale.

Coutaz (1995) distinge doua abordari în evaluarea teoretica: tehnici bazate pe teorie si tehnici euristice, bazate pe proprietati cunoscute ca având efecte pozitive asupra utilizabilitatii. Evaluatorul cauta sa detecteze asemenea trasaturi în proiectul interfetei.

O tendinta actuala în evaluarea interfetelor om-calculator este reconsiderarea metodologiei în scopul gasirii unor solutii eficace, din punctul de vedere al cerintelor de calitate, dar si eficiente, realiste, care sa conduca la o implementare efectiva a ingineriei utilizabilitatii în procesul de dezvoltare al sistemelor informatice. Un exemplu în acest sens este proliferarea unor metode care se încadreaza în categoria inspectiilor de utilizabilitate.

3. Evaluare în cadru participativ

Orientarea tehnologiei software spre cerintele societatii actuale necesita luarea în considerare a unui cadru de lucru extins, care sa asigure participarea tuturor factorilor interesati în proiectare si evaluare. Evaluarea în cadru participativ determina atât o orientare catre factorul uman dar si o orientare catre activitate si context. De asemenea, aceasta abordare conduce la un proces iterativ, prin care proiectul devine subiect de evaluare si îmbunatatire în fiecare etapa.

O modalitate de extindere este atragerea utilizatorului în acest proces, care are ca scop atât eficientizarea dezvoltarii, întrucât permite identificarea problemelor de utilizabilitate în fazele incipiente, atunci când costurile corectarii sunt reduse, cât si apropierea proiectantului de perspectiva conceptuala specifica profesiei si activitatii celui care va utiliza produsul.

O alta tendinta, determinata de caracterul interdisciplinar al abordarii procesului de interactiune, este atragerea în echipa de evaluare si, dupa unii autori, mai radicali chiar în echipa de dezvoltare, a unor specialisti în sociologie, psihologie, tipografie si design industrial. În acest sens, Norman (1998) arata ca în prezent, datorita conservatorismului firmelor producatoare, aportul unor specialisti din domeniile amintite are

numai un rol cosmetic, fapt care conduce la o utilizabilitate scazuta a tehnologiei informatice.

4. Inspectia de utilizabilitate

Modelele analitice de evaluare, bazate pe o analiza formala, nu au atins încă nivelul de dezvoltare necesar pentru utilizarea în proiecte reale, iar volumul de munca necesar pentru elaborarea unei specificatii este considerabil. De aceea, în practica curenta cel mai frecvent a fost utilizata evaluarea experimentală. Din pacate, în multe cazuri nici aceasta metoda nu este la îndemâna din diferite motive cum sunt lipsa specialistilor care sa stie sa conduca un experiment sau timpul insuficient.

Pe de alta parte, utilizarea metodelor analitice, aplicabila în diferite etape ale ciclului de viata, cade, aproape exclusiv, în sarcina proiectantilor care sunt tentati sa repete, în evaluare, aceleasi greseli. Atragerea utilizatorilor în acest proces este utila dar, din pacate, mai dificila în cazul unor dezvoltari strategice (de tip oferta), fara un context de utilizare concret.

Din aceste motive, una din tendintele actuale în evaluare este reconsiderarea inspectiei (expertizei) de utilizabilitate, care marcheaza o specializare profesionala la nivelul unei firme producatoare. Expertizele au avantajul de a estima, încă din faza de proiectare, probleme care pot sa apara în timpul utilizarii. O trecere în revista a acestor metode de evaluare permite identificarea a trei categorii mai largi: introspectii cognitive, inspectii specializate si evaluari euristice.

5. Evaluarea euristica

Evaluarea euristica a fost definita de Molich si Nielsen (1990) ca metoda de inginerie a utilizabilitatii care are ca scop identificarea problemelor de utilizabilitate în proiectarea unei interfete utilizator. Metoda se bazeaza pe utilizarea unui set restrâns de evaluatori care examineaza interfata si o evalueaza cu ajutorul unor principii recunoscute de utili-

zabilitate. În general, un singur evaluator nu poate identifica toate problemele potentiale de utilizabilitate pe care le poate avea o interfata. Mai mult chiar, este greu de gasit cel mai bun evaluator pentru a se baza numai pe aprecierile acestuia. Din acest motiv, este necesara implicarea unui numar mai mare de evaluatori, de regula 3-5.

