

# **DEN FLERSPRÅKIGA BIBLIOTEKSKATALOGEN**

**Hantering av icke-latinsk skrift på bibliotek**

**Maria Sjöberg  
Anna Striberger**

Examensarbete (20 poäng) för magisterexamen i Biblioteks- och informationsvetenskap vid Lunds universitet.

Handledare: Göran Gellerstam

BIVILs skriftserie 2006:20

ISSN 1401:2375

© Lunds universitet. Biblioteks- och informationsvetenskap 2006.

## **Title**

The Multilingual Library catalogue – Handling of non roman script in libraries

## **Abstract**

The main purpose of this thesis has been to investigate librarians opinions about the need for and possibility to use non-roman script in library catalogues. We have also investigated the routines of handling materials in non-roman scripts in two Swedish libraries. The investigation was made in the light of the recent shift in character set in the Swedish National Library Catalogue Libris. This change will enable use of multiple writing systems in the future.

Our findings are that a multicultural society requires special consideration to the needs of citizens who speak other languages than the majority of the population. This becomes difficult when the languages spoken have a different writing system than the one used by the majority. The solution to transcribe or transliterate these languages has shown problematic especially in terms of information retrieval. Finding and hiring library educated staff with knowledge of different languages is also a problem, especially for libraries who cater to many diverse ethnic groups. Since handling non-roman script in computer systems has progressed in recent years it is now possible to use multiple writing systems in library catalogues. The most wide-spread character set for using multiple writing systems is unicode, which is why we have focused on this character set in our thesis. The use of unicode in Swedish library systems is progressing slowly however, and some libraries are developing local solutions to the problem with non-roman script in the catalogue. This despite the fact that a uniform standard is more efficient for sharing information in a computer based environment.

Our empirical study showed that there is a clearly stated need to adjust library services to fit the needs of users whose languages differ from the majority in Sweden, especially then these languages have non roman script.

Our method is based on empirical research through qualitative interviews as well as extensive literature studies to give our study a firm theoretical background.

## **Keywords**

Icke-latinsk skrift

Unicode

Romanisering

Flerspråkig biblioteksverksamhet

Standardisering

# INNEHÅLLSFÖRTECKNING

<b>Inledning .....</b>	<b>4</b>
<b>Syfte och frågeställningar .....</b>	<b>6</b>
Syfte .....	6
Frågeställning .....	6
Avgränsningar .....	6
Disposition .....	7
Definitioner .....	7
<b>Metod.....</b>	<b>9</b>
Urval.....	9
Metod .....	10
Genomförande.....	11
Etiska aspekter.....	11
<b>Flerspråkighet på bibliotek.....</b>	<b>13</b>
Demografi och statistik .....	13
Mångkulturell och flerspråkig biblioteksverksamhet.....	13
Språkets betydelse .....	18
<b>Romanisering.....</b>	<b>20</b>
Transkribering och translitterering som metod .....	20
Historik.....	20
Kritik mot romanisering.....	21
<b>Unicode som teckenuppsättning.....</b>	<b>25</b>
Utvecklingen före unicode .....	25
Unicode .....	28
<b>Standardisering .....</b>	<b>35</b>
Utbyte av bibliografiska poster .....	35
MARC-formatet .....	35
MARC och unicode.....	39
<b>Empirisk studie .....</b>	<b>42</b>
Presentation av biblioteken, bibliotekssystemen och informanterna .....	42
Resultat av intervjuerna.....	47
Libris' uppgradering till unicode .....	61
<b>Analys .....</b>	<b>63</b>
Implementering av unicode .....	63
Standardisering – MARC-formatet och dess förutsättningar.....	64
Transkribering och translitterering.....	66
Språkkompetens hos personalen .....	67
Brister i kommunikationen mellan avdelningar .....	68
Kritik mot samt alternativ till unicode för teckenhantering .....	69
<b>Slutsatser och slutdiskussion .....</b>	<b>70</b>
<b>Litteraturförteckning .....</b>	<b>74</b>
<b>Bilagor .....</b>	<b>78</b>
Bilaga 1 – Ordlista .....	78
Bilaga 2 – Intervjuguide.....	79



”If you are a programmer working in 2003 and you don't know the basics of characters, character sets, encodings, and Unicode, and I *catch* you, I'm going to punish you by making you peel onions for 6 months in a submarine. I swear I will.”<sup>1</sup>

---

<sup>1</sup> “The Absolute Minimum Every Software Developer Absolutely, Positively Must Know About Unicode and Character Sets (No Excuses!)” av Joel Spolsky, 2003

## Inledning

Vi lever idag i ett samhälle där mångkultur och flerspråkighet är ett faktum. I Malmö kommun fanns år 2005 så många som 169 länder representerade och 26 % av invånarna var födda i utlandet (Malmö stad -02, 2006). Vid Lunds universitet studerar över tusen internationella studenter varje år<sup>2</sup>. Detta ger en miljö där flera olika språk möts. I mötet mellan det allt mer globaliserade samhället och biblioteket finns en del intressanta problemområden, både vad gäller kulturella skillnader och språkliga.

I vår uppsats har vi valt att fokusera på den språkliga aspekten, och framför allt det skrivna språket. Ett exempel där språkproblemet blir tydligt är inom det område inom biblioteks- och informationskunskap som kallas kunskapsorganisation och då framför allt vid hantering av bibliografiska poster vid katalogisering, särskilt när det gäller material på språk som inte använder latinska bokstäver. När ett språks skrivsystem inte är baserat på latinska skrivtecken kräver de flesta svenska bibliotekskataloger idag att originalskriften eller originaltecknen transkriberas eller translittereras. Ett problem uppstår då formen för denna transkribering eller translitterering inte alltid är standardiserad, vilket innebär att det kan finnas mer än ett system för samma språk. Det finns även en risk att transkriberingen eller translittereringen blir felaktig på grund av bristande kunskaper hos den personal som hanterar materialet.

Då bibliografiska poster är mycket starkt styrda av standardisering genom användandet av exempelvis MARC-formatet uppstår ett problem i mötet mellan transkribering eller translitterering och katalogisering. De olika transkriptionsscheman som används kan ge en rad olika varianter på stavningar av exempelvis en författares namn, vilket kräver att den som söker i katalogen vet vilken av stavningarna som katalogisatören använt sig av, för att få träff på sin sökning.

Som ett led i den ökade globaliseringen finns idag den tekniska möjligheten att använda sig av alla världens teckensystem i datorsammanhang, något som tidigare inte varit möjligt då datorerna från början var anpassade efter det latinska alfabetet. I och med uppkomsten av unicode, en teckenuppsättning som möjliggör användandet av all världens skrivsystem, finns i princip inga hinder mot att använda sig av vilket skrivsystem som helst i vanlig datormiljö. Detta innebär att det även är möjligt att använda sig av originalskrift i bibliotekskataloger som har implementerat unicastandarden. Möjligheten att använda sig av originalskrift i bibliotekskataloger är något som flera bibliotek sedan länge utnyttjat runt om i världen, framför allt har det varit nödvändigt i de länder som har språk med icke-latinska skrivsystem. Ofta har det

---

<sup>2</sup> Mejl från programhandläggare på Internationella sekretariatet på Lunds universitet.

dock rört sig om lokala lösningar i respektive land, vilket förhindrar utbyte av information mellan länder i datormiljö. Tekniken och praxisen har funnits länge i länder där man använder sig av icke-latinska tecken, exempelvis i Kina och Japan, men förekommer nu även i bibliotekskataloger i exempelvis Frankrike, USA och England.

För att kunna dra nytta av de katalogposter som görs på bibliotek över hela världen krävs att de olika systemen för katalogisering och teckenkodning är kompatibla. I dagsläget är MARC-formatet det rådande för katalogisering och unicode spås bli den rådande teckenuppsättningen. För att förena icke-latinsk skrift med MARC-formatet krävs en stor mängd anpassningar och regler, något som Library of Congress i USA arbetat med inför sin övergång till en unicodeanpassad katalog och därmed även visat vägen för de bibliotek som önskar följa denna utveckling.

