

Till NUTEK, ITYP-programmet, att: Ulf Eklund

Forskningsrådsnämnden, programmet för Biblioteks- och informationsvetenskaplig forskning, att: Margareta Larsson

Skolverket, verksstaben, att: Lena Landgren

Utveckling av verktyg för hantering av strukturerade dokument

Projektet Datorstöd för kunskapsarbete, slutredogörelse

3 september 1993

Donald Broady och Hasse Haitto
IPLab, NADA, KTH
100 44 Stockholm

Sammanfattning

Datorstöd för kunskapsarbete är ett treårigt forskningsprojekt som avslutades i juni 1993. Det har finansierats av NUTEK, Forskningsrådsnämnden och Skolverket, och bedrivits som ett samarbete mellan KTH och HLS.

Syftet har varit att pröva och utveckla redskap som fungerar som stöd för hantering av dokumentbaser. De tänkta användargrupperna är lärare, forskare, dokumentalister, handläggare, redaktörer, läromedelsförfattare och andra som kan tänkas behöva bygga upp och handskas med omfångsrika och komplexa dokumentsamlingar. Ett centralt användningsområde är framställning av skräddarsytt undervisningsmaterial.

Det mest påtagliga resultatet av projektarbetet är dokumenthanteringssystemet SGML Darc, avsett för material som följer den internationella standarden SGML (Standard Generalized Markup Language, ISO 8879).

Innehåll

| | |
|--|----|
| 1. Institutionsplacering, finansiering, personal..... | 3 |
| 2. Huvudlinjer i projektarbetet | 3 |
| 3. Om de SGML-baserade verktyg som utvecklats inom projektet | 5 |
| 4. Samarbetspartners och kontakter, Sverige | 13 |
| 5. Samarbetspartners och kontakter, internationella..... | 15 |
| 6. Av projektet arrangerade sammankomster (urval) | 16 |
| 7. Deltagande i övriga konferenser och seminarier i Norden | 18 |
| 8. Deltagande i internationella konferenser | 18 |
| 9. Rapporter, uppsatser | 19 |
| Bilagor | 20 |

1. Institutionsplacering, finansiering, personal

Projektet "Datorstöd för kunskapsarbete" har varit ett samarbete mellan å ena sidan IPLab, Institutionen för numerisk analys och datalogi, KTH, å andra sidan Institutionen för pedagogik, Högskolan för lärarutbildning i Stockholm. Det startade den 1 juli 1990 och avslutades den 30 juni 1993.

Projektet har under tre år (bå 1990/91, 1991/92 och 1992/93) finansierats av NUTEK (ITYP projekt 90-02737P, den officiella benämningen på projektet är "Datorstöd för författande och samarbete i utbildning"). Under två år (bå 1991/92 och 1992/93) har projektet även erhållit stöd från Forskningsrådsnämnden (Dnr 910176, A 16-5/137) och från Skolverket (delprojektet "Kunskapsverkstaden"). Medlen från NUTEK har förvaltats av KTH och medlen från Skolverket och FRN har förvaltats av HLS.

Under projektarbetets inledningsfas erhöles smärre donationer av hårdvara från Compaq Computer AB och Apple Computer AB. Från SICS (Swedish Institute of Computer Science) har projektet erhållit en donation av tre grafiska arbetsstationer.

Följande personer har varit anställda inom projektet:

- Donald Broady, docent, tf professor vid HLS, forskare vid NADA, KTH (projektledare).
- Hasse Haitto, civ.ing., doktorand vid NADA, KTH (ansv. f. teknisk koordinering)
- Peter Lidbaum, civ.ing. (programmerare).
- Magnus Tobiasson, civ.ing. (programmerare).
- Eva Trotzig, fil lic, chefsbibliotekarie vid Statens Pedagogisk-Psykologiska Bibliotek, doktorand vid HLS (har inom projektet gjort begränsade insatser som försöksledare).
- Annika Ullman, fil lic, doktorand vid HLS (har inom projektet gjort begränsade insatser som försöksledare).

2. Huvudlinjer i projektarbetet

Den följande framställningen är översiktlig och behandlar framför allt systemutvecklingsaspekter.

2.1. 1990/91: preludeer

Projektarbetet föregicks av en förstudie, som var beställd av STU (nuv. NUTEK) och genomfördes av D. Broady under våren 1990. Denna förstudie ledde till att STU anslog medel till projektet "Datorstöd för författande och samarbete i utbildning" (även benämnt "Datorstöd för kunskapsarbete"), som startade 1 juli 1990 som ett av de första projekten inom STUs insatsprogram ITYP ("Informationsteknik inom Tjänsteproduktion för att utveckla Yrkesskicklighet och Produktivitet").

Projektet syftade till att utvärdera och utveckla datorstöd för hantering av dokumentbaser. De användargrupper vi hade i åtanke var en lärargrupp, en grupp läromedelsförfattare, en forskargrupp, en tidskriftsredaktion, en grupp handläggare etc. Ett centralt användningsområde var framställning av skraddarsydda läromedel. Medarbetarna måste kunna använda olika datorer och tillämpningsprogram, och de skall kunna ta emot, redigera och exportera strukturerade dokument. Ett välfungerande samarbete kräver nämligen att medarbetarnas personliga preferenser i val av arbetsverktyg respekteras, vilket leder till stor variation i fråga om tillämpningsprogram och dokumentformat. Därför var det väsentligt att under ett tidigt skede av projektarbetet beakta en rad problem rörande bryggorna mellan skilda dokumentrepresentationer.

Arbetet koncentrerades till tre plattformar: MS Windows (och MS DOS) på PC, unix på Sun Sparcstation, samt Macintosh OS på Apples datorer. Genom valet av dessa plattformar täckte vi de i utbildningsväsendet och arbetslivet vanligaste persondator- och arbetsstationsmiljöerna. Som väntat fick en hel del arbete ägnats åt tekniska problem (utvärdering av maskin- och programvara, integrering av program etc) som måste lösas för att fungerande installationer i enlighet med projektets intentioner skulle komma till stånd.

Inledningsvis ägnades en hel del arbete åt att granska existerande programvara, prestanda hos dessa, samt jämförelser mellan produkterna. Detta är ett nytt och expansivt område; det är först på allra sista tiden som en rik flora av produkter för arbete med ett generaliserat dokumentbeskrivningsspråk blivit tillgänglig. Ett antal program prövades mer ingående, i första hand för översättning till och från SGML, men även program/hårdvara för bl. a. optisk läsning och för versionshantering.

Ett kärnproblem med alla mer omfattande komplexa dokumentbaser gäller svårigheterna att orientera sig, att skaffa sig överblick, att inte gå vilse. Här erbjuder hypertextsystemen vissa möjligheter. De hypertextsystem som mest intensivt prövats inom projektet har varit *NoteCards* från Xerox (på Xerox Lisp-maskiner), *Intermedia* från Brown University, Providence, Rhode Island, USA (på vissa Macintoshmaskiner under Apple unix version 1.1.1), *DynaText* från Electronic Book Technologies, Providence, Rhode Island, USA (på Sun Sparcstation och på PC under Windows), *Guide* från University of Kent, England, samt OWL Ltd, Edinburgh, UK (versionen från University of Kent körs på SUN-maskiner, den kommersiella versionen från OWL Ltd på PC under MS Windows och på Apple Macintosh), samt *KMS* (Knowledge Management System). Till dessa hypertextsystem arbetade vi bl.a. med importrutiner av sekventiella texter kodade i RTF-format (ett väl definierat format, som understöds av flera vanliga textbehandlare).

Vart och ett av de nämnda hypertextsystemen har sina begränsningar och fördelar. *DynaText* tillåter presentation på skärmen av SGML-märkade dokument (inkl programmerbara hypertextlänkar); nackdelen är att *Dynatext*—förutom att vara förhållandevis dyrt—framför allt är avsedd för presentation av statiskt material, samt att dokumenten initialt måste göras om till "elektroniska böcker" i ett förberedande steg. Däremot tillåter *Dynatext* t.ex. annotering av dokumenten vid läsning på skärm. *Guide* är lämpligt för mer interaktiv användning. Både *Dynatext* och *Guide* kan f.n. användas såväl under unix som på DOS-maskiner under Windows. *Intermedia* är av flera skäl det intressantaste systemet för undervisningsbruk, inte minst för att lärarna och eleverna kan skapa egna sätt att organisera en och samma dokumentbas med hjälp av personliga vävar av hypertextlänkar. Tyvärr har *Intermedia* numer enbart historiskt intresse, eftersom utvecklingen kom i otakt med hårdvarans och operativsystemets utveckling och pengar saknades för att driva *Intermedia*-projektet vidare på andra plattformar. *KMS* är sidorienterat och därmed inte särskilt lämpat för arbete med omfattande och komplexa dokument (många av de enklare kommersiella hypertextsystemen lider av samma svaghet, de inbjuder till att informationen styckas sönder för att passa in på små "sidor" eller "kort").