Evaluarea euristica se desfasoara pe baza inspectiei individuale, facute de catre fiecare evaluator în parte. Numai dupa ce fiecare evaluare individuala a fost încheiata este permisa comunicarea între evaluatori si integrarea aprecierilor. Rezultatele evaluarii pot fi înregistrate fie într-un raport individual de evaluare, fie comunicate verbal, unui observator, pe masura inspectarii interfetei.

Pe parcursul evaluarii, sunt examinate pe rând elementele de dialog si evaluate în raport cu setul de principii de utilizabilitate, denumite euristici. Euristiciile sunt reguli generale care descriu proprietatile unei interfete utilizabile. În afara acestor principii generale, evaluatorul poate considera si alte reguli de detaliu pe care el le considera ca fiind relevante pentru dialogul respectiv. Mai mult chiar, se pot dezvolta euristici specifice, aplicabile unor clase de produse, care sa suplimenteze lista de principii.

Evaluarea se poate face pe un prototip sumar, caz în care este posibila în etapele de început ale ciclului de viata, sau pe un sistem complet, caz în care trebuie furnizate evaluatorului scenarii tipice de utilizare. Outputul evaluarii este o lista de probleme de utilizabilitate adnotate cu principiile de utilizabilitate care nu au fost respectate.

Un set initial de 9 euristici este dat de Molich si Nielsen (1990): dialog simplu si natural, fara elemente care nu sunt absolut necesare, exprimare în limbajul utilizatorului, minimizarea încarcarii memoriei utilizatorului, consistenta, feedback, scurt-circuitari (acceleratori), mesaje de eroare de calitate si prevenirea situatiilor de eroare.

Ulterior, Nielsen (1994a) a rafinat aceste euristici si a propus urmatoarele 10 principii de utilizabilitate:

- 1) vizibilitatea starii sistemului;
- 2) potrivire între sistem si lumea reala;
- 3) controlul si libertatea utilizatorului;
- 4) consistenta si standarde;
- 5) prevenirea erorilor;
- 6) recunoastere vs. reamintire;
- 7) flexibilitate si eficienta a utilizarii;
- 8) estetica si design minimal;
- 9) ajutarea utilizatorilor, recunoasterea, diagnosticarea si recuperarea în caz de eroare;
- 10) asistenta si documentatie.

Pentru fiecare euristica se specifica problema de conformitate (cum trebuie sa fie /ce trebuie sa faca interfata), elementele care probeaza conformitatea si motivatia evaluatorului (impactul criteriului asupra utilizabilitatii).

6. Insertia evaluarii în ciclul de viata

Unul din motivele pentru care metodele ingineriei utilizabilitatii nu sunt utilizate este costul ridicat. Nielsen (1994a) arata ca într-un proiect de dimensiuni medii (32000 linii sursa) se poate economisi de la 128000 \$ pâna la 65000 \$, prin utilizarea metodologiei "discount" care se bazeaza pe utilizarea a trei tehnici: scenariu; varianta simplificata de gândire cu voce tare; evaluare euristica.

În mod additional, principiul orientarii initiale pe utilizator poate fi sustinut cel puțin prin vizite la locul de munca al clientului.

Scenariile sunt un tip particular de prototipuri, care au atât trasaturi de prototipizare orizontala cât si verticala. Fiind prototipuri de dimensiuni mici, pot fi schimbate frecvent, mai ales daca sunt combinate cu studii bazate pe gândirea cu voce tare. În acest fel sunt utile pentru implicarea utilizatorului în proiect.

Teoretic, gândirea cu voce tare constituie un studiu realizat cu ajutorul unor psihologi sau experti în HCI, care utilizeaza o camera video pentru înregistrarea utilizatorilor. În cazul multor specialisti acesta poate constitui un cadru stresant, care intimideaza. Este totusi posibil sa se utilizeze o varianta simplificata a gândirii cu voce tare, fara a face apel la tehnica de laborator si experti,

invitând utilizatorul sa execute sarcina si sa spună ceea ce gândesc în acelasi timp.

7. Ingineria utilizabilitatii la nivel de firma

O tendinta actuala în evaluarea interfetelor om-calculator este integrarea mai puternica a metodelor de evaluare a proiectului în ciclul de viata al produselor informatice. În acest sens, se constata o largire a cadrului de lucru pentru evaluare, prin reconsiderare a ingineriei utilizabilitatii la nivel de firma.