Då arbetet med denna uppsats påbörjades tidigare i år var den svenska nationella bibliotekskatalogen Libris inte unicodeanpassad. En övergång var dock planerad och genomfördes i slutet av maj i år. Detta innebär att man nu har öppnat upp för möjligheten att katalogisera med originalskrift, vilket inte har varit möjligt i den gemensamma katalogen tidigare. Vi betonar dock att möjligheten finns, men att Libris än så länge avråder från att aktivt katalogisera med icke-latinsk skrift då man inte till fullo utrett hur detta kommer att påverka de kringssystem som finns (Libris katalogiseringsklient – nyheter).

Vi önskar med den här uppsatsen utreda hur biblioteken ser på den förändring det skulle innebära att använda sig av böckernas originalspråk i bibliotekets katalog. Vi avser att undersöka bibliotekspersonalens attityd till att använda originalskrift i bibliotekskatalogen, vilket skulle öka sökmöjligheterna av material på språk med andra teckensystem än det latinska, men troligtvis även innebära en större arbetsbelastning för biblioteken. Med detta följer även frågor kring vad gemensamma standarder som MARC-formatet och enhetliga teckenuppsättningar betyder för bibliotekets verksamhet samt hur biblioteken lever upp till sin roll att serva alla medborgare när en stor del av användarna på grund av språkbarriärer har svårt att hitta material i bibliotekens kataloger. För att kunna undersöka hantering av icke-latinsk skrift i bibliotekskatalogen, som är vårt huvudfokus, måste vi även lyfta blicken och undersöka den kontext som finns, nämligen hur material med icke-latinsk skrift hanteras på bibliotek vad beträffar exempelvis förvärv och selektion, transkribering och translitterering samt sökhjälp för besökare med annat modersmål än majoritetsbefolkningen.

## Syfte och frågeställningar

### Syfte

Syftet med uppsatsen är att *undersöka bibliotekspersonalens attityder och inställningar till hantering av material med icke-latinsk skrift samt utreda förekomsten icke-latinsk skrift i bibliotekskatalogen.*

Detta är ett högaktuellt ämne då en unicodeanpassning av Libris kommer att göras i slutet av maj 2006, alltså strax efter det att vår undersökning ute på biblioteken avslutats. Tyvärr kommer inte unicodeanpassningen i detta skede innebära att biblioteken får möjlighet att använda icke-latinsk skrift i sina kataloger då Libris avråder från detta i dagsläget, men strategier eller åsikter kring fenomenet med icke-latinsk skrift i bibliotekssammanhang finns ute på de aktuella biblioteken, och dessa avser vi fånga upp med vår uppsats.

### Frågeställning

Med vår undersökning önskar vi kartlägga de tankar och funderingar som finns bland bibliotekspersonalen på de aktuella biblioteken kring den problematik det innebär att hantera originalskrift i bibliotekssystemet. Vi vill även utreda huruvida det finns några uttalade strategier eller handlingsplaner för att i framtiden arbeta med icke-latinsk skrift i katalogen. Vår frågeställning lyder alltså som följer:

Hur ser bibliotekspersonalen på hantering av material på språk med icke-latinsk skrift på de aktuella biblioteken? Vad skulle det innebära att använda icke-latinsk skrift i bibliotekskatalogen och vilka är inställningarna till detta ute på biblioteken?

### Avgränsningar

För att kunna hantera icke-latinsk skrift i datormiljö krävs en teckenuppsättning som kan visa denna skrift. Tekniken bakom en sådan teckenuppsättning är oftast komplicerad och den teckenuppsättning som vi har valt att se närmare på i denna uppsats, unicode, är inget undantag. Vi är inte i första hand intresserade av att utreda de tekniska förutsättningarna i denna uppsats, då detta redan gjorts i tidigare arbeten, men har tagit upp grunderna för att ge en förståelse för vad som ligger bakom teckenuppsättningar som unicode samt hantering av bibliografisk data i exempelvis MARC-formatet. Huvudfokus för uppsatsen ligger på hantering av material på språk med icke-latinsk skrift, vilket vi utreder med hjälp av tidigare forskning samt genom en empirisk undersökning i form av kvalitativa intervjuer.



Vi utreder förutsättningarna i Sverige men har även tagit exempel från framförallt USA eftersom det här finns en stor mängd undersökningar inom mångkulturella frågor för bibliotek och dess verksamhet, vilket har visat sig komplettera de slutsatser vi dragit om de svenska förhållandena.

Vår empiriska undersökning hade kunnat göras med två olika perspektiv, nämligen ur personalens alternativt användarnas perspektiv. Vi har valt att fokusera på bibliotekspersonalens perspektiv i första hand, eftersom denna har en inblick i verksamheten och systemens möjligheter på ett sätt som inte användarna har. För vårt syfte hade det inte varit lika fruktbart att göra denna undersökning med användare då dessa troligen inte är lika införstådda med problematiken, vilket antagligen skulle ha gett ett allt för magert material för en analys.

## Disposition

Uppsatsen är indelad i tio huvudavsnitt. De tre första delarna består av en inledning, syfte och frågeställningar samt metod. Därefter följer fyra delar som kan ses som en teoretisk bakgrund till vår empiriska undersökning. De fyra delarna behandlar nyckelområden för uppsatsen och är följande: flerspråkighet på bibliotek, romanisering, unicode som teckenuppsättning samt standardisering. Efter den teoretiska bakgrunden följer den empiriska studien där våra informanter presenteras tillsammans med de bibliotek och bibliotekssystem som är aktuella för den här studien samt en redovisning av de resultat som kommit av våra intervjuer. I delen som följer, analysen, sammanfogas den teoretiska bakgrunden med resultatet av intervjuerna för att på så sätt ge ett djup åt de resultat som framkommit. I en avslutande del redovisas våra slutsatser baserade på analysen samt en kortare slutdiskussion för att sätta uppsatsen i ett större sammanhang.

## Definitioner

Vi använder oss av en del termer och uttryck i uppsatsen som vi anser att vi behöver definiera för att undvika eventuella missuppfattningar. Vi har även sammanställt en ordlista för några av de termer som förekommer i uppsatsen, se bilaga 1.

*Romanisering* – vi använder termen romanisering som ett samlingsbegrepp för de två termerna transkribering och translitterering, då vi talar om språk där båda formerna är aktuella.

*Lokala systemet/katalogen* – när vi talar om det lokala systemet/katalogen avser vi det bibliotekssystem som används av det aktuella biblioteken i vår undersökning, nämligen Millennium (Malin) på Malmö stadsbibliotek och Rosengårds stadsdelsbibliotek samt Virtua (Lovisa) på biblioteken inom Lunds universitetsbibliotek. Vi talar om det lokala systemet/katalogen tillskillnad från den nationella, där vi avser Voyager (Libris).

*Icke-latinska skrivsystem* – detta begrepp innefattar alla de alfabet och skrivsystem som inte använder sig av det latinska alfabetet med västeuropeiska modifieringar, som exempelvis bokstäver med diakriter<sup>3</sup>.

---

<sup>3</sup> Tecken som skiljer snarlika bokstäver åt, exempelvis accenter (ó), tilde (õ) och cirkumflex (ô).

*Originalspråk* – originalspråk avser det språk som materialet i fråga är skrivet på. När vi talar om att använda originalspråk i denna uppsats menar vi även det skrivsystem som språket normalt använder sig av, alltså ej i romaniserad form.

*Standardisering* – med standardisering avser vi en internationellt accepterad samling riktlinjer och regler att utgå ifrån för att på så sätt försäkra sig om att så många aktörer som möjligt arbetar utifrån samma förutsättningar. Vi avser först och främst standardisering inom datorvärlden, som exempelvis MARC21 och Unicode.

## Metod

### Urval

Vi har valt att genomföra vår undersökning på fyra bibliotek där problematiken är högaktuell. Dessa är Malmö stadsbibliotek och Rosengårds stadsdelsbibliotek med sina invandrartäta användargrupper samt Lunds universitets bibliotek, med särskilt fokus på Asienbiblioteket och Språk- och litteraturcentrums bibliotek, där studier och forskning inom olika språk förekommer. Vid Lunds universitet studerar även en stor mängd utbytesstudenter med modersmål på språk med icke-latinska skrivsystem.

Med vår undersökning önskar vi kartlägga de tankar och funderingar som finns bland bibliotekspersonalen på de aktuella biblioteken kring den problematik det innebär att hantera originalskrift på bibliotek.