Inom parentes kan tilläggas, att de europeiska installationerna av de stora "klassiska" hypertextsystemen *Intermedia* och *NoteCards* varit lätt lätt räknade. I Sverige är möjligen i dag installationerna som finns kvar från projektet Datorstöd för kunskapsarbete de enda som existerar (periodvis har det funnits installationer i Linköping och någon enstaka *NoteCards*-installation i Göteborg). Det har sitt värde att dessa system, som många talar om men få sett i levande livet och som bara kan köras på maskiner som tagits ur produktion och under föråldrade versioner av operativsystemen, finns till påseende i vårt land — här skulle man önska sig ett slags K-märkning av klassiska datormiljöer.

Det inledande projektåret (1990-1991) användes således framför allt till att skaffa kunskap om och pröva tillgängliga systemlösningar. Projektets utgångspunkter sammanfattades hösten 1991 i den programmatiska uppsatsen *Kunskapsverkstaden — om lokala dokumentbaser som arbetsverktyg för lärare*. Resultaten av en inledande granskning av produkter av relevans för SGML-relaterade tillämpningar redovisades i rapporten *SGML-relaterade verktyg för intermediära former och dokumenthantering*.

2.2. 1991/92: verktygen tar form

Från och med budgetåret 1991/92 har Skolverket och Forskningsrådsnämnden bidragit med stöd till projektet.

Tack vare stödet från Skolverket kunde arbetet bättre inriktas mot tillämpningar inom utbildningsväsendet.

Tack vare stödet från FRN (programmet för Biblioteks- och informationsvetenskaplig forskning) kunde vi ägna uppmärksamhet bl.a. åt modeller för arbete i anslutning till bibliotekssystem. Vi tänkte oss inledningsvis att bygga moduler för hantering av fulltextdatabaser som kunde integreras med det unix-baserade bibliotekssystemet TinLib. Vi övervägde att som databashanterare använda TinMan, som ingår i detta bibliotekssystem, och vi erhöll efter vissa besvärligheter från IME (som utvecklar bibliotekssystemet TinLib) löften om långtgående samarbete (och fri tillgång till relevant programvara och teknisk dokumentation). Efter utprovningar och funderande kom vi fram till en mer generell lösning; det framstod efter hand som rimligare att inte knyta systemutvecklingen till ett visst bestämt bibliotekssystem. Redan de astronomiska kostnaderna för slutanvändare gjorde det orimligt att bygga verktyg som förutsatte licensiering av TinMan.

Efter den inledande kunskapsinsamlingen, anskaffning av nödvändiga program och försök i mindre skala, inleddes 1992 arbetet med att på allvar börja bygga verktyg för dokumenthantering enligt riktlinjerna i den nämnda uppsatsen. Därför anställdes fr.o.m. januari 1992 två programmerare inom projektet, Peter Lidbaum och Magnus Tobiasson, båda nyutbildade civilingenjörer med datalogisk inriktning.

Det arbete som skulle komma att realiseras avvek i flera avseenden från det som ursprungligen planerades. Från början var ambitionerna blygsamma. Det tekniska utvecklingsarbetet var inriktat på att skapa bryggor (huvudsakligen i form av konverteringsrutiner) mellan olika färdiga program och att utpröva och dokumentera arbetsmodeller. Vi tänkte oss närmast att skapa en prototypomgivning med tillgänglig programvara som byggstenar: TinMan som databashanterare, kanske Intermedia som hypertextmiljö, och SGML-kodning och viss översättningsprogramvara därtill som sammanhållande kitt. Genom en gemensam intermediär form som bevarar strukturen hos materialet kan man mimimera mängden av översättningsrutiner som behöver skapas; för varje nytt format som tillförs miljön definieras då endast två rutiner för översättning till och från det egna intermediära formatet. Vi valde tidigt att arbeta med SGML (Standard Generalized Markup Language, ISO 8879) som intermediärt format, eftersom denna standard lämpar sig mycket väl för beskrivning av strukturerade dokument.

Denna ambition framstod efter hand som otillräcklig och av alltför begränsat praktiskt värde för de tillämpningar i arbetslivet vi siktade mot. Det kändes föga meningsfullt att bygga en prototyp enbart för laboriemiljön, och som var beroende av programvara som förmodligen har begränsad livslängd och som i vissa fall (TinMan, DynaText) var så orimligt dyr att den aldrig kunde tänkas komma till bredare användning. Över huvud taget är det olyckligt att alltför hårt binda lösningarna till de tillämpningsprogram som för tillfället råkar finnas på marknaden. I stället är mer generella lösningar är att föredra. Vi ville med andra ord följa principen om öppna system och leverantörsberoende.

Inter medias hädanfärd illustrerar riskerna med att vara alltför leverantörsberoende. Intermedia från Brown University var då projektarbetet inleddes det i många avseenden bästa hypertextsystemet för undervisningsbruk. Dock var Intermedia vid denna tidpunkt bundet till en enda plattform, Apples Macintoshdatorer, och till Apples egen tidiga unix-version. Denna bindning gjorde att Intermedia gick i graven då samarbetet med Apple började gnissla och då utvecklingen av Intermedia därmed kom i otakt med utvecklingen av Apples hårvara och Apples eget unix. Än mer beklagligt var att de dokumentbaser som skapats inom Intermedia inte kunde användas av andra system. Inter medias öde gav oss en tankeställare. Att knyta projektarbetet till alltför specifik hårdvara och mjukvara innebär att hela kedjan brister när någon av dess länkar kommer i otakt med utvecklingen.

Den lösning vi inledningsvis eftersträvade — närmast att skapa kitt mellan befintliga kommersiellt tillgängliga tillämpningar — visade sig således vara mindre välbetänkt. Vi hade länge trott att Intermedia skulle kunna användas inom det svenska utbildningsväsendet. I stället satsade vi nu på

generella lösningar genom ett massivt utnyttjande av de möjligheter som den internationella dokumentbeskrivningsstandarden SGML erbjuder. Dokumentbaser som ansluter till SGML-standarden kommer mänskligt att döma att kunna användas och återanvändas för olika syften under överskådlig framtid. Även om de verktyg och system (däribland de som utvecklats inom vårt projekt) som använts för att skapa och hantera dessa baser skulle bli föråldrade, finns ändå dokumentbaserna själva kvar, omedelbart färdiga att bearbetas med nya verktyg och system. Detta är den kanske viktigaste fördelen med SGML-märkningen: att informationen erhåller en sådan form att den överlever de tillämpningsprogram och datormiljöer i vilka de skapats. Register, innehållsförteckningar, hypertextlänkar etc kan vid behov extraheras fram ur dokumenten själva, i stället för att bindas till särskilda datormiljöer och tillämpningsprogram.

Vi ser det över huvud taget som en viktig uppgift att främja utvecklingen av SGML-baserade tillämpningar. SGML-märkningen är i dag den internationellt sett helt dominerande standarden för beskrivande dokumentmärkning. Ju förr svenska dokumentproducenter (förlag, tryckerier, företag, myndigheter, organisationer) övergår till ett standardiserat dokumentbeskrivningsspråk dess bättre. För övrigt är många ledande svenska företag, olika enheter inom det svenska försvaret och vissa statliga myndigheter just nu i färd med att utreda metoder för att SGML-märka sin dokumentation.

2.3. 1992/93: SGML Darc

Det mest påtagliga resultatet av projektets arbete är dokumenthanteringssystemet SGML Darc (Document ARchive Controller), avsett som ett stöd för uppbyggnad och användning av företrädesvis SGML-märkta dokumentbaser. En preliminär version av SGML Darc visades upp för en internationell publik i samband med konferensen ECHT '92 (fourth ACM-conference on hypertext) i Milano, 30 nov—4 dec 1992. Den första skriftliga presentationerna var rapporten *Darc — Document Archive Controller*, som publicerades i mars 1993, gjordes tillgänglig via de internationella näten och presenterades i relevanta elektroniska konferenser.