O clasificare a companiilor dupa atentia acordata criteriilor de utilizabilitate, facuta de Nielsen (1994b), cuprinde 8 categorii:

1. utilizabilitatea nu are importanta;
 2. utilizabilitatea este importanta dar o interfata buna se poate asigura cu personalul existent;
 3. utilizarea unui expert HCI pentru inspectia de utilizabilitate a interfetelor om-calculator;
 4. utilizarea mai multor experti în anumite momente cheie, de schimbare a tipului de interfata;
 5. metode economice de ingineria utilizabilitatii aplicate sporadic;
 6. idem, aplicate sistematic;
 7. grup de experti sau laborator de utilizabilitate;
 8. acoperirea ciclului de viata al produsului.
- Raportat la clasificarea lui Erlich si Rohn (1994) nivelurile 1 si 2 corespund unei atitudini sceptice, 3 si 4 curiozitatii, 5 si 6 acceptarii si 7 si 8 parteneriatului. În prezent se constata o crestere a interesului factorilor de decizie pentru inserarea activitatii de evaluare la nivel de firma. Astfel, studii citate de catre Nielsen (1994b) arata o crestere a bugetului alocat utilizabilitatii la nivel de proiect pâna la un nivel de 4-6% în 1993.

În ceea ce priveste organizarea specialistilor în grupuri de utilizabilitate exista doua variante: un departament centralizat sau distribuirea în cadrul proiectelor. Pentru fiecare varianta exista avantaje si dezavantaje. Chiar daca exista un laborator de utiliza-

bilitate, ambele variante sunt aplicabile.

Principala ratiune de a fi pentru un laborator de utilizabilitate este conducerea experimentelor de testare a produselor de catre utilizatori. Odata înfiintat, poate fi utilizat si pentru activitati conexe: analiza sarcinii de lucru, înregistrarea sedintelor de proiectare, în special în cadru participativ (cu implicarea utilizatorilor) sau evaluarea euristica a utilizabilitatii.

8. Concluzii

Evolutiile din ultimii ani în domeniul evaluarii interfetelor om-calculator arata aparitia unor noi strategii si abordari, atât în ceea ce priveste managementul unui proiect, datorita insertiei unor tehnici si metode specifice, orientate pe procesul de interactiune, cât si la nivel de firma, datorita cerintelor de reconsiderare a prioritatilor privind alocarea resurselor umane si tehnice.

O trasatura caracteristica a acestor evolutii este utilizarea pe scara tot mai larga a inspectiilor de utilizabilitate datorita caracterului pragmatic, eficient si a valentelor de integrare la nivelul managementului de proiect si firma.

Bibliografie

Card, S.K., T.P.Moran si A.Newell, 1983, *The Psychology of Human-Computer Interaction*, Lawrence Erlbaum Associates, Hillsdale, New Jersey.
Coutaz, J. (1995) "Evaluating techniques: exploring the intersection of HCI and software engineering," In: Taylor,R and J.Coutaz (Eds.) *Software Engineering and HCI*, Springer.

Dix,A., J.Finlay, G.Abowd & R.Beale (1993) *Human-Computer Interaction*, Prentice Hall.

Ehrlich, K., and Rohn, J. (1994). Cost-justification of usability engineering: A vendor's perspective. In Bias, R.G., and Mayhew, D.J. (Eds.), *Cost-Justifying Usability*. Academic Press, Boston, MA.

ISO/DIS 9241-11:1994 Information Technology - *Ergonomic requirements for office work with visual display terminal (VDTs)-Guidance on usability*.

Johnson, P. (1992) *Human-Computer Interaction*, McGraw Hill, London.

Molich, R., and Nielsen, J. (1990). Improving a human-computer dialogue, *Communications of the ACM* 33, 3 (March), 338-348.

Nielsen, J. (1994a) Guerrilla HCI: Using Discount Usability Engineering to Penetrate the Intimidation Barrier. In Bias, R.G., and Mayhew, D.J. (Eds.), *Cost-Justifying Usability*. Academic Press, Boston, MA.

Nielsen, J. (1994b) "Usability Laboratories: A 1994 Survey" *Behaviour & Information Technology*, vol. 13, nos. 1-2.

Norman, D. A. (1998). *The invisible computer: Why Good Products Can Fail, the Personal Computer Is So Complex, and Information Appliances Are the Solution*. Cambridge, MA: MIT Press.

Payne,S.J. si T.R.G.Green (1986). "Task action grammars: a model of the mental representation of task languages". *Human-Computer Interaction*, 2, 93-133.

Whitefield, Wilson & Dewell (1991) "A framework for human factors evaluation". *Behaviour & Information Technology*, 1, 65-79.