Inom ramarna för denna uppsats bestämde vi oss för att intervjua sex personer totalt, tre från Lund och tre från Malmö. Ett viktigt kriterium vid valet av informanter var att de hade aktuell erfarenhet av katalogisering samt hantering av material med icke-latinsk skrift. Vi försökte finna informanter som innehar liknande ansvarsområden på båda biblioteken trots skillnaderna i hur bibliotekens organisation är uppbyggd. Några av informanterna visste vi tidigt var relevanta för vår undersökning och genom dessa blev vi rekommenderade och fick kontakt med andra lämpliga informanter.

### **Biblioteken i undersökningen**

Vid Lunds universitet efterfrågas material på många olika språk av lärare och forskare så väl som av studenter. Framför allt vid två nätverksbibliotek är hanteringen av denna typ av material stor, nämligen vid språk- och litteraturcentrums bibliotek samt Asienbiblioteket, och det är dessa två bibliotek som vi valt ut för vår studie. Lunds universitets bibliotekskatalog Lovisa har redan fullt stöd för unicode och därmed möjlighet att använda icke-latinsk skrift. Vid Asienbiblioteket har man experimenterat med att använda sig av originaltecken i katalogen när det gäller kinesiska och japanska. Det finns alltså ett intresse för att lägga in icke-latinsk skrift i bibliotekskatalogen och de möjligheter som en unicodeanpassning även av Libris kan medföra. Detta gör Lunds universitets bibliotek och då i synnerhet de bibliotek i nätverket som är kopplade till institutioner med språk med icke-latinsk skrift till mycket lämpliga föremål för vår undersökning då hanteringen av material med icke-latinsk skrift och de problem och möjligheter som finns kring detta är väldigt aktuella frågor på dessa bibliotek.

Även på Malmö stadsbibliotek och Malmös olika stadsdelsbibliotek är hantering av icke-latinsk skrift en aktuell problematik. Malmö är en stad där en stor del av

befolkningen har ett annat modersmål än svenska och många av de språken det rör sig om är språk med icke-latinsk skrift, som t ex arabiska, persiska och kinesiska. Till skillnad från många andra folkbibliotek satsar också Malmö stadsbibliotek på att köpa in litteratur på främmande språk från olika agenter för att kunna erbjuda ett bredare utbud. I Malmö stadsbiblioteks biblioteksplan för 2005 står t ex ”Ett kvalitativt, aktuellt och allsidigt medieurval garanterar medborgarnas möjlighet till en säker utveckling samt en fri och jämlik tillgång till information och upplevelser” (Malmö stadsbiblioteks biblioteksplan 2005, s. 2).

På Malmö stadsbibliotek har man arbetat mycket med att nå ut till grupper som har ett annat modersmål eller inte kan svenska bland annat genom att anställa gymnasieungdomar med olika modersmål som biblioteksguider som visar och berättar om biblioteket på de mest förekommande språken (Informant 5). Att biblioteken i Malmö kommun skall vara tillgängliga och öppna för alla oavsett bakgrund och ge jämlik tillgång till information och upplevelser är ett av bibliotekets mål med verksamheten. ”Faktorer som boende, ålder, arbetstider, funktionshinder, språk och ekonomi ska inte utgöra något hinder” (Malmö stadsbiblioteks biblioteksplan 2005, s. 3).

Malmö stadsbiblioteks bibliotekssystem är Millennium, som i dagsläget inte är unicodeanpassat. Det finns dock Millenniumsystem på andra bibliotek i Sverige som är unicodeanpassade. Man arbetar vid undersökningstillfället på Malmö stadsbibliotek med ett projekt där katalogisatörer lägger in bilder på tecken i t ex kyrillisk skrift i katalogen, men man hade ännu inte kommit så långt med detta projekt (Informant 6). I biblioteksplanen står även att ”Biblioteken skall fortsatt ligga i frontlinjen med hänsyn till utveckling och deltagande i virtuella/digitala tjänster både regionalt, nationellt och internationellt.” (Malmö stadsbiblioteks biblioteksplan 2005, s. 5) På Malmö stadsbibliotek katalogiserar man i det lokala systemet men importerar ofta poster från bland annat Libris.

Hanteringen av material med icke-latinsk skrift är alltså en högst aktuell problematik även på Malmö stadsbibliotek och stadsdelsbiblioteken om än med andra förutsättningar än Lunds universitets bibliotek vad gäller vilka grupper av användare man vänder sig till och vilka tekniska förutsättningar man har. Vi tycker därför att det kunde vara fruktbart att undersöka två olika typer av bibliotek med samma grundläggande problematik vad gäller material med icke-latinsk skrift men med olika perspektiv och förutsättningar.

## Metod

Den metod som vi ansåg vara mest lämplig för vår undersökning var kvalitativa intervjuer, eftersom vi ville undersöka attityder och åsikter hos våra informanter. Då vi skulle intervjua personer med olika arbetsuppgifter och olika roller inom de olika biblioteken bestämde vi oss för att en halvstrukturerad intervju, där frågeställningen belyses med i förväg bestämda frågeområden och följdfrågor och med en kombination av öppna och bundna svar, var den mest lämpliga intervjuformen (Lantz 1993, s. 68). Detta gav oss flexibilitet och möjlighet att följa upp olika teman som kom upp under intervjuerna. Samspelet mellan informant och intervjuare möjliggör en nyansering av

svaren och informanten får utrymme att verbalisera sitt tänkande och söka nyanser och differentieringar med en intresserad och ickevärderande person.

Det är viktigt att tänka på att intervjuaren spelar en stor roll i denna process, då han eller hon kan tänkas väcka frågor som informanten själv inte tidigare reflekterat över. Ickevärderande ska inte jämföras med "objektiv" utan bör betraktas som intervjuarens strävan att inte värdera det informanten säger utifrån sig själv, trots att det i praktiken inte är möjligt för en intervjuare att vara helt ickevärderande. Genom tolkning av informanternas svar ökar förståelsen av det undersökta fenomenet, i vårt fall hantering och inställning till icke-latinsk skrift i bibliotekssammanhang. Fokus ligger på den subjektiva erfarenheten, undersökaren närmar sig fenomenet utifrån individen och ur dennes perspektiv. Målet är att försöka fånga hur respondenten gestaltar fenomenet och tillskriver det mening. Meningen är alltid sammanhangsbestämd och det är individens person och erfarenheter som utgör den kontext mot vilken fenomenet får mening och kan förstås (Lantz 1993, s. 34 f).

Eftersom det ämne vi valt att undersöka är brett och berör både tekniska frågor, personalfrågor samt mediehantering men framförallt attityder och inställningar valde vi att ta ett brett grepp över många olika områden och föra samman det till en ny förståelse med hjälp av fördjupad litteraturstudie och en empirisk undersökning i form av kvalitativa intervjuer.

### Genomförande

Efter att ha inhämtat relevant kunskap genom en litteraturstudie kontaktade vi de utvalda informanterna via e-post och ringde upp dem ett par dagar senare för att boka intervjuer. Intervjutiden beräknades till cirka en timme per informant och intervjuerna genomfördes på informantens arbetsplats, dels för att balansera ut maktfördelningen i intervjusituationen för att informanten skulle känna sig bekväm, dels av rent praktiska skäl. I en intervjusituation finns det alltid en maktasymmetri där intervjuaren definierar situationen, introducerar vilka ämnen man skall tala om och styr genom följdfrågor hur intervjun fortlöper (Kvale 1997, s. 126). Utöver dessa sex intervjuer ställdes även en del frågor till den ansvarige för bibliografisk kontroll på Libris via e-post.

Alla intervjuerna spelades in på kassetband och transkriberades sedan. Intervjuguiden, som finns bifogad som bilaga 2, följdes i stort, men med vissa undantag. Följdfrågorna varierade något beroende på informantens svar samt skillnader i informanternas kunskap och yrkesroll. Under varje intervju var båda författarna med, men en tog rollen som huvudintervjuare medan den andra antecknade och kom med följdfrågor. Rollen som huvudintervjuare roterade mellan författarna.

Eftersom en av författarna arbetat med vissa av informanterna ansåg vi det viktigt att hon inte intervjuade dessa personer eftersom det antagligen skulle ha påverkat informanterna och hur de hade svarat på frågorna.