Gensvaret blev överväldigande. Redan en vecka efter publiceringen hade vi fått in ett femtiotal seriösa förfrågningar från intresserade världen över, och sedan dess ytterligare många fler — vilket illustrerar att SGML-verktyg av denna sort är fåtaliga och synnerligen efterfrågade. Förfrågningar har framför allt kommit från dokumentationsansvariga vid industriföretag, förlag, universitet och högskolor samt från militära organisationer, statliga myndigheter etc. — i samtliga fall företag eller organisationer som har att hantera stora mängder strukturerade dokument. Vi har någorlunda kontinuerliga kontakter med det fåtal svenska företag där SGML redan sedan något år används i produktionen, och även kontakter med ett antal företag och organisationer som f.n. utreder möjligheter att lägga över sin dokumentation i SGML-format.

Som nämnts har arbetet inom projektet utvecklats i riktning mot allt mer generella lösningar. Det innebär inte att vi tar mindre hänsyn till de användarkategorier vi ursprungligen främst hade i åtanke, lärare, men vi tror att alla, även lärare, tjänar på att systemen är öppna och generella — det innebär nämligen att systemkonstruktörerna gör så få antaganden som möjligt om användarens behov.

Den viktigaste egenskapen hos SGML Darc är att systemet till fullo utnyttjar den information som SGML-kodningen bär med sig. Som ett exempel kan nämnas att bibliografiska data på basis av den i dokumentet befintliga SGML-kodningen extraheras helt automatiskt när ett dokument lagras — en funktion som välkomnas av bibliotekarier och dokumentalister. Dokumenten blir därför självlagrande och självindexerande. Till skillnad från t.ex. DynaText kan dokumenten omedelbart presenteras för läsning på skärm. Vidare finns hypertextfunktioner som t.o.m. är mer utvecklade än Intermedia (bl.a. tillåter SGML Darc att man har flera vävar öppna samtidigt). I SGML Darc finns annoterings- och överstrykningsmöjligheter, vilket är av värde inte minst för samarbetet inom en arbetsgrupp.

Samtidigt som vi utvecklat dessa funktioner har vi lagt oss vinn om att behålla den viktigaste fördelen med SGML-kodningen: de återanvändbara dokumenten. De dokument som SGML Darc hanterar är vanliga standardiserade SGML-dokument. De kan därför även hanteras av andra tillämpningsprogram, konverteras till vanliga ordbehandlingsformat etc.

För en kort presentation av de olika modulerna i SGML Darc, se nedan.

2.4. Fortsatt arbete

Vi har kontakt med ett antal svenska företag och organisationer som eventuellt vill använda systemet SGML Darc, antingen i dess helhet eller moduler eller anpassningar därav. Vi har redan en livlig korrespondens med företag och organisationer runt om i världen som önskar prova systemet. Den beta-version (för PC under Windows) som blev färdig i juli 1993 beta-testas när detta skrives (i början av sept 1993) vid Nice Technologies, Schweiz; USIT, Oslo universitetet; SoftQuad, Toronto, Canada; Oxford University; Telub; Falch Hurtigtryck AB, Norge; Cork University College, Storbritannien; CAPS (ett EG-projekt); Jordbruksdepartementet; SAAB Service partner; samt Skolverket. (Med beta-testning avses en fältmässig utprovning innan ett program är moget för reguljär användning.) Resultaten från denna beta-testning kommer att påverka utformningen av de definitiva versioner för både MS Windows och unix som beräknas bli färdiga under loppet av hösten 1993.

För att möjliggöra en fortsatt produktutveckling och anpassningen till behoven inom olika företag har ett avknopningsbolag från KTH bildats. Detta bolag, Synex Information AB, kommer att försälja SGML Darc. Förhandlingarna med företag som uttryckt intresse för att använda SGML Darc berör i huvudsak tre möjligheter 1) slutanvändning som dokumenthanteringssystem, 2) integration med annan mjukvara, samt 3) licensiering av delkomponenter, i första hand Darcs presentationsverktyg. En del av de eventuella intäkterna kommer enligt avtal att återbetalas till NUTEK. (Det omfattande programmeringsarbete som lett fram till produkten har möjliggjorts tack vare stödet från NUTEK, och NUTEK har aktivt uppmuntrat "produktifieringen" av SGML Darc och tillkomsten av avknopningsföretaget).

Samtidigt är vi angelägna om att SGML Darc inte enbart skall bli en angelägenhet för näringslivet. Vi vill även fortsättningsvis bidra till den informationsteknologiska utvecklingen inom vissa områden — forskningen, undervisningsväsendet, biblioteken — där det vore orimligt och oetiskt att låta kommersiella hänsyn råda, särskilt med tanke på att FRN och Skolverket bidragit till att finansiera projektarbetet.

Därför tänker vi oss att försäljningen av SGML Darc sker genom det nämnda avknopningsbolaget, samtidigt som vi även fortsättningsvis i våra forskningsmiljöer (IPLab, Nada, KTH samt Institutionen för pedagogik, Högskolan för lärarutbildning) kommer att arbeta på så sätt att forskare, vissa lärarkategorier, forskningsbibliotek etc får tillgång till redskapen till rimlig kostnad (eller gratis). Det finns exempel på att det slaget av arbetsdelning är fruktbart: hypertextsystemet *Guide* har under många år utvecklats i två parallella versioner, dels en version som tillhandahålles av University of Kent och är främst avsedd för akademiska miljöer, dels en kommersiell version som säljs av OWL Ltd.

3. Om de SGML-baserade verktyg som utvecklats inom projektet

Den följande framställningen är en smula teknisk. Vi kommer självfallet att skriva mer lättillgängliga presentationer. Tilläggas bör, att Darc kan användas även av den som inte vet något alls om SGML eller andra teknikaliteter. Däremot har vi sett det som väsentligt att Darc skall fungera flexibelt och väl även i händerna på den som är väl förtrogen med alla möjligheter som SGML erbjuder. Man kan t.ex. tänka sig grupp handläggare där den projektansvarige gör fullt bruk av systemets funktionalitet medan somliga medarbetare nöjer sig med att arbeta enligt de rutiner som den projektansvarige förberett.

3.1. Vad är SGML?

SGML Darc (fortsättningsvis för korthets skull benämnt Darc) är ett dokumenthanteringssystem som utnyttjar de nya möjligheter som skapats i och med SGML (Standard Generalized Markup Language, ISO 8879), den internationella standard för dokumentbeskrivning och dokumenthantering som antogs 1986.

Intresset för SGML har stadigt vuxit under de senaste åren, i Sverige och än mer i vissa andra länder, kanske främst USA och Japan och i Europa Frankrike, Tyskland och Holland. Att SGML-standarderna så snabbt accepterats beror i hög grad på att den mycket tidigt antogs och främjades av den amerikanska försvarsmakten, den amerikanska förläggarföreningen, Europakommissionen, IBM, Cern, och andra inflytelserika organisationer och företag. SGML har på sistone fått allt starkare stöd av nybildade inflytelserika organisationer som konsortiet SGML Open, där tiotalet stora internationella företag avser att uppmuntra spridningen och användningen av SGML. Även de tongivande datorproducenterna i Open Software Foundation har annonserat sitt stöd för SGML. Standarderna har också fått draghjälp genom spridningen av nya standarder såsom ISO 9000 (en standard för kvalitetssäkring, som förutsätter att de olika stegen i produktionsprocessen etc dokumenteras på ett standardiserat sätt) och ISO 10744 (HyTime, antagen 1992), den enda existerande internationella standarden för Multimedia och tidsberoende information (ljud, video). HyTime bygger på SGML, vilket även gäller för en kommande ISO-standard, DSSSL, som när den blir färdig utan tvivel kommer att få stor betydelse när SGML-dokument skall typograferas och tryckas på papper.

Hittills har arbete med SGML krävt specialistkompetens och programmeringskunskaper, men nu börjar allt fler mer lättanvända verktyg att erbjudas. Så kommer t.ex. de ledande textbehandlingsprogrammen att erbjuda SGML-stöd. Det innebär att användaren kan spara sina textbehandlingsfiler i SGML-format, och i viss mån även gå motsatt väg, dvs importera SGML-dokument till textbehandlingsprogrammet. WordPerfect saluför redan en sådan konverteringsprogramvara, Intellitag, för MS DOS och unix, och utlovar en version för MS Windows under första kvartalet 1994. MS Word lär följa efter (under första kvartalet 1994, enligt informella upplysningar från Microsoft); redan nu använder Microsoft SGML internt, t ex för att skapa Microsoft Bookshelf. Vidare finns det allt fler program (Darc är ett exempel) där kontaktytan med användaren är så utformad att denne inte behöver veta något alls om SGML.

SGML-märkning innebär att man förser dokumenten med information om deras innehållsliga uppbyggnad. På så sätt kan dokumentens byggstenar, t ex titel, författarnamn, rubriknivåer, litteraturreferenser, listor, fotnoter, bilder etc, identifieras och hanteras. På senare tid har SGML blivit den dominerande standarden för sådan s.k. beskrivande märkning av dokument.

