



1860-2010

150



Universitatea din Iași
„Alexandru Ioan Cuza”

Laudatio

Profesor Nancy IDE

*Cu ocazia acordării titlului de Profesor de Onoare
al Universității „Alexandru Ioan Cuza” Iași*

LAUDATIO

Doamnei profesor **Nancy IDE**,
de la Vassar College, Statele Unite, cu ocazia acordării
titlului de *Profesor de Onoare* al Universității
„Alexandru Ioan Cuza” din Iași

25 martie 2010

www.uaic.ro

Laudatio

În onoarea

Doamnei prof. univ. dr. **Nancy IDE**, de la Vassar College, Statele Unite

Doamna prof. dr. Nancy Ide deține specializări în neuropsihologie, literatură engleză, lingvistică și informatică. În lucrarea sa de disertație, susținută în 1982, domnia sa utilizează metode statistice pentru determinarea modelelor de coeziune lexico-semantică în text, cu aproape un deceniu înainte ca abordarea statistică a analizei de text să fie îmbrățișată de comunitatea specialiștilor în lingvistică computațională.

În perioada 1985 – 1995 este președinte al *Association for Computers and the Humanities*, iar în 1996 preia conducerea redacției la *Association's journal, Computers and the Humanities*. Câțiva ani mai târziu, Nancy Ide restructurează revista pentru a releva schimbările în domeniu, redenumind-o *Language Resources and Evaluation*. Își continuă până în prezent activitatea, ocupând funcția de redactor șef adjunct al publicației.

În 1995 participă la înființarea seriei de cărți Springer - *Text, Speech and Language Technology*, fiind editor adjunct al acestei serii, care a publicat până acum peste 30 de volume, în care se prezintă cele mai recente metode de cercetare în domeniul lingvisticii computaționale.

În 1987 doamna profesor Ide a dat viață proiectului *Text Encoding Initiative* (TEI) prezidând periodic consiliul director, până în 1996. A lucrat de asemenea intens în cadrul a numeroase grupuri de activități TEI – printre cele mai remarcabile fiind redactarea părții relative la Dicționar, unde a propus standarde pentru reprezentarea informațiilor lexicale. Acestea au servit apoi drept bază pentru standardele extensive pentru date lexicale și terminologice dezvoltate ulterior de *International Standards Organization* (ISO). TEI este în continuare ținut la zi și actualizat, rămânând încă, și după 22 de ani de la inițierea proiectului, standardul de facto pentru codificarea textelor din științele umaniste. El a contribuit în mod considerabil la dezvoltarea limbajului de marcare extensibil (*Extensible Markup Language, XML*).

Doamna profesor Ide are publicații însemnate în domeniul dezambiguizării sensurilor și al semanticii lexicale. La începutul anilor 1990 a lucrat în domeniul extragerii relațiilor semantice din dicționare electronice, dezvoltând un model de rețea neuronală pentru utilizarea în WSD. Doamna sa s-a aflat de asemenea printre primii care au explorat și evaluat utilizarea informațiilor multilingvistice pentru WSD. A fost coautoare a unor lucrări care au deschis drumuri noi în domeniu, printre care amintim introducerea la un număr special dedicat WSD în prestigioasa revistă *Computational Linguistics* din anul 1998 – care rămâne până în prezent una din cele mai citate sinteze din domeniu. A redactat, de asemenea, în colaborare cu Yorick Wilks, un capitol mai recent în care analizează cercetarea actuală în WSD și tendințele ei viitoare. De curând a dezvoltat o metodă și o metrică de determinare a similitudinii semantice pentru a fi utilizată la

soluționarea listelor de cuvinte înrudite din punct de vedere semantic și totodată la determinarea similitudinilor dintre sensuri.