### Etiska aspekter

Då informanternas yrkesroll och arbetsuppgifter var viktig bakgrundsinformation för undersökningen vad det nödvändigt att redovisa dessa uppgifter. Detta förhindrade total

anonymitet för informanterna, vilket de informerades om vid intervjutillfället. Alla informanter informerades också om att de givetvis hade rätt att när som helst avbryta intervjun och sin medverkan i undersökningen, vilket ingen valde att göra.

Som tidigare nämnt så har en av författarna arbetat tillsammans med några av informanterna. Detta kan vara en fördel, eftersom det kan bli en familjär och avslappnad stämning mellan informant och intervjuare, men det finns också en risk för att det uppstår ett tyst samtal mellan informanten och intervjuaren, vilket kan leda till att viktig information inte tas upp. Detta löstes genom att den aktuella författaren aldrig var huvudintervjuare vid intervjun med dessa informanter.

Att båda författarna var närvarande vid alla intervjutillfällen gav en möjlighet till insyn i arbetet med intervjuerna från intervjutillfället till färdig analys.

## Flerspråkighet på bibliotek

För att kunna sätta vår egen undersökning i ett sammanhang är det viktigt att veta något om hur flerspråkig och mångkulturell biblioteksverksamhet ser ut idag samt har sett ut ur ett historiskt perspektiv. Vi tar i detta kapitel upp hur bibliotek agerat för att tillgodose den nya befolkningsstruktur som finns i Sverige idag. Kapitlet innehåller även ett avsnitt om den betydelse språket spelar i informationssökningsprocessen.

### Demografi och statistik

Den faktor som mer än något annat påverkar behovet av utländsk litteratur på framförallt folkbiblioteken är vilka etniska språkgrupper som finns företrädda i landet. För vår undersökning är det mest relevant att se på hur det ser ut i Malmö kommun samt vid Lunds universitet vad gäller andelen människor med annan etnisk bakgrund och utbytesstudenter.

I Malmö var i början av 2006 ca 26% av befolkningen född i utlandet, 8% var födda i Sverige med båda föräldrarna födda i utlandet. 169 nationaliteter fanns representerade men de största grupperna kommer från Jugoslavien, Irak, Polen och Bosnien Hercegovina (Malmö stad -02, 2005). Sedan 1990 har Malmös andel födda i utlandet ökat från 16 % till 25 % medan Stockholms andel har ökat från 15% till 20%. Av Malmös utlandsfödda kommer 62% från Europa, att jämföra med Stockholms 49%. Malmö kommer på tredje plats i Sverige när det gäller kommunernas andel födda i utlandet. Av storstäderna Stockholm, Göteborg och Malmö har Malmö högst andel invandrare. 28% av Malmös utlandsfödda kommer från Asien. Av barn mellan 0 och 15 år har 59% svensk bakgrund och 41% utländsk bakgrund (Malmöbor 2005).

Vid Lunds universitet studerade läsåret 2005/2006 1338 utbytesstudenter, varav 83 från länder med ett språk med icke-latinsk skrift. Dessa kommer bland annat från, Kina, Japan, Singapore, Grekland och Ryssland<sup>4</sup>. Säkerligen finns även ett visst mörkertal eftersom en del studenter som av olika skäl inte räknas som utbytesstudenter eller kommer från ett land med ett språk med latinsk skrift, men i själva verket har en icke-latinsk språkbakgrund. Till exempel kan utbytesstudenter från USA ha asiatisk etnisk bakgrund.

### Mångkulturell och flerspråkig biblioteksverksamhet

Antalet medborgare och invandrare i Sverige med rötter i länder utanför Europa har ökat sedan sextiotalet. Det har blivit allt viktigare att bemöta dessa förändringar i demografin

---

<sup>4</sup> Mejl från programhandläggare på Internationella sekretariatet på Lunds universitet.

inom alla samhällsinstitutioner, och då inte minst i biblioteken. 1983 gavs en handbok för invandrarverksamheten ut i de nordiska länderna som fick titeln ”KITAB - kulturell identitet tryggad av biblioteken”. Boken skrevs utifrån författarnas egna erfarenheter av biblioteksarbete med invandrare under sextio- och sjuttitalen. Dåtidens invandring bestod främst av människor som invandrat som arbetskraft med relativt låg utbildning från hemlandet. Biblioteken råddes att köpa in olika genrer inom skönlitteraturen, så många fackområden som möjligt, barnböcker, ljudböcker och musik, efter bästa förmåga eftersom språkkompetensen inte alltid fanns. KITAB värnade om människors rätt att behålla sitt eget språk och kultur i det nya landet och särskilt religiösa urkunder ansågs viktigt. KITABs syn på bibliotekens invandrarverksamhet följde den grundsyn som gällde för de nordiska folkbiblioteken. I biblioteksverksamhet liksom i all annan offentligt finansierad verksamhet utgick man ifrån att alla människors behov var lika. Målsättningen var en rättvis, jämlik service. Biblioteken byggdes upp efter en mall där det uppstod väldiga likheter mellan de olika biblioteken och urvalet av samlingar byggde på samma principer. På 1980-talet i Sverige var inte användarnas efterfrågan på viss litteratur något man lade någon större vikt vid (Brunnström 2005, s. 11).

I Sverige är denna problematik med medier på främmande språk för invandrare relativt ny till skillnad från USA, ett land där en stor del av befolkningen är invandrare från flera olika kulturer och länder. Redan i slutet av 1800-talet talade bibliotekarier i USA om bibliotekets roll som utbildare och skapare av upplysta, produktiva och självständiga medborgare (Du Mont 1994, s. 25).

The free library is the most promising of all measures for social amelioration because more than any other, it teaches and leads to self help...[reading library books causes] increased productiveness of our mechanics and artisans, in the lessening of crime and disorder among us, in the influx of the most desirable class of citizens, in the greater sobriety, industry, morality and refinement through-out the community that must necessarily result.

*Crunden 1890, s. 79 citerad i Du Mont 1994, s. 24*

För att nå ut till de grupper som skulle kunna dra nytta av bibliotekets service utvecklade man filialer. Dessa sågs som en bekväm metod för att ge service till speciella grupper som till exempel industriarbetare och icke-engelsktalande grupper som inte utnyttjade de centraliserade biblioteken. Trots vissa problem med finansiering och brist på utbildade bibliotekarier tillgängliggjorde filialerna för första gången böcker för många arbetande människor och gjorde det möjligt för grupper som tidigare ignorerats vad gällde biblioteksservice att använda sig av den service och tillgångar filialerna hade (Du Mont 1994, s. 25).

Huvudmålet med verksamheten var assimilation av de olika invandrargrupperna in i den amerikanska majoritetskulturen. Bibliotekarierna skapade egna informella metoder för att utbilda och socialisera, de samarbetade med dag- och kvällsskolor genom att tillhandahålla böcker som rekommenderades av lärare och hjälpte studenter genom att arrangera lektioner i engelska i biblioteket och att tillgängliggöra böcker på främmande språk och i översättning till engelska när det fanns tillgängligt. De höll medborgarskapskurser i biblioteket och tog fram broschyrer skrivna av bibliotekspersonal på främmande språk som beskrev regler och lagar och annan samhällsinformation (Du Mont 1994, s. 27). Att ge speciell service till



minoritetsgrupper och att tillhandahålla litteratur på främmande språk är alltså inget nytt i biblioteksvärlden. Dock har det skett en betydande skiftning i hur bibliotekarier ser på denna service under 1990-talet i USA:

Rather than responding to diversity through individual programming initiatives, libraries are beginning to view diversity more holistically. The goal is becoming how to improve basic library service for targeted groups. [---] The few initiatives begun in California in the last few years reflect a new emphasis on diversity as part of the basic fabric of public library service. This approach may become a model for libraries in the future. Changing demographics mean that all kinds of libraries need to consider an approach that views service to culturally diverse groups as one of the basic directions and goals they must strive to meet. The reality is that the United States is rapidly becoming ethnically diverse as a nation. Libraries have the choice of addressing this inevitable diversity or maintaining the "status quo" that will eventually lead them to serve only a relatively limited portion of the populace in their communities.