SGML-standarderna är viktiga av flera skäl:

Flyttbarhet: dokument kan hanteras och distribueras oberoende av hårdvaruplattform, programvara och nationella teckenuppsättningar. För första gången existerar nu en internationellt accepterad standard för framställning och distribution av strukturerade dokument.

Återanvändning: ett och samma dokument kan användas för olika ändamål, t ex för läsning på skärm, som databas, som underlag för elektronisk överföring eller för publicering på papper. SGML-märkningen garanterar att dokumenten får längre livslängd än dagens maskinpark, operativsystem och tillämpningsprogram.

Väldefinierad regelstyrd märkning: dokumenten kan både läsas av människor och hanteras av många slag av program. Exempel på användningsområden är automatisk katalogisering och indexering av dokumentbaser, konvertering till textbehandlings- eller layout-program.

Några ord bör sägas om några centrala begrepp inom SGML-vokabulären: DTD, instans, element, entitet och attribut.

SGML bygger på grundidén att dokument kan indelas i klasser, var och en karaktäriserad av gemensamma egenskaper i fråga om innehållslig struktur. Affärsbrev kan utgöra en klass, tekniska manualer en annan, romaner en tredje, akademiska avhandlingar en fjärde, språklexika en femte, osv. Varje dokument som tillhör en viss klass säges vara en "instans" av en för hela klassen gemensam struktur, vilken i SGML beskrivs med en dokumenttypdefinition (DTD). I en DTD namnges de element — författarnamn, rubriker, fotnoter, listor, glosor, citat på grekiska eller vadhelst som är relevant i detta slag av dokument — som kan ingå i dokument som tillhör klassen, och där definieras på vilket sätt elementen kan förekomma. För varje klass av dokument måste en särskild DTD skapas. En DTD är således ett regelverk som entydigt fastställer ramarna för hur dokumentets struktur kan realiseras.

Elementen som definieras i en DTD är behållare för informationen i dokumenten. Det innebär att ett SGML-dokument består av två delar: information och struktur. Elementen kan kapsla in varandra (elementet för efternamn kan ingå i elementet för författarnamn, elementet för stycke kan ingå i elementet för kapitel), men de kan inte överlappa varandra. Därmed utgör dokumentets struktur en hierarki av element. Sådana hierarkier brukar ofta representeras grafiskt som ett träd.

Vad man vill kalla ett element är valfritt. I en viss DTD kan elementet för titel benämnas som `<title>`, i någon annan DTD — förutsatt att ett titelelement överhuvudtaget definierats i denna DTD — kan benämningen `<ti>` förekomma.

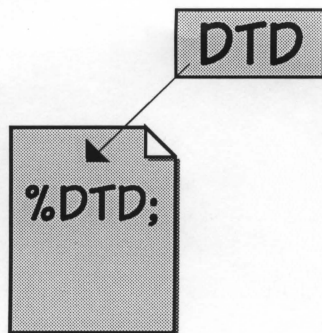
SGML är en neutral standard som är helt maskin- och mjukvaruoberoende. Detta är möjligt genom att standarden är definierad genom en rad abstraktioner. Så existerar exempelvis inget filbegrepp utan i stället en generell mekanism som kallas för entitet, där en entitet kan vara allt från ett komplett dokument till ett enstaka tecken.

Till varje element är det möjligt att definiera attribut. Som termen antyder är det frågan om egenskaper som hör ihop med elementet; attributen är med andra ord platsbyllare för information som hör till elementstrukturen. En vanlig användning av attribut är för att skapa korsreferenser, dvs samband mellan olika ställen i dokument.

3.2. Kort presentation av Darc

Darc är ett dokumenthanteringssystem avsett för dokument kodade i SGML (Standard Generalized Markup Language, ISO8879), en internationell standard för dokumentbeskrivning. Systemet kombinerar nutida informationsteknologi från olika områden: strukturerad dokumentation, SGML, databasteknologi, hypertext. Det är ett stöd för arbetet med att ordna, indexera, genomsöka (både vanlig fritextsökning och sammanhangsberoende sökning), återvinna och exportera (publicering elektroniskt eller på papper, presentation på skärm) material ur dokumentbaser, vilka kan var tämligen stora (flera tiotusental dokument). Darc finns f.n. i betaversioner för Sun Open Windows och PC/Microsoft Windows 3.1.

Varje SGML-dokument är alltid försett med en DTD som entydigt definierar dokumentstrukturen i termer av element. Vad som är element, och vad elementen kallas, skiljer sig från DTD till DTD. Det faktum att varje dokument har en DTD utnyttjas till fullo i Darc för att åstadkomma smidig dokumenthantering. Förutsatt att tillämplig DTD finns lagrad i systemet kan ordinära SGML-dokument kan utan någon som helt modifiering omedelbart användas av Darc. Darc accepterar alla DTD:er som bygger på den s.k. Standard Reference Syntax, vilken är en minsta gemensam nämnare för det stora flertalet SGML-tillämpningar.

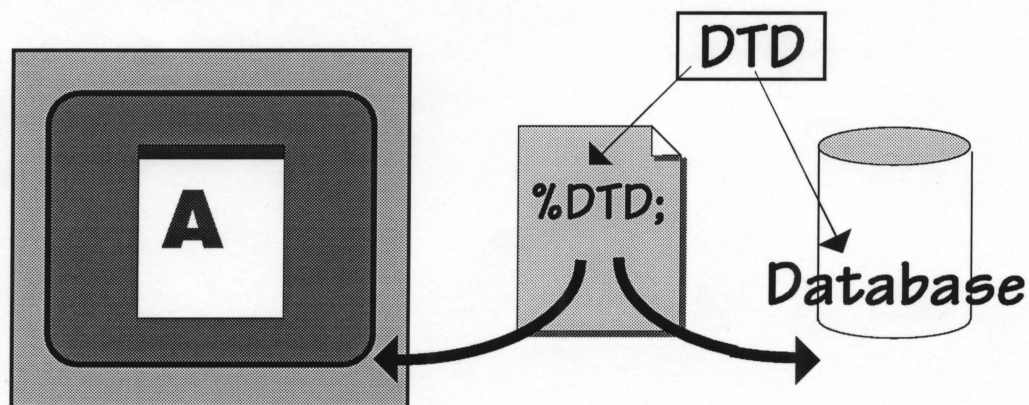


- Darc förutsätter att dokumentinstanserna inkluderar en DTD med hjälp av den mekanism som kallas PUBLIC identifier.

Darc består av tre huvudkomponenter:

1. en databashanterare för SGML-dokument,
2. ett presentationsverktyg som gör det möjligt att läsa, utforska, och förse dokument på skärm med egna anteckningar, samt
3. vyer, en mekanism för att ordna filerna i en hierarkisk trädstruktur.

Darc är i grund och botten en databashanterare, men specialiserad för SGML-dokument. Som i vilken databas som helst behövs det index — söknycklar — för åtkomst av dokumenten. Indexeringen i Darc är dock osedvanligt flexibel: Först registreras DTD:n för varje nytt dokumentslag. I denna process anger användaren vilka element som ska användas för indexering. Man kan alltså skapa index som författare, titel, ISBN, etc och göra kopplingen till motsvarande SGML-element för varje DTD. Detta medför ett konsistent gränssnitt mot användaren: varje dokument kan sökas genom samma indexbegrepp, trots att dokumenten har olika DTD:er. Den största fördelen uppstår sedan vid lagring av dokument: indexen byggs upp ur det SGML-kodade innehållet och dokumenten lagrar sig själva fullständigt automatiskt.



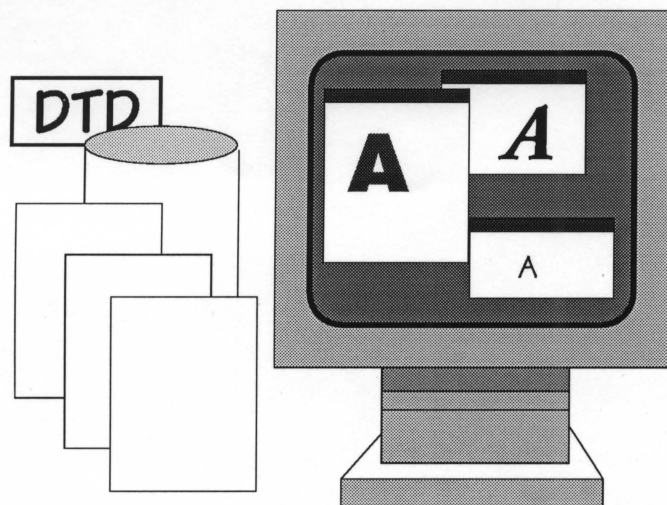
- Indexering och lagring av dokument definieras genom villkor relaterade till DTD:n.