Doamna profesor Ide a desfășurat timp de mai mulți ani activități de cercetare în Europa, iar din 1993 până în 1997 s-a implicat în coordonarea proiectului MULTEXT, care era la vremea respectivă cel mai amplu proiect din UE inițiat vreodată în domeniul lingvisticii computaționale. Proiectul MULTEXT își propunea să dezvolte pentru prima dată, pe scară largă, resurse lexicale, inclusiv corpusuri, și instrumente de prelucrare pentru mai multe limbi, utilizând un format unitar. Proiectul MULTEXT a propus numeroase soluții folosite și azi în elaborarea și proiectarea lexicoanelor și a corpusurilor, precum și arhitecturi de prelucrare a limbajului, printre care arhitectura *pipeline* pentru instrumente NLP și conceptul de adnotare separată de text, pentru suprapunerea nedistructivă de informație lingvistică unui text. În cadrul proiectelor MULTEXT și EAGLES doamna Ide s-a ocupat de dezvoltarea Standardului de Codificare a Corpusurilor (CES, devenit mai târziu XML-CES sau XCES), care este încă larg răspândit și utilizat pentru reprezentarea corpusurilor lingvistice și adnotarea acestora. În 1994 profesor Ide a implementat și coordonat proiectul european de tip *follow-on* MULTEXT-EAST, care a aplicat metodele și instrumentele dezvoltate în cadrul MULTEXT la șase limbi din Europa de Est. Faimosul corpus „Orwell 1984”, care include în prezent versiuni paralele adnotate, în peste 20 de limbi, a fost unul din cele mai importante produse ale proiectului respectiv. Este folosit în continuare pe scară largă, fiind unul dintre puținele corpusuri paralele de dimensiuni mari. Ca exemplu, ediția românească a corpusului include marcaje XML pentru sintaxă și morfologie, segmentarea frazei, grupuri nominale,

coreferință și, parțial, pentru structura de discurs și teoria nervurilor (la care, de altfel, doamna Ide este coautoare).

Prof. Ide lucrează la elaborarea de standarde în reprezentarea și managementul resurselor lingvistice de mai bine de 20 de ani, după cum aminteam, primele realizări fiind TEI și standardul de codificare a corpusurilor. La mijlocul anilor 1990 a lucrat la mai multe standarde în cadrul proiectului european EAGLES. În 2000 s-a alăturat echipei de experți din cadrul Organizației Internaționale a Standardelor (ISO) TC 37 SC4 Managementul Resurselor Lingvistice coordonând grupul de lucru pentru descriptori și mecanisme de bază pentru resurse lingvistice. În acest context a creat recent Cadrul Lingvistic de Anotare (*Linguistic Annotation Framework - LAF*) și Formatul de Anotare Grafică aferent (*Graph Annotation Format - GrAF*), un formalism puternic de reprezentare a datelor lingvistice și a informațiilor lingvistice asociate. În prezent, doamna Ide este Coinvestigator Principal la proiectul „*Sustainable Interoperability for Language Technology*” (SILT), un proiect amplu finanțat de Fundația Națională de Științe a Statelor Unite, care și-a propus perfecționarea interoperabilității resurselor lingvistice, un proiect „înfrățit” cu proiectul cu finanțare europeană FLaReNet. În acest context, doamna Ide lucrează la elaborarea unei infrastructuri software ce va facilita interoperabilitatea resurselor și este preocupată de oferirea accesului liber al cercetătorilor la date lingvistice de diferite tipuri.

În 1998, doamna profesor Ide a inițiat proiectul de creare a Corpusului Național American (*American National Corpus*), comparabil prin dimensiuni și structură cu *British National Corpus*, care este folosit intens în cercetările de lingvistică computațională. Acest proiect este finanțat de mai mulți ani de NSF și are ca țintă dezvoltarea și

îmbunătățirea metodelor de adnotare prin învățarea automată, dezvoltarea reprezentării și a mecanismelor de acces pentru date lingvistice, incluzând adnotări la scară largă. Proiectul implică echipele FrameNet și WordNet, cu preocupări de armonizare a acestor două resurse și îmbunătățirea concordanței dintre adnotări în privința sensurilor.