*Du Mont 1994, s. 32-33*

Även i Sverige har denna holistiska syn som Du Mont talar om blivit allt vanligare på bibliotek, framförallt på folkbibliotek i områden där en stor del av befolkningen har en annan etnisk bakgrund än svensk. Att arbeta mot olika invandrargrupper har blivit en allt viktigare uppgift för dessa bibliotek.

#### *Att möta flerspråkiga användares behov*

I undersökningen "Mer nytta än nöje - en undersökning av invandrade göteborgares syn på biblioteksverksamheten", som bygger på intervjuer med olika fokusgrupper av invandrare, efterfrågar deltagarna i undersökningen skyltar, broschyrer och litteraturlistor på olika språk i biblioteken. Förteckningar över vilket material som finns på de olika språken är ett önskemål som återkommer flera gånger. Enligt Brunnström finns det bland invandrare liksom bland många svenskar en bristande kännedom om vad biblioteken kan erbjuda, och flera av deltagarna i undersökningen säger sig aldrig ha vänt sig till informationsdisken, trots att de besökt biblioteket i flera år.

Kontakterna mellan personal och användare verkar fungera bättre på de mindre biblioteken. Personalen tycker sig känna användarna bättre och stämningen är mer avslappnad. Det är också lättare att hitta någon i närheten som kan fungera som tolk. Just språkliga problem verkar vara ett hinder för att de aktuella grupperna ska vända sig till informationsdisken, de vänder sig i så fall hellre till personal som är ute i biblioteket (Brunnström 2005, s. 13).

Vad gäller bibliotekskatalogen säger sig de flesta av deltagarna i undersökningen inte använda den. Att bibliotekskatalogen skulle ha möjlighet att använda både latinsk och icke-latinsk skrift var det främst de nyast anlända till Sverige som efterfrågade. Det faktum att en bibliotekskatalog med det egna språkets skrift inte efterfrågas mer kan bero på att så få av de i undersökningen använder bibliotekskatalogen överhuvudtaget. Det bekräftas från bibliotekspersonalen att de som söker litteratur på främmande språk oftast tittar i hyllorna istället för att använda sig av katalogen, vilket resulterar i att bibliotekarierna inte heller får möjlighet att visa hur man söker i katalogen. Försök på stadsbiblioteket i Göteborg tyder dock på att en flerspråkig katalog skulle vara en bra

ide, framförallt om det finns personal som behärskar de aktuella språken som kan vägleda användarna (Brunnström 2005, s. 14).

Då undersökningar eller forskning om utbytesstudenter vid högskolebibliotek är begränsade i Sverige har vi i första hand vänt oss till den forskning som gjorts på detta område i USA för att underbygga vår förståelse. Det största intresset för forskningen har dock varit kulturella faktorer och inte främst språk, även om språket anses vara en betydande barriär för utbytesstudenters förmåga att söka information (Jeong 2004, s. 384).

Whether on or off campus, Korean graduate students have little interaction with Americans. Instead, they tend to socialize only with other Korean students. A fundamental reason for this exclusiveness is the language barrier. Because of their weak conversational skills, many Korean graduate students feel insecure speaking English with Americans. Consequently, they avoid situations in which they must speak English. [...] intense fear of communicating in English may actually prevent some Korean students from asking questions and receiving valuable information [...].

*Jeong 2004, s. 391*

I boken *Multiculturalism in libraries* från 1994 presenteras en undersökning genomförd av Richard Hyman där katalogisatörer fick uttala sig om lokala anpassningar i katalogen. Trots att språket visat sig spela en viktig roll för högskolestuderandes informationssökning fann Hyman vid intervjuer med katalogisatörerna på biblioteken vid City University of New York att ingen av dem följde någon generell policy för att anpassa informationen i katalogen efter olika användargrupper. Hyman frågade katalogisatörerna huruvida de såg ett behov av att göra lokala anpassningar för att underlätta för vissa grupper, angavs brist på personal som det främsta skälet till att detta inte gjordes. Det ansågs mer kostnadseffektivt och mindre tidskrävande att bara importera poster utan att göra några lokala anpassningar (Du Mont 1994, s.113).

#### *Bibliotekspersonalen och dess roll*

I boken *Diversity and multiculturalism in libraries* av Mandernak, Lin & Hovde kan man läsa följande om bibliotekspersonalens roll i en mångkulturell verksamhet:

That staff is a key source in the provision of library service goes without saying. Whether at the Reference Desk or in a classroom setting, the interaction and communication between the staff and the patron often determines the success of that individual's comprehension and understanding of the information and concepts being presented. Effectively addressing the needs of a diverse user population introduces additional concerns, both in the hiring process and in training. [...] A diverse staff, by its very nature, is much more likely to recognize differences in sensitivities, thus providing a 'teaching laboratory' in the everyday work of the unit. Moreover, it also conveys to the users a greater invitation to utilize the resources of the library.

*Mandernak, Lin & Hovde 1994, s.91*

Att det finns invandrare bland personalen uppfattas som mycket viktigt av de flesta grupperna i undersökningen från Göteborgs stadsbibliotek. Detta är inte i första hand en språkfråga, även om de flesta skulle uppskatta hjälp från en landsman eller någon som talar samma språk. Som skäl anges främst att man känner sig tryggare med människor

med liknande bakgrund. Personal med utländsk bakgrund ses också som en symbol för möjligheten att lyckas få ett arbete som kräver högre utbildning. Både de som deltog i undersökningen och de ur personalen som intervjuades i Göteborg ansåg att anställda med invandrarbakgrund var bättre på att se behoven hos olika grupper och många invandrare uppgav att de hellre vände sig till en i personalen som inte såg ut som majoritetsbefolkningen även om de inte hade samma etniska bakgrund. Språket spelar dock också en viktig roll då det uppfattades som oerhört värdefullt att träffa på någon som pratar ens eget språk. Flera personer ansåg också att det var lättare att förstå en person som pratar bruten svenska än en infödd svensk som talar utan brytning (Brunnström 2005, s. 15).

På BIBLIST, en e-postlista för svenska bibliotekarier, har det under våren förts en kortare diskussion om betydelsen av att rekrytera bibliotekarier med invandrarbakgrund och man efterlyser här ett uppfångande av invandrade bibliotekarier genom att antingen anordna en speciell utbildning eller genom praktikplatser på bibliotek för att på så sätt ge de invandrade bibliotekarierna de extra kunskaper som behövs för att arbeta på ett bibliotek i Sverige.<sup>5</sup>

Åsikten att bibliotekarier med minoritetsbakgrund kan ge bättre service till andra ur samma minoritetsgrupp och även andra etniska grupper är inte unik för Sverige. Även i USA har man länge uppmärksammat detta. Amerikansk forskning visar att en bibliotekarie med samma etniska bakgrund som majoriteten av låntagarna i närområdet bäst kan uppmuntra till flitigt användning av en filial (Du Mont 1994, s. 90). Trots detta betjänas få låntagare av personal med samma kulturella bakgrund. Det är vanligt att personal som har minoritetsbakgrund inte är utbildade bibliotekarier utan assistenter eller frivilliga. Dessa utför dock ofta arbetsuppgifter som i normala fall skulle utföras av bibliotekarier.

It is as paraprofessionals in libraries that minority persons have tended to be employed. There is evidence that indicates that many of the minority librarians currently employed in the profession came into the profession because of early nonprofessional work experiences.

*Du Mont 1994, s. 90*

Efterhand som minoritetspopulationen växer blir effekterna av detta allt tydligare. Du Mont menar att minoriteterna inte får biblioteksservice likvärdig den som majoriteten får:

As public libraries decide to enhance services and collections in response to ethnic and racial diversity, they are in need of staff members who can select and process diverse materials. This need suggests a profile of a librarian who is in many cases bilingual and who has knowledge of materials reflecting diversity in various fields. Libraries will have to compete to hire and retain workers who have special language skills and/or the knowledge of how to acquire scarce but needed resources related to minority concerns. As racioethnic minorities increase in proportional representation in the labour pool, it is to this group to which libraries can turn for this kind of knowledge and skills.

*Du Mont 1994, s. 43*

---

<sup>5</sup> Tre inlägg på BIBLIST har under våren 2006 diskuterat frågan.