Ytterligare en fördel är att ett dokument inte behöver vara lagrat i databasen för att kunna presenteras och läsas på skärmen.

När en användare definierar indexelementen i DTD:n fastställs även vilka element — titel, författare, etc — som ska presenteras som bibliografiska data. Om dokumentklassen även har element för abstract eller nyckelord är det naturligt att visa innehållet i dessa element här. Bibliografiska data används bl.a. för att ge användaren en snabb uppfattning om dokumentinnehållet.

Presentationen av dokumenten på skärmen kan varieras efter behov. Formateringen styrs genom användardefinierade stilmallar. Dessa definieras med avseende på elementen och knyts till motsvarande DTD, dvs alla dokument med samma DTD kan utnyttja en gemensam uppsättning

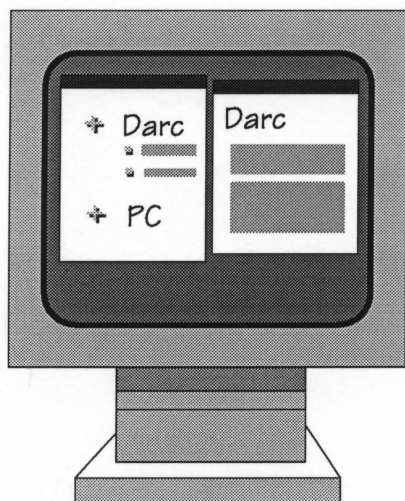
stilmallar. Det är även möjligt att visa ett och samma dokument med olika layout samtidigt på skärmen. Denna formatering sker dynamiskt och användare kan således medan de läser ändra utseendet på dokumentet. Redigering av stilmallar sker interaktivt (WYSIWYG) men de lagras SGML-kodade med FOSI-liknande syntax.



- Presentation av dokument sker med stöd av stilmallar knutna till DTD:n.

Formateringen av elementen är kontextberoende, det går alltså att styra formateringen genom att formulera villkor rörande elementen (motsvarande t ex "endast första stycket i en lista", "alla förekomster utom den första", etc). Presentationsverktyget har även stöd för att tolka korsreferenser kodade med ID och IDREF(S)-attribut. Sådana korsreferenser visas med hypertextlänkar som kan följas från båda hållen.

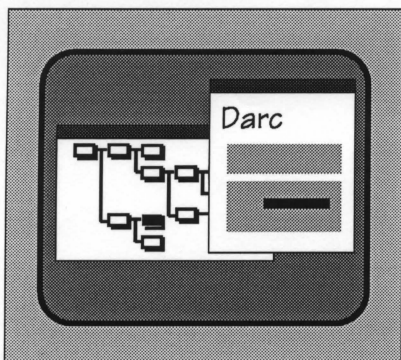
En särskild form av stilmallar är navigeringshjälpmedlen. Dessa är generaliserade innehållsförteckningar och definieras så att de endast visar särskilda element. Den vanligaste användningen torde vara att visa rubriker på olika nivåer, men man kan även tänka sig att visa alla bildtexter i ett dokument, alla tabellrubriker etc. När man klickar i en sådan "innehållsförteckning" bläddras aktuellt fulltextfönster fram till motsvarande avsnitt.



- Generaliserade innehållsförteckningar som navigeringshjälpmedel.

Det finns funktioner för utskrift på papper av hela eller delar av dokument, inklusive tabeller, bilder etc. Vid behov skapas korsreferenser som hänvisar till rätt sida i utskriften.

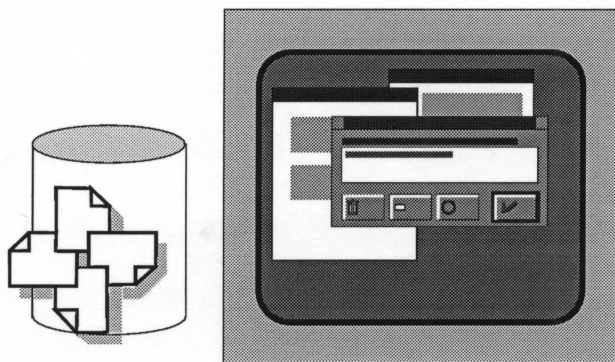
Förutom förflyttningar med hjälp av rullist, innehållsförteckningar och korsreferenser (dvs via hypertextlänkar), kan Darc visa en interaktiv grafisk bild av dokumentets innehållsliga struktur, dvs. SGML-trädet. Trädet används för navigering och för enkel åtkomst av stilmallsdefinitioner. (Här finns en potentiella pedagogiska vinster: genom att använda Darc blir den som så önskar allt mer förtrogen med både med SGML i allmänhet och de specifika DTD:er som tillhör de dokumenttyper han eller hon handskas med.)



- SGML-trädet visar strukturen hos dokumentet. Det används för navigering och för modifiering av stilmallarna.

Ytterligare ett sätt att förflytta sig i dokument är fritextsökning, där sökresultaten visas i en interaktiv lista; varje förekomst visas i sitt sammanhang och kan användas för att bläddra fram till motsvarande rad i fulltextfönstret. Dessutom kan fritextsökning kombineras med s.k. struktursökning, genom att användaren begränsar sökningen till bestämda SGML-element. Exempelvis: sök alla "Smith" i elementet författarnamn, eller i elementet litteraturreferens, eller i elementet abstract eller i...

En speciell funktion hos Darc är vävarna. Vävar är en sorts behållare för information som knyts till utsnitt i dokument. (Grundidén liksom termen, "webs", är lånad från Intermedia, men i Darc utvecklad till betydligt större flexibilitet.) Vävarna kan innehålla tre ting: "marginalanteckningar", "understrykningar" och "korshänvisningar". Man kan således förse dokument med fristående noteringar knutna till bestämda platser i dokumentet, man kan markera partier med avvikande färg, ungefär på samma sätt som med en överstrykningspenna, och man kan infoga hypertextlänkar som fungerar som vägvisare inom dokument eller mellan olika dokument. All sådan extern information skapas interaktivt och lagras i vävar; om motsvarande väv inte öppnas visas inte denna information, vävarna stör således inte dokumentens integritet. Flera vävar kan vara öppna samtidigt: presentationen på skärmen är då summan av alla anteckningar och länkar i vävarna. Vävinformationen kan vidare undersökas och sällas med ett eget verktyg.



- Vävar används för anteckningar, överstrykningar och interaktiv hypertextlänkning mellan utsnitt.

En nackdel med databaser är att de brukar utgöra ett slags svart låda, där åtkomsten bygger på indexsökningar och användaren saknar möjlighet att bilda sig en uppfattning om databasens samlade

innehåll. Som ett komplement till index (som allt efter användarens önskemål kan vara fulltextindex eller mer speciella DTD-beroende index) har Darc därför försetts med en mekanism kallad vyer, med vars hjälp användaren skapar hierarkiska strukturer av dokument. Dokument i databasen kan placeras i noder, vilka i sin tur kan innehålla andra noder och/eller dokument. Noderna kan liknas vid bokhyllor i ett bibliotek — med den viktiga skillnaden att ett och samma dokument i Darc kan placeras på alla de ställen där det logiskt hör hemma.

I indexsökningen finns en återkoppling till vymekanismen så att dokument kan sökas med både traditionella sökmetoder eller genom att användaren interagerar med vyer. Presentationsverktyget å sin sida känner till vilka index som är definierade; ett ord kan markeras i fulltextfönstret och genast sökas upp i valfritt index, låt säga en förteckning över alla författarnamn eller alla publiceringsorter.

Databashanteraren, vyerna och presentationsverktyget är således tätt integrerade. Ett rättesnöre vid konstruktionen av Darc har varit att så långt som möjligt försöka göra fullt bruk av de potentiella fördelar som SGML-märkningen erbjuder.

4. Samarbetspartners och kontakter, Sverige

Samarbete med redaktionen för Nationalupplagan av August Strindbergs samlade verk. Våra kontakter har varit med professor Gunnel Engwall, ordförande i nationalupplagans ledningsgrupp, samt med docent Lars Dahlbäck, nationalupplagans huvudredaktör. Donald Broady har utformat en plan för ett projekt "Det nya handbiblioteket", med Lars Dahlbäck som medsökande. Detta nya projekt skall bl.a. ägnas åt utprövning av en metodik för SGML-märkning av Strindbergsmaterialet. Forskningsrådsnämnden beviljade i juni 1993 forskningsmedel för detta projekt som skall pågå i tre år.