În perioada 2002-2004, doamna profesor Ide a lucrat la un alt proiect finanțat de NSF în scopul exploatării tehnologiilor de web semantic pentru studiul documentelor istorice. Proiectul a clasificat anterior texte din Biblioteca Franklin D. Roosevelt privitoare la relațiile americano-japoneze din perioada care preceda bombardarea Pearl Harbor în 1941. Textele au mai multe niveluri de adnotare și au fost aliniate la o ontologie de entități și evenimente, ceea ce a făcut posibil, pentru prima dată, un tip de acces de o natură complet nouă și a facilitat descoperirea unor relații și caracteristici nerecunoscute până atunci. Doamna Ide a dezvoltat în cadrul acestui proiect un model de „ontologie de istorie” care își sistematizează reprezentarea de relații între elemente ontologice în funcție de cadrul temporal la care se raportează cererea adresată de utilizator.

În 2007, doamna Nancy Ide a devenit președintele Grupului Special de Interes pentru Adnotare din cadrul Asociației de Lingvistică Computațională (SIGANN). SIGANN susține financiar Workshop-ul pentru Adnotare Lingvistică, pe care domnia sa îl coordonează de patru ani.

Doamna profesor Ide a organizat și coordonat numeroase ateliere de lucru de-a lungul ultimului deceniu pe o varietate largă de teme, de la arhitecturi pentru prelucrarea limbii, cadre și metode pentru adnotare lingvistică, până la dezambiguizarea sensurilor cuvintelor și

diferențierea sensurilor din perspectivă psiholingvistică. Domnia sa a înființat de asemenea conferința asupra Interoperabilității Globale pentru Resurse Lingvistice, organizată o dată la doi ani la Hong Kong.

Doamna prof. dr. Ide este membru al Facultății de Informatică din cadrul Vassar College din 1982, unde predă cursuri de Teoria Calculului, Teoria Limbajelor și Automatelor, Proiectarea Compilatoarelor și Lingvistică Computațională. În ultima parte a anilor 1980 a contribuit la dezvoltarea curriculei de informatică pentru asemenea instituții, aceasta servind de atunci ca bază pentru curricula de informatică în universitățile de arte liberale din Statele Unite. A preluat coordonarea programei de informatică ce oferea la nivelul anului 1989 doar specializarea Matematică-Informatică. Conduc de domnia sa, Departamentul de Informatică a fost înființat la Vassar în 1991, oferind pentru prima dată specializarea independentă Informatică. A fost șefa departamentului timp de 20 de ani până în 2009, perioadă în care a îndrumat dezvoltarea programei spre a deveni una dintre cele mai bune programe de informatică într-o universitate de arte liberale din Statele Unite.

Anul 1997... Suntem la Tușnad și așteptăm ca profesorii invitați să fie aduși cu microbuzul universității de la Otopeni la Eurolan – școala de vară dedicată Prelucrării Limbajului Natural. Este a treia ediție într-o serie de școli internaționale bienale organizate de Facultatea de Informatică a universității noastre începând cu anul 1993 și continuate până în prezent. Pentru prima dată organizatorii au decis să mute Școala de Vară Eurolan în afara Iașului, într-un loc care să-i oblige pe profesori să interacționeze intens cu studenții, datorită programului dar și datorită configurației locului.

În afară de puținii lectori români (Dan Tufiș, Liviu Ciortuz și Dan Cristea) printre profesori se mai numără: Paula Monachesi – de la Utrecht, Massimo Poesio – de la Essex, Nicoletta Canzolari – din Pisa, John Sinclair – de la *Tuscan Word Centre*, Montecatini, directorul seriei de dicționare Collins, inventator al corpusurilor lingvistice, profesorul Aravind Joshi din Philadelphia – inventator al gramaticilor de adjuncții pe arbori și al teoriei *Centering* și... doamna Nancy Ide – deja celebra inventatoare a TEI – din Vassar, Poughkeepsie.