## Språkets betydelse

Kulturella skillnader och etnisk bakgrund, hur användaren ser på bibliotek, hur språk är konstruerade, hur användaren tänker och ser på omvärlden är faktorer som kan ha en stor inverkan på framgång i informationssökning (Zoe & DiMartino 2000, s. 295). Zoe och DiMartino fann i en undersökning av studenters informationssökning ett samband mellan modersmål och framgång med sökningar på engelska i en fulltextdatabas. De studenter som hade engelska som modersmål fick mest positiva resultat: 71% uppgav att de funnit den information de sökte medan endast 51% av de som hade ett östasiatiskt språk som modersmål ansåg sig ha uppnått ett positivt resultat av sin informationssökning. 67% av de som talade ett östasiatiskt språk var missnöjda med sökningen, mot endast 38% av de engelsktalande.

Zoe och Dimartino menar att språket är en viktig faktor för hur väl studenterna lyckas med sökningen, inte bara på grund av ordförståelse utan även på grund av förståelse av språkets konstruktion och grammatik:

The construction of English and Asian languages are extremely different. For example, in Chinese, there are no function words like prepositions, conjunctions, or articles. [...] Further, the physical layout of documents, and the ways in which they are read and segmented, may also cause difficulties for students used to associating, reading and conceptualizing printed letters or characters in different orders. The construction of the original language has underlying implications in the development of more appropriate and accurate instructional methods.

*Zoe & DiMartino 2000, s. 302*

### *Engelska som norm i datorvärlden*

Engelska har varit och är fortfarande till stor del det dominerande språket när det gäller datorer och Internet. Datasystem och datorer är anpassade för engelsk skrift i första hand, och det uppstår ofta problem när man försöker introducera icke-latinsk skrift i datorsystem (se kapitlet om unicode). Detta är inte längre en hållbar situation när exempelvis de asiatiska länderna får allt större tillgång till datorer och människor med asiatiska språk som modersmål använder sig av Internet. Arsky och Cherny menar i sin artikel "The Ethno-cultural, linguistic and ethical problems of the 'infosphere'" att interaktionen mellan olika språk i det expanderande internationella utbytet av information är ett allvarligt problem som inte har någon lätt lösning (Arsky & Cherny 1997, s. 256). De ifrågasätter den dominans som engelska språket har i informationssamhället idag;

If we wish to extend co-operation among the different peoples of the world on the basis of information exchanges, we must do more than create the global information infrastructure and highway which form the technological basis for the information society of the 21<sup>st</sup> century. An effort must also be made to resolve sensitive issues involved in mutual understanding between peoples belonging to different cultures and speaking different languages. This will not come about if the tendency for some languages and cultures to dominate others is allowed to persist.

*Arsky & Cherny 1997, s. 251*

Några av lösningarna som föreslås av Arsky och Cherny är utveckling av datorsystem för översättning och samarbete kring inkludandet av nationellt material i

internationella databaser. De påtalar också vikten av att kunna söka i databaser och bibliotekskataloger på flera språk (Arsky & Cherny 1997, s. 258).

## Romanisering

En viktig del av denna uppsats är att påvisa den problematik som kommer med att transkribera eller translitterera skrift, en problematik som informanterna i vår undersökning visade sig var väl medvetna om. I det här kapitlet går vi igenom denna företeelse ur ett historiskt perspektiv samt redogör för den kritik som framförts genom bland annat Hans H. Wellisch.

### Transkribering och translitterering som metod

För att kunna återge olika språks skrivsystem med andra tecken än de ursprungliga krävs att man använder sig av translitterering eller transkribering, beroende av vilken typ av skrivsystem det aktuella språket använder sig av. Om avsikten är att representera exempelvis arabiska med latinska bokstäver handlar det om translitterering, det vill säga ett språks bokstäver ersätts av ett annat, tecken för tecken. Om det i stället är frågan om språk som inte har alfabet måste man utgå från ljud i stället för tecken, vilket innebär att man använder sig av transkribering. Detta är aktuellt då till exempel japanska eller kinesiska ska representeras med latinska bokstäver. För att kunna tala om både transkribering och translitterering används samlingsbegreppet romanisering, vilket innefattar de båda systemen och innebär ett främmande språk med icke-latinsk skrift skrivet med latinska bokstäver, oavsett metod.

### Historik

Även om romanisering används i olika former genom historien var det först under 1800-talet som man kan tala om mer organiserade standarder. Mot slutet av 1800-talet utvecklade och publicerade katalogisatorer på de stora biblioteken i Storbritannien, Amerika och Tyskland transkriptionssystem för sina kataloger. När sedan fler bibliotekarier i andra europeiska länder följde deras exempel, utgick de oftast från någon av föregångarna men med lokala avvikelser för att bättre korrespondera med det egna språket. "Inevitably, these changes were labeled 'minor' by the editors of codes, but they led to almost complete incompatibility between many of the codes and the catalog entries produced according to their rules" (Wellisch 1976, s. 246). Problemet med system som inte var kompatibla ledde till försök att utveckla internationella standardiserade transkriberingssystem, och även om internationellt accepterade system kunde utvecklas, var de långt ifrån de standarder som från början avsågs (Wellisch 1976, s. 246).

På grund av att mängden material med icke-latinsk skrift fortfarande var begränsad på västerländska bibliotek under första halvan av 1900-talet utgjorde bristen på enhetliga romaniseringssystem inte något större hinder. Detta förändrades dock då samlingarna

började växa och behovet av att kunna hantera detta material på ett smidigt sätt blev allt större på biblioteken (Wellisch 1976, s. 224).

Kring 1950-talet hade frågorna kring romanisering samt smidig hantering av materialet växt och kritik mot romanisering som metod uppstod på många håll. Författare, forskare och bibliotekarier började ifrågasätta allt från felaktigheter i specifika transkriptionsscheman till svårigheterna för biblioteksanvändarna att återfinna materialet utifrån dess romaniserade form (Wellisch 1976, s. 226).

### Kritik mot romanisering

Inom biblioteksbranschen förekommer ifrågasättande av romanisering sällan utanför specialistkretsar enligt Aliprand: "[...] many nonspecialist librarians are unaware of the deficiencies and still regard romanization as adequate for access. Language experts reject this view [...]" (Aliprand 1992, s. 105). Det är först när man börjar titta närmare på romanisering av ett språk som de brister som finns tydliggörs, varför det är lätt för de som inte är insatta i problematiken att anta att romanisering är en fullgod metod att hantera språk med icke-latinska skrivsystem.

Det finns dock de som mycket starkt kritiserar denna metod. En av dessa kritiker är Hans H. Wellisch, som 1978 skrev boken *The Conversion of Scripts – its Nature, History, and Utilization*. I boken beskriver han mycket ingående historiken och praxisen bakom romanisering och ifrågasätter starkt metoden i bibliografiska sammanhang. Han menar att romaniseringen i bibliografiska sammanhang snarare förhindrar än förenklar sökningen av litteratur med icke-latinska skrivsystem, då romaniserade dokument oftast är i det närmaste obegripliga för de med det aktuella språket som modersmål (Wellisch 1978, s. 230). Detta är något som även Joan Aliprand kritiserar i sin artikel "Nonroman scripts in the bibliographic environment": "A reader faced with text rendered in an unfamiliar way find it incomprehensible" (Aliprand 1992, s. 106).

Det finns idag två förhärskande system för transkribering av kinesiska, nämligen Wade-Giles och pinyin. Wade-Giles är traditionellt det mest använda systemet i Nordamerika medan Europa och Australien i huvudsak använder pinyin. 1997 beslutade Library of Congress att övergå till pinyin, då denna form kommit att bli den mest accepterade av de ledande institutionerna i världen. Den kinesiska staten har erkänt pinyin som den officiella transkriptionsformen och även Förenta Nationerna och ISO utgår från pinyin vid transkribering av kinesiska (Arsenault 2002, s. 45). Trots att beslutet att byta transkriptionssystem kom redan 1997 skedde övergången till pinyin först under 2001 (Library of Congress pinyin conversion project: coordinated timeline).