Kontakter med Nationalencyklopedins redaktion (Kari Marklund, Arne Ekman). Projektmedarbetare har besökt redaktionen i Höganäs för att studera det dokumenthanteringssystem som används för produktionen av Nationalencyklopedin. Kari Marklund har vid flera tillfällen besökt oss i Stockholm för att studerat de dokumenthanteringsverktyg som utvecklats inom projektet.

Ett samarbete har inletts med Ericsson Telecom (Peter Witt, Gunbritt Jonsson m fl). Närmast planerar Ericsson Telecom att beta-testa SGML Darc.

Sammanträffanden med ett flertal befattningshavare vid Skolverket. En tidig version av Darc demonstrerades på Skolverket hösten 1992. Det mest konkreta samarbetet hittills rör redskap för att ordna och använda dokumentbaser bestående av de tillståndsbeskrivningar som undervisningsråden sänder in till Skolverket; vår kontaktman i detta arbete är Göran Isberg.

Rörande frågor om upphovsrätt till elektronisk information samarbete med professor Peter Seipel, Institutet för rättsinformatik, Stockholms Universitet. Bl.a. genomför f.n. en av professor Seipels doktorander, Martin Brinnen, en utredning rörande de upphovsrättsliga problem som aktualiserats inom projektet. Särskilda medel för denna utredning har erhållits från NUTEK och BIBSAM/Kungliga biblioteket. Närmare bestämt rör det sig om frågan om hur t.ex. lärare skall kunna använda maskinläsbar litteratur som "råmaterial" för framställning av skraddarsydda läromedel.

Inom projektet har sedan något år förberedelser pågått för att skapa en svensk intresseförening för SGML, ansluten till motsvarande internationella organisation. Härvid har vi samarbetat med företag inom tryckeribranschen (särsk. Ulf Larson, Elanders tryckeri). Inledande diskussioner om SGML-användarföreningens syfte och organisering fördes under ett seminarium om elektronisk dokumenthantering anordnat av SIS-STG (Standardiseringskommissionen i Sverige - Allmänna

Standardiseringsgruppen) i Stockholm den 14 april 1994. Föreningen konstituerades vid en SGML-konferens på KTH den 26 augusti 1993.

Cap Gemini Logic och Cap Programator har länge arbetat med ett projekt för en regelbundet utkommande CD-ROM med EG-dokumentation, text för skolbruk. Här har vi deltagit i diskussionerna om lämpliga dokumentformat och presentationsverktyg; våra kontakter har varit Roland Lundgren, Peter Eriksson och André Ekespong. Bland övriga intressenter bör främst nämnas Svenska Dagbladet, vars journalister förväntas bidra med själva artikelmaterialet. Nyligen har ett annat företag övertagit huvudansvaret för detta projekt, men CAP Programator och SvD skall enligt planerna alltså vara delaktiga, och arten av vår medverkan diskuteras f.n.

Elektronisk utgivning av material av relevans för skolväsendet har diskuterats med representanter för CE Fritzes, Liber och Norstedts, företag som numer tillhör samma koncern. Våra kontakter har varit Patric Hamilton och Dan-Gunnar Sahlström, CE Fritzes; Barbro Folke, Liber; samt Mikael Wallgren, Norstedts.

Samarbete med Institutionen för ämnesdidaktik i lärarutbildningen, Göteborgs universitet, rörande material för matematikundervisning. Våra kontakter har varit matematikdidaktikerna Bengt Johansson, Göran Emanuelsson och Ronnie Ryding. Under vintern 1992-93 företogs inom projektets ram en utvärdering av ett nytt amerikansk elektroniskt läromedelspaket, MathFINDER CD-ROM, innehållande material till drygt tusen lektioner i matematik. Dessa lektioner skall illustrera de undervisningsprinciper (vanligen helt enkelt kallade "the Standards") som utarbetats av NCTM National Council of Teachers of Mathematics. Materialet levereras på en kompaktskiva, som utöver lektionsförslagen innehåller hela "the Standards"-texten. Skivan har framställts med anslag från National Science Foundation. Utvärderingen genomfördes av matematiklärarna Bo Ekbäck och Jan Wahlgren, knutna till Institutionen för ämnesdidaktik i lärarutbildningen, Göteborgs universitet, och resultaten publicerades som en artikel "CD-skiva i stället för räknebok" i matematiklärarnas tidskrift *Nämnamn* nr 2 1993.

Deltagande i en arbetsgrupp för elektronisk publicering av historiskt källmaterial för bl.a. skolbruk, främst material som belyser Stockholms historia. I denna arbetsgrupp deltar representanter för Stockholms skolor, Stockholms stadsarkiv (Stockholms historiska databas), Vasamuseet, Krigsarkivet, Nordiska museet, Historiska museet, Stockholms Stadsmuseum. F.n. sysslar arbetsgruppen med frågan om elektronisk publicering på CD-ROM av Stockholms Stads tänkeböcker, till att börja med 1626 års tänkebok. Tänkeböcker innehåller rättegångsprotokoll och tillhör de förnämsta källorna till kunskap om människors vardagsliv i Stockholm i gångna tider.

Kontakter med Ellemtel. Våra kontakter har varit Pär Karlsson och Jan Olov Lindberg. Emmettel önskar betatesta SGML Darc.

Tillsammans med Institutionen för Teleinformatik (prof Björn Persson) och Institutionen för data- och systemvetenskap, Stockholms universitet (tf prof Carl-Gustaf Jansson) planering av ett nytt forskningsprojekt "Networked education. Course development and education in a distributed multimedia environment". Avsikten är att utveckla verktyg för en tidsenlig "distansundervisning", där lärare och deltagare är geografiskt spridda och har kontakt via video, ljud, etc och dessutom tillgång till gemensamt kursmaterial. En ansökan härom inlämnades till NUTEK i mars 1993. Denna ansökan kommer att behandlas inom ramen för ett nytt forskningsprogram som startar 1 januari 1994.

Telub har för ändamål inom svenska försvaret föreslagit ett gemensamt utvecklingsprojekt som utnyttjar den teknologi vi tagit fram i Darc. Våra kontakter har varit Lars Jernbäcker och Thomas Ragnarsson. Vidare har vi haft kontakt med Erik Häggblad från Försvarets Materielverk. Likaledes från FMV har Sven-Åke Lindberg, byrådirektör för teknisk dokumentation, varit i kontakt med oss.

Vi har haft många besökare som kommit till KTH eller HLS för att se demonstrationer av de verktyg för dokumenthantering som utvecklats inom projektet. Utöver svenska och utländska forskare och lärare som är för många för att här räknas upp (bl.a ett par hundra datalärare), kan nämnas representanter för Ellemtel, Ericsson Telecom, Telub, Försvarets materialverk, Infologics, Liber och Fritzes, Skolverket, Nationalencyklopedien.

Vi har även själva tagit med de verktyg för dokumenthantering som utvecklats inom projektet och demonstrerat dessa på annan ort, bl.a. i Göteborg för representanter Volvo Data, Volvo personvagnar samt Volvo lastvagnar, i Oslo för den norska SGML-användarföreningen, i Italien för expertis inom hypertextområdet.

5. Samarbetspartners och kontakter, internationella

Vid Oxford University pågår sedan länge sammanställningen av en omfattande corpus (det kan bli flera gigabyte) av SGML-märkta engelskspråkiga texter som skall göras elektroniskt tillgänglig för forskning och undervisning. Vi kontaktades i juli 1993 av Lou Burnard (ä. ansvarig för samordningen av forskarnätverket TEI) som önskar utprova möjligheten att SGML Darc levereras som dokumenthanterare tillsammans med denna textcorpus. Diskussioner pågår, och en preliminär överenskommelse har träffats. Detta samarbete är självfallet av stort intresse för oss, det skulle innebära att våra verktyg skulle spridas till och prövas av många användare i kvalificerade forsknings- och undervisningssammanhang.

Samarbete med EG-projektet TIDE (the European Communities DG XIII TIDE project), via CAPS, som leds av prof Jan Engellen, Leuven, Belgien. Övriga deltagare är Katholieke Universiteit, Leuven, Belgien; University of Bradford, UK; Royal National Institute for the Blind, Peterborough, UK; Infologics, InfoVox Department, Sverige. Eventuellt kommer SGML Darc att användas inom CAPS, utprovning pågår.

Kontakter rörande hypertextproblem med forskare, lärare och systemutvecklare vid Brown University, Providence, Rhode Island, bl.a. åtskilliga sammanträffanden med Paul Kahn.