Ca de obicei, școala s-a desfășurat intens pe o durată de două săptămâni, timp în care au fost predate cursuri și susținute sesiuni practice pentru doctoranzi și studenți talentați din România și din alte șase țări, în fața unor calculatoare (pentru prima oară conectate la Internet în mica stațiune din Carpați, prin eforturi gigantice din partea organizatorilor). Doamna Nancy Ide este impresionată de atmosfera degajată și are timp să discute cu studenții și chiar să inițieze proiecte de colaborare cu organizatorii. Este atrasă așa de mult de ideea acestei serii de școli, încât se numără de atunci printre membrii consiliului director, contribuind la consolidarea importanței acestui eveniment în Europa și în întreaga lume. De atunci, din 1997, doamna Nancy Ide a mai venit de șase ori în România (de patru ori la Iași) pentru a se implica în organizarea acestor școli și pentru a împărtăși studenților și nouă experiența ei didactică și de cercetare.

Prof. dr. Nancy Ide are contribuții atât de însemnate în domeniile înrudite ale lingvisticii computaționale și ale prelucrării limbajului natural, încât numele domniei sale va rămâne unul dintre cele mai remarcabile. A păstrat până în prezent relații strânse de colaborare științifică cu Facultatea de Informatică a Universității „Alexandru Ioan Cuza” din Iași. În luna iunie vom avea a opta generație de studenți

care finalizează masteratul de Lingvistică Computațională la Facultatea de Informatică. Mulți dintre ei au rămas fideli acestui domeniu, continuându-și studiile la nivel de doctorat la Facultatea noastră sau în străinătate (am numărat 17 absolvenți cu titlu de doctor și 12 doctoranzi în momentul actual, dar sunt convins că numărul lor este mai mare). Răspândiți în toată lumea, acești tineri desfășoară în domeniul nostru activități științifice și de cercetare deosebite alături de mulți alți români nativi. Dacă îi întrebăm pe oricare dintre ei cine este Nancy Ide, fiecare va spune o istorie diferită. Au cunoscut-o de când erau studenți ai facultății noastre, dar au întâlnit-o din nou la conferințe mari, la întâlniri științifice importante, în amfiteatre, în cadrul unor proiecte internaționale, sau pe web, citindu-i lucrările. Într-un fel sau altul, doamna Nancy Ide este prezentă în mod continuu în viața științifică a celor care lucrează în acest domeniu.

Acordarea distincției de *Profesor de Onoare* al Universității „Alexandru Ioan Cuza” reprezintă privilegiul de a recunoaște atât meritele științifice de excepție ale doamnei prof. dr. Nancy Ide, cât și îndelunga colaborare cu universitatea noastră și cu alte instituții de cercetare din România.

Membrii comisiei îi urează o viață lungă și activă!

COMISIA DE ELABORARE A MEMORIULUI

Președinte,

Prof. univ. dr. **Vasile IȘAN**,

Rectorul Universității „Alexandru Ioan Cuza” din Iași;

Membri:

Prof. univ. dr. **Dan CRISTEA**,

Prodecan al Facultății de Informatică,

Universitatea Alexandru Ioan Cuza din Iași ;

Prof. univ. dr. **Horia Nicolai TEODORESCU**,

Prorector al Universității „Gheorghe Asachi” din Iași,

Director al Institutului de Informatică Teoretică al Academiei Române –
Filiala Iași,

Membru corespondent al Academiei Române

Prof. univ. dr. **Eugen MUNTEANU**,

Facultatea de Litere, Universitatea Alexandru Ioan Cuza din Iași,

Director al Institutului de Filologie Română „Alexandru Philippide” al
Academiei Române,

Director al Centrului de Studii Biblico-Filologice „Monumenta Linguae
Dacoromanorum” din cadrul Universității „Alexandru Ioan Cuza” Iași

Iași, 25 martie 2010