Arsenaults artikel "Pinyin romanization for OPAC retrieval – is everyone being served?" redovisar resultatet i en undersökning där Arsenault avser att utreda skillnader i informationsåtervinning beroende på vilka transkriptionssystem som använts. Deltagarna i Arsenaults undersökning delades in i tre grupper, Wade-Gilesbaserad samt två olika former av pinyin (med och utan mellanrum mellan orden), och ombads söka efter en rad titlar skriva med kinesisk skrift. Undersökningen visade att romaniseringsfel står för mellan 33 och 50 % av missarna i sökningen, beroende på vilken romaniseringsgrupp som utförde sökningarna. Problemen med romaniseringen var

följande: ett tecken uttalades eller lästes fel, pinyin användes istället för Wade-Giles eller tvärt om, fel fonetisk transkribering (till exempel chen istället för zhen, fu istället för hu) samt övriga fel (wuan istället för wan) (Arsenault 2002, s. 48).

Arsenault avslutar sin artikel med att kommentera resultatet så här:

[...] a large number of search failures are directly attributed to romanization errors in the query strings input by the participants. The main conclusion to be drawn from this is that using romanization as the primary retrieval technique in OPACs still remains rather problematic for many end users. It would therefore be desirable to investigate the possibility of providing other means of retrieval, most probably character-based, as an alternative for retrieval of Chinese-language records in OPACs. [...] The adoption of the global UCS/Unicode character set in MARC 21 should facilitate and improve multiscript searching in Web-based bibliographic databases.

*Arsenault 2002, s. 49*

Vid transkribering måste hänsyn tas till de två språkens fonologi, det vill säga deras ljudbilder, för att på bästa sätt foga samman dessa. Vid jämförelse mellan svenska och japanska finns skillnaden att japanskans ljud delas in i stavelser i stället för enskilda bokstäver, fonem, som svenskan. I japanskan uttalas inte en konsonant utan vokal, vilket innebär att ett ord på svenska med flera konsonanter i följd inte kan få en motsvarighet på japanska. Det svenska ordet "klara" skulle skrivas "kurara" på japanska då k och l skulle uttalas ku och ra i detta ordet medan stavelsen ra redan finns i japanskan. Anledningen till att l skrivs r är att det svenska ljudet r inte har någon direkt motsvarighet i japanskan och det ljud som mest liknar l i japanskan skrivs r (Kockum 2001).

Att omvandla japanskans eller kinesiskans tecken till latiniserade ord innebär mer än bara ett byte av skrivsystem. På grund av den stora mängden homofoner<sup>6</sup> i dessa språk försvinner lätt betydelsen när ordet skrivs med bokstäver istället för tecken. Detta tar Joan Aliprand upp i sin artikel "Nonroman scripts in the bibliographic environment":

Chinese contains a considerable number of homophones, and the differences between characters with the same sound cannot be maintained in a romanization scheme based on the rendering of pronunciation.

*Aliprand 1992, s. 106*

Ett problem som gäller både transkribering och translitterering är den stora mängd konverteringsschema som finns. Beroende på vilket schema som används kan en arabisk författares namn skrivas på följande sätt:

Gibran, Kahlil  
Gibran, Khalil  
Gibrán, Jalil  
Gabran, Halil

---

<sup>6</sup> Homofon är ord som uttalas exakt likadant men skrivs med olika tecken, exempelvis ordet shō i japanskan som kan betyda liten, avdelning, handlare, kapitel, motstånd och pris, beroende på vilket tecken som används för att skriva ordet (i detta fallet ett av sex olika tecken). För att i romaniserad text förstå vilket av dessa som avses måste man se till sammanhanget i meningen i övrigt.



Jabran, Halil Jabran  
 Jubran, Jubran Khalil  
 Gebran, Gebran Khalil  
 Jibran, Jibran Khalil  
 Yubran, Yubran Jalil  
 Chi-po-lun  
 Yubran, Yubran Jalil  
 Gibran, Gibran Khalil  
 Gibran, Halil Gibràn  
 Gibran, Halil

(Exemplen hämtade från Risberg 2003, s. 15)

I detta exempel blir det mycket tydligt hur olika transkriberingsschema kan påverka hur ett namn skrivs, vilket innebär att det blir nästintill omöjligt att få en rimlig sökträff på denna författares namn i en bibliotekskatalog om personen som söker inte känner till vilket transkriberingsschema som använts. Det är givetvis möjligt att lägga in fält och hänvisningar för alternativa stavningar vid katalogisering, men då en del språk har ett så stort antal transkriptionssystem, som dessutom kan variera beroende på vilket språk katalogisatören själv talar, skulle det generera en orimlig mängd alternativa stavningar för att täcka in alla varianter.

I slutändan går det inte att komma ifrån att ett fält med originalskrift skulle erbjuda en mycket bättre möjlighet att få relevanta och korrekta sökträffar i katalogen. Risberg uttrycker sig så här:

Om det är svårt för bibliotekarien att söka den här litteraturen effektivt så är det ofta en omöjlighet för biblioteksbesökaren. Visst kan man lära sig systemet och det kan man kanske förvänta sig av studenter och forskare på forskningsbiblioteket, men man kan inte kräva att låntagarna på folkbiblioteket ska ha specialkunskaper i romaniseringsregler. Romaniseringen är ett hinder för dem och ingen hjälp.

*Risberg 2003, s. 17*

Wellisch citerar Sommer som ställer sig frågan för vem romaniseringen egentligen görs:

For whose benefit is the transliteration made? Is it primarily for the readers or for the staff? On consideration, there can be but one answer: it is for the staff, or, more generally, for those who are unable to read the original script....As to the foreign readers, they naturally prefer the original script and derive practically no benefit from the transliteration.

*Sommer 1934, se Wellisch 1976, s. 224*

Wellisch går så långt i sin kritik av romanisering på bibliotek att han menar att romaniseringen motsvarar icke läskunniga människors försök att markera böckerna med kryss, cirklar och färgade prickar för att sedan be andra människor att leta reda på boken de behöver utan att använda titlar eller författare (1976, s. 230). Även om detta uttalande kan tyckas vara att gå lite för långt i kritiken av romanisering är det tänkvärt. För de flesta människor skulle det troligtvis vara mycket svårt att känna igen sitt eget språk skrivit enligt ett annat språks regler och med skrivtecken som är anpassade efter en annan ljudbild. Faktum kvarstår dock att ingen människa kan lära sig alla språk och

skrivsystem i världen, varför det är orimligt att förvänta sig att bibliotekspersonal ska besitta sådana kunskaper. Som Risberg uttrycker det i sin uppsats *Originalskrift i flerspråkiga bibliotekskataloger*:

[...] romanisering har en enda fördel och det är att den *i bästa fall* ger tillgång till bibliografisk data för den som inte kan läsa språket i fråga i originalskrift. Det innebär att material på ett enkelt sätt kan integreras i ett och samma system och kan hanteras i t ex cirkulation och återvinning. Men frågan är om denna fördel uppväger de nackdelar som metoden innebär för dem som vill komma åt litteraturen?

*Risberg 2003, s. 19*

## Unicode som teckenuppsättning

Då en av avsikterna med denna uppsats är att utreda förekomsten med icke-latinsk skrift i bibliotekskatalogen ser vi ett behov av att förklara vad som krävs för att använda en större mängd skrivsystem i datormiljö vad gäller teckenuppsättning. Den idag absolut mest utbredda teckenuppsättningen för den typen av hantering är unicode, varför detta kapitel kommer att behandla utvecklingen före och efter unicode samt hur utvecklingen inom biblioteksvärlden ser ut vad gäller implementering av unicode, detta gör vi framförallt med hjälp av en omfattande undersökning gjord av Laura Tull. Vi kommer även att ta upp en del av den kritik som finns mot unicode för att på så sätt belysa fler aspekter.

### Utvecklingen före unicode

#### *Från engelska till all världens språk*

När de första datorerna utvecklades fanns enbart stöd för representation av ett språk, nämligen engelska, vilket föll sig naturligt då denna utveckling skedde i USA. Många år har dock passerat och datoriseringen har blivit global, med följden att en stor mängd användare av datorer idag inte har engelska som modersmål, många är inte ens engelsktalande. Detta har tvingat fram en utveckling där det nu är möjligt att använda sig av en mängd skriftsystem för representation av all världens språk i datormiljö. Redan 1994 skrev Glenn Adams så här i sin bok *Introduction to Unicode*:

Enter the global market – English may be the lingua franca of modern times, but there are many more desktops waiting for computers that speak other languages; consider for a moment that Indonesia, Thailand, and Vietnam together are more populous than the United States. Just think about market growth for a moment.