Beta-testning av SGML Darc vid University College, Cork, Irland, där vår kontaktman är Peter Flynn som arbetar med medeltida handskrifter och är en specialist inom både SGML-området och T_eX-området.

Beta-testning av SGML Darc vid Falch Hurtigtrykk, Norge. Kontaktman Steve Pepper.

Kontakter med Information Management & Engineering Ltd, London (som tillverkar bibliotekssystemet TinLib), bl.a. sammanträffanden med företags vd Peter Noerr.

Kontakter inom det internationella forskarnätverket TEI (Text Encoding Initiative), som utarbetar riktlinjer för beskrivande märkning av humanistisk och samhällsvetenskaplig litteratur.

Inledande diskussioner om framtida informationsteknologisk forskning inom ramen för EG's biblioteksforskningsprogram (DG XIII/E/3). Här har vi haft kontakter med bl.a. Chris Rowley vid Open University, London. Närmare bestämt rör det sig om utveckling av verktyg för att koppla dokumentbeskrivningsstandard (SGML) till typografisk standard (L_aT_eX), vilket är ett viktigt område exempelvis vid produktion av skraddarsydda läromedel. Vi gjorde i årsskiftet 1992/93 ett försök att sy ihop ett internationellt samarbete mellan forskarmiljöer i Tyskland, England, Sverige och ytterligare några länder, vilket dock inte lyckades inte i tid för den ansökningsomgång för EG-programmet som avslutades 5 februari 1993; en bidragande orsak till detta misslyckande var att ESS-

avtalet sköts på framtiden (eftersom schweizarna röstade nej), vilket fick till följd att svenska forskare inte kunde bli fullvärdiga partners i EGs forskningsprogram. Dock avser vi att återkomma. Projektmedarbetare har deltagit i seminarier i detta ämne (en slags grundkurs i konsten att söka medel från EG's forskningsprogram) i Stockholm 6 nov och 12 dec 1992; arrangör var BIBSAM, Kungl. biblioteket.

Diskussioner om samverkan med Electronic Book Technologies, Providence, Rhode Island, USA, ett ledande företag på SGML-området, specialiserat på presentationsverktyg för "elektroniska böcker" (systemet DynaText). Vår kontakt har varit företagets vd Lou Reynolds.

Det kanadensiska företaget SoftQuad (ett av de ledande företagen inom SGML-mjukvarubranschen) har dels hört sig för om licensiering av presentationsverktyget i Darc, dels föreslagit utvecklingssamarbete; vi inleder kontakten genom att SoftQuad får utprova Darc med start i september 1993.

Vi har själva kontaktat Exoterica Corporation, Ontario, Kanada, ytterligare ett av de fåtaliga ledande företagen för SGML-programvara. Vår kontakt är företagets vd John McFadden.

Förutom dessa rent konkreta kontakter får vi ständigt förfrågningar om Darc från hela världen: Nokia i Finland, US Navy, US Air Force, Del Marva Power Company (New Jersey), Interleaf; samtliga i USA. Diwan Tech i England, som enligt egen uppgift gör program för tidningsbranschen, har hört sig för om licensiering av Darc. I samband med seminarier, föredrag och liknande har vi haft samtal med bl a representanter för det tyska företaget Atlas Elektronik, samt Blinden-Anstaltung i Frankfurt am Main, Geco-Prakla i Holland, Fabricius i Norge m.fl. Till detta tillkommer ytterligare brevkontakter från t ex Japan, Australien, Italien, Singapore, Ungern och Sydafrika samt från olika forskningscentra.

Slutligen kan nämnas att vi påverkat utformningen av de SGML-mallar som Open Software Foundation nyligen offentliggjort, samt att vi anlitas av National Science Foundation i USA för att avge utlåtande om en SGML-relaterad projektansökan.

6. Av projektet arrangerade sammankomster (urval)

Seminarium "Kunskapsverkstaden - om hypertextsystem för undervisningsändamål", HLS, Stockholm, 22 okt 1990. Deltagare: ett drygt femtiotal lärare, forskare och studerande vid HLS.

Seminarium "Datorstöd för kunskapsarbete", HLS, Stockholm, 24 okt 1990. Ca 40 deltagare. Syftet var att sammanföra representanter för utbildningssektorn med datalogiska forskare och representanter för informationsteknologiska företag. Deltagare kom bl.a. från Utbildningsdepartementet; Skolöverstyrelsen; Svenska Kommunförbundet; Landstinget, Stockholm; Statskontoret; Teknisk databehandling, Umeå universitet; Högskolans Råd för grundutbildning, UHÄ; ITYP-programmet, STU; Naturens hus, Stockholm; Nationalencyklopedin, Bokförlaget Bra Böcker, Höganäs; IBM Svenska AB; SUN Microsystems Svenska AB; Programator Stockholm AB; Apple Computer AB; Rank Xerox AB; Lunds Universitet; Institutionen för ADB och Datalogi, Uppsala universitet; Institutionen för nordiska språk, Stockholms universitet; Odontologiska fakulteten, Lunds universitet; HCID-Studium, Informationsbehandling, Umeå universitet; Nada, KTH; Datorrådet, KTH; Grafisk teknik, KTH; Didaktikcentrum HLS; Inst. f pedagogik, HLS; ITC, HLS; Rektorsämbetet, HLS; Biblioteket, HLS.

Workshop om copyright och elektronisk publicering, Högskolan för lärarutbildning i Stockholm, 23 sept 1992. Denna workshop anordnades som ett samarbete mellan projektet Datorstöd för

kunskapsarbete, där frågorna väcktes, och FRN (Jan Hagerlid) och Institutet för rättsinformatik, Stockholms Universitet (Peter Seipel). Frågorna gällde hur upphovsrättsfrågor och avtalsjuridiska frågor skall hanteras för att maskinläsbar information skall kunna göras tillgänglig för lärare, forskare, studerande etc. För utbildningsväsendet är det ett akut behov att råda bot för den upphovsrättsliga och avtalsjuridiska oklarheter som i dag hindrar förlag, författare, bildproducenter och andra från att göra sitt material tillgängligt för spridning och användning i maskinläsbar form. Deltagarna representerade NADA, KTH; Datorrådet, KTH; Institutionen för lingvistik, Stockholms universitet, Högskolan för lärarutbildning i Stockholm; LIBLAB-IDA, Linköpings universitet; BIBSAM, Kungl. Biblioteket; Bibliotekstjänst, Lund; Forskningsrådsnämnden; Centrum för biblioteks- och informationsvetenskap, Göteborg; Institutet för rättsinformatik, Stockholms universitet; Libris-avd., Kungl. Biblioteket; Centrum för interaktiv video, Tandläkarhögskolan; Sveriges Lantbruksuniversitet, Uppsala. Workshopen leddes av professor Seipel, som torde vara Sveriges främste expert i dessa frågor, och arbetet fortsätter i form av några mindre utredningar, finansierade av NUTEK och BIBSAM/Kungl. biblioteket.

Seminarium, Stimdi (Sveriges tvärvetenskapliga intresseförening för människa-dator-interaktion), sammankomst i Stockholm, 4 febr 1993, föredrag och demonstration av projektets verktyg för dokumenthantering.

Konferens SGML Sweden '93, KTH, Stockholm. Projektmedarbetare deltog i organisationskommittén och i ansvaret för förberedelser och genomförande. Bakgrunden är följande: För att uppmuntra spridningen av SGML beslöt vi hösten 1992 att starta en SGML-användarförening. Ulf Larson vid Elanders Tryckeri hade samma planer vilket mynnade ut i ett fruktbart samarbete där vi tillsammans med Karita Thomé från SIS-STG och Roswitha Graham från KTH:s Datorråd den 26 augusti 1993 arrangerade en konferens där föreningen konstituerades.

Intresset för denna konferens översteg alla förväntningar. Bland de ca 130 deltagarna fanns representanter för AAB Automation AB, AAB Metallurgy AB, AB Volvo Teknisk utveckling, Albert Bonniers förlag, Alfa Laval Separation AB, Arméns Tekniska Centrum, Bibliotekstjänst AB, Bokförlaget Natur och Kultur, Bonnierförlagen AB, DATA, CelsiusTech Electronics, CBG Konsult, CE Fritzes AB, Computer Publishing Systems AB, DataOriginal, Datorrådet/KTH, Digital, Elanders tryckeri, Ellemtel, Ericsson Programatic AB, Ericsson Programatic Sweden AB, Ericsson Radar Electronics, Ericsson Radio Systems, Ericsson Telecom AB, Ferrator AB, FFV Aerotech AB, Försvarets Materielverk, Grafisk Teknik/KTH, Grafiska Företagen, Graphics Systems, Hägglunds Vehicle AB, IBM Svenska AB, IDA Linköpings universitet, Informationslogik AB, Institutet för Verkstadsteknisk Forskning IVS, Institutionen för lingvistik/Stockholms universitet, Interleaf Scandinavia AB, IOS Svenska AB, ITS, IVF, Jordbruksverket, Karlskronavarvet, KTH Bibliotek, Kungl. Biblioteket/BIBSAM, LM Ericsson, Microsoft AB, NADA/KTH, Next Publishing AB, Norstedts förlag, Norsys Technology AB, Optosoft AB, PCD-konsult/CD-ROM Plaza, Physica Scripta, Posten Systemservice, Publishing Development, Regeringskansliets förvaltningskontor, SAAB Automobile AB, SAAB Service Partner AB, SAAB-Scania AB, Scandface, Scandinavian Data Resource AB, SCANIA DATA AB, SJ-DATA Forskn. & utveckl. AB, Skolverket, Standardiseringskommissionen i Sverige - Allmänna Standardiseringsgruppen SIS-STG, Standardiseringskommissionen i Sverige SIS, Studentlitteratur AB, SUNET/KTH, Sörman Information AB, Telia Research AB, Telub Inforum, TeXcel Data AB, Translator Teknikinformation AB, TS DATA/KTH, Utbildningsbyrån AB, Volvo Data AB, Volvo Lastvagnar, WordPerfect Sverige AB.

Den stora tillströmningen till konferensdagen och användarföreningen utgjorde ett tydligt bevis för att ledande företag och organisationer i Sverige just nu är synnerligen intresserade av SGML. Samtidigt var det upplysande att flertalet som svar på en fråga i plenum genom handuppräknning intygade att de hade föga egen erfarenhet av arbete med SGML. Utöver bildandet av användarföreningen ägnades konferensdagen åt föredrag. Vidare demonstrerades SGML-relaterade programvaror av ett drygt dussin företag, vilket gav en god överblick över de redskap som f.n. är mest utvecklade på området.

7. Deltagande i övriga konferenser och seminarier i Norden

ITYP-Forum 20 nov 1991. Projektet presenterades med posters.

ITYP-forum 11 dec 1991. Projektet presenterades med posters.

ITYP-forum 6 maj 1992. Projektet presenterades med posters.

Dataläraryrökens 4:e Ostkustkonferenskonferens, Stockholm, 27—28 nov 1992. Föredrag samt demonstration av projektets verktyg för dokumenthantering.

Lärardag Stockholm den 19 mars 1993. Projektet presenterades med föredrag för gymnasielärare.

ITYP-forum 25 mars 1993. Projektet presenterades med poster samt demonstration av projektets verktyg för dokumenthantering.

Digital dokumenthantering, 31 mars 1993, seminarium för industrirepresentanter anordnat av NFI (Nordic Forum for Information), Stockholm. Projektet presenterades med föredrag.

Seminarium om elektronisk dokumenthantering anordnat av SIS-STG (Standardiseringskommissionen i Sverige - Allmänna Standardiseringsgruppen) 14 april 1994.

STIMDI (Sveriges tvärvetenskapliga intresseförening för människa-dator-interaktion) årsmöte, Lund, 3-4 maj 1993. Projektet presenterades med föredrag.

Årsmöte, SGML-brugergrupp i Norge, Oslo, 22 juni 1993. Projektet presenterades med föredrag, samt demonstration av projektets verktyg för dokumenthantering.

8. Deltagande i internationella konferenser

ECHT '90 (1^{ère} conférence européenne sur les hypertextes), Versailles, nov 1990.

SGML-symposium, Paris, maj 1991.

Hypertext '91, San Antonio, Texas, USA, dec 1991.

8:e Nordiska konferensen för information och dokumentation, Helsingborg, maj 1992. Föredrag om projektet.

SGML '92, Danvers, Mass., USA, okt 1992.

Nordic TeX, Köpenhamn, Danmark, nov 1992. Föredrag om projektet.

ECHT '92 (The fourth ACM Conference on Hypertext 1992), Milano, 30 nov-4 dec 1992. Demonstration av de dokumenthanteringsverktyg som utvecklats inom projektet.

Documentation Europe '93, Rotterdam, 12-14 maj 1993. Föredrag om projektet.

9. Rapporter, uppsatser

D. Broady: *Datorstöd för kunskapsarbete. Om hypertexttillämpningar i utbildningssammanhang*, KTH, NADA, okt 1990.

D. Broady: "Informasjonsteknologi for et desentralisert skolesystem". *Uddanningskvalitet — styrbar eller ustyrbar? Om målstyring og kvalitetsvurdering av norsk skole* (red. Marit Granheim, Ulf P. Lundgren och Tom Tiller). Oslo: Tano, 1990, pp. 300-310.

H. Haitto: *SGML-relaterade verktyg för intermediära former och dokumenthantering*. IPLab-rapport nr 37, 1991.

H. Haitto: *Rapport från SGML-symposium i Paris maj 1991*. IPLab Arbetspapper nr 27, 1991.

D. Broady: "Lokala dokumentbaser för undervisningsbruk", *Teknologi och kompetens/Technology and Competence*. Proceedings from the 8th Nordic Conference on Information and Documentation, pp. 59-63.

D. Broady: "Kunskapsverkstaden — om lokala dokumentbaser som arbetsverktyg för lärare", *Datorn i Utbildningen*, årg. V, nr 2 1992, pp. 22-28. [Tidigare utgiven under samma titel som IPLab-rapport nr 47, Institutionen för numerisk analys och datalogi, KTH, 1991].

D. Broady: "Lärarnas kunskapsbanker", *Datorn i utbildningen*, nr 4 1992.

D. Broady, H. Haitto, P. Lidbaum och M. Tobiasson: *Darc — Document Archive Controller*, IPLab-rapport 63, TRITA-NA-P9306, 1993.

B. Ekbäck och J. Wahlgren: "CD-skiva i stället för räknebok", *Nämnnaren, tidskrift för matematikundervisning*, årg XX, nr 2 1993, pp. 35-37.

D. Broady och H. Haitto: "Flexible management of SGML-encoded documents, *Documentation Europe '93. Conference Proceedings*. Alexandria, Vancouver, USA: Graphic Communications Association, 1993. (elektron. publikation som DynaText-bok)

D. Broady, Hasse Haitto, Peter Lidbaum och Magnus Tobiasson: *Navigational Issues and Multi-User Layered Hypertext in the Darc System*. Prel. version. IPLab, Institutionen för numerisk analys och datalogi, KTH, 1993.

Samt kortare presentationer av projektet i bl.a. Medlemsbladet för STIMDI (Sveriges tvärvetenskapliga intresseförening för människa-dator-interaktion), tidningen *Datornytt* mm.

Bilagor

1. D. Broady: "Informasjonsteknologi for et desentralisert skolesystem". *Uddanningskvalitet — styrbar eller ustyrlig? Om målstyring og kvalitetsvurdering av norsk skole* (red. Marit Granheim, Ulf P. Lundgren och Tom Tiller). Oslo: Tano, 1990, pp. 300-310.
2. H. Haitto: *SGML-relaterade verktyg för intermediära former och dokumenthantering*. IPLab-rapport nr 37, 1991.
3. D. Broady: "Lokala dokumentbaser för undervisningsbruk", *Teknologi och kompetens/Technology and Competence*. Proceedings from the 8th Nordic Conference on Information and Documentation, pp. 59-63.
4. D. Broady: "Kunskapsverkstaden — om lokala dokumentbaser som arbetsverktyg för lärare", *Datorn i Utbildningen*, årg. V, nr 2 1992, pp. 22-28. [Tidigare utgiven under samma titel som IPLab-rapport nr 47, Institutionen för numerisk analys och datalogi, KTH, 1991].
5. D. Broady: "Lärarnas kunskapsbanker", *Datorn i utbildningen*, nr 4 1992.
6. D. Broady, H. Haitto, P. Lidbaum och M. Tobiasson: *Darc — Document Archive Controller*, IPLab-rapport 63, TRITA-NA-P9306, 1993.
7. D. Broady och H. Haitto: "Flexible management of SGML-encoded documents, *Documentation Europe '93. Conference Proceedings*. Alexandria, Vancouver, USA: Graphic Communications Association, 1993. (elektron. publikation som DynaText-bok)
8. D. Broady, Hasse Haitto, Peter Lidbaum och Magnus Tobiasson: *Navigational Issues and Multi-User Layered Hypertext in the Darc System*. Prel. version. IPLab, Institutionen för numerisk analys och datalogi, KTH, 1993.