*Adams 1994, s. 3*

Adams menar att system anpassade efter en viss teckenuppsättning ofta är dyra och svåra att sköta, varför idealet borde vara system som är ”completely transparent to languages” (1994, s. 4). Detta kan man uppnå genom att öppna upp systemet för möjligheten att känna igen alla världens tecken. Ett sätt att göra det är genom unicode. Janet C. Erickson skriver 1997 så här i sin artikel ”Options for presentation of multilingual text: use of the unicode standard”:

Service to a growing international audience is imperative. Digital libraries are springing up and their communities span the globe. The world needs to come to an agreement on how the information they provide can be understood unambiguously by the computer systems of their newfound audience.

*Erickson 1997, s. 172*

Tankegångarna bakom dessa två citat är tydliga: det finns ett behov av att använda olika skrivsystem i datorbaserade miljöer, och metoden för detta bör vara standardiserad. Även in på 2000-talet fortsätter diskussionen genom bland andra Laura Tull som skriver följande i sin artikel "Library systems and Unicode: a review of the current state of development":

Unicode implementation has been gaining momentum in recent years especially in the software and computer industry. Academic libraries with collections of materials in multiple languages will want to take advantage of Unicode for display and searching of materials in non-Latin scripts such as Arabic, Hebrew and Chinese.

*Tull 2002, s. 181*

Det är tydligt att den utveckling som Adams förutspådde 1994 och Erickson 1997 fortfarande pågår år 2002. Tull menar dock att även om arbetet med unicode startade redan 1991 så är det först en bit in på 2000-talet som unicode börjat användas i större utsträckning (Tull 2002, s. 181). Library of Congress gjorde en fullständig unicodeanpassning först 2005, och många stora tillverkare av bibliotekssystem har först nyligen förstått vikten av att byta teckenuppsättning på sina produkter (se undersökningen gjord av Tull längre fram i detta kapitel).

### *ASCII och ISO*

I datormiljö lagras information i enheter som kallas bits. Bakom en vanlig bokstav ligger oftast en kod bestående av en byte, vilket är en kombination av åtta bits. Traditionellt sett består skrivtecken av enkla bytes. På grund av detta, kombinationen av två tal (etta eller nolla) i gruppering om åtta, ges 256 kombinationer att bygga tecken av (Risberg 2003, s. 20). För att representera det engelska alfabetet behövs 26 plus 26 för alla bokstäver i alfabetet, stora och små. För att lägga till siffror krävs ytterligare nio. Till detta kommer sedan alla vanliga symboler och tecken som punkt, utropstecken, frågetecken, parenteser, plus- och minustecken, asterisker, kolon, semikolon och så vidare, vilket innebär att de mest basala tecknen för att skriva engelska i en dator kräver lite över hundra tecken.

Det första teckensystemet för datorer var ASCII (American Standard Code for Information Interchange), som var anpassat för en engelskspråkig miljö. ASCII var från början uppbyggt på 7-bitar, vilket ger 128 tecken. Detta var fullt tillräckligt för språken baserade på det latinska alfabetet med de vanligaste numeriska tecknen och symbolerna, i alla fall om man inte försökte använda sig av mer än ett språk i systemet. Antalet tecken räcker dock inte till för att använda sig av exempelvis både engelska och franska i samma system (Adams 1994, s. 7). Eftersom detta system byggt på 7-bitar inte var tillräckligt flexibelt utvecklades en 8-bitars "extended ASCII", vilket gav möjligheten till 256 tecken i stället för föregångarens 128. "Extended ASCII" fungerade dock bara på PC (DOS), vilket ledde till att Apple introducerade ett egen system, MAC Roman, som bara fungerade på Macintosh. Även ISO, International Standard Organisation, utvecklade en teckenuppsättning, ISO 8859, som kunde användas av UNIX och Microsoft Windows (Risberg 2003, s. 20).

Medan ASCII är en amerikansk standard anpassad för de amerikanska förutsättningarna är ISO utformad för internationellt bruk. Detta innebär att ISO tvingats hantera kraven på flerspråkiga lösningar, vilket i sin tur lett till ett antal versioner med avsikt att täcka de europeiska språken (varav två med icke-latinska skrivsystem: grekiska och kyrilliska) samt arabiska. Dessa standarder fick benämningarna ISO 8859-1 till 8859-10 (Risberg 2003 s. 20).

#### *Tidiga bibliotekssystem med icke-latinsk skrift*

Behovet av att kunna lagra icke-latinsk skrift i datorer har funnits länge, inte minst i bibliotekssammanhang. De två stora amerikanska agenterna (OCLC) och Research Libraries Information Network (RLIN) började redan under 1980-talet utveckla system som kunde hantera icke-latinska skrivsystem (Risberg 2003, s. 26). Båda kom senare att bli aktiva medlemmar i Unicode Consortium (se nedan).

På grund av svårigheterna att utveckla teckenuppsättningar som fungerar för många språk samtidigt har en del andra lösningar uppstått. En är att använda sig av bilder av text, för att på så sätt komma runt teckenhantering och teckenstöd, då en bild för en dator alltid är en bild, oavsett vad den föreställer. Denna metod har dock en stor brist, nämligen att text i bildform inte är sökbar. Erickson tar upp detta i sin artikel ”Options for presentation of multilingual text: use of the unicode standard”:

On first glance, images seem the ideal solution; most if not all browsers understand their content and they are fairly simple to produce. However, these images are not searchable, and what is seen on the screen is what is copied to a local drive when the image is “saved”. A person saving an image-filled page receives only a reference to the image, not the image itself. [...] while a particular application may solve the problems of display in this way, images in and of themselves do not provide a satisfactory general solution.

*Erickson 1997, s. 175*

Att försöka kringgå ett fungerande teckensystem på detta sättet innebär alltså enbart en kortsiktig och halvhjärtad lösning på problemet med representation av icke-latinsk text. De första flerspråkiga bibliotekssystemen med teckenstöd för icke-latinsk skrift var det israeliska systemet ALEPH (Automated Library Expandable Program Hebrew University of Jerusalem) och det kanadensiska systemet MINISIS. ALEPH började experimentera med användandet av två skrivsystem redan 1981, och använde sig av ett chip i den aktuella terminalen, vilket möjliggjorde användandet av både hebreisk och latinsk skrift. 1996 gick man över till en unicodeuppsättning, och systemet används i dag i ett 40-tal länder. Parallellt med ALEPH utvecklades MINISIS av International Development Research Center i Kanada, som på 1970-talet experimenterade med olika skrivsystem baserade på först på 8-byte (256 tecken), exempelvis arabiska och kyrilliska, för att sedan gå vidare med double-byte (16-byte) för att kunna skriva kinesiska, japanska och koreanska. 1996 kunde man använda sig av tio språks skrivsystem utöver de latinska, nämligen: kinesiska, arabiska, grekiska, koreanska, thailändska, devanagari, kyrilliska, bengali, gujarati och punjabi (Risberg 2003, s. 26-27).

Trots stora framgångar med egenhändigt utvecklade system både i Europa, USA och Asien finns ett stort hinder när det gäller utbyte av information och möjligheten till

import och export av katalogposter då systemen ofta är uppbyggda oberoende av varandra och därför inte kompatibla.

## Unicode

### *Hur allt började*

Arbetet med unicode-standarden påbörjades redan 1983 då ISO inledde arbetet med den teckenuppsättningsstandarden som senare fick namnet ISO 10646. Parallellt med detta arbete utvecklade även Joseph Becker vid Xerox PARC samt Lee Collins och Mark Davis vid Apple Computer en helt ny teckenuppsättning som fick namnet unicode (Risberg 2003, s. 28).

Unicode Consortium bildades 1991 för att arbeta vidare med standarden. Medlemmar i konsortiet var framför allt företag och grupper inom datorbranschen, exempelvis IBM, Apple Computer, Adobe Systems och Microsoft, men även biblioteksvärlden var representerad genom Research Libraries Group (RLG), Online Computer Library Center (OCLC) och en rad andra företag som utvecklade bibliotekssystem (Tull 2002, s. 181). Kort efter att Unicode Consortium bildades gick man samman med ISO 10646 vilket ledde till att de två standarderna gjordes kompatibla och unicode accepterades som internationell standard 1992 (Risberg 2003, s. 28).

### *Unicode som teckenuppsättning*

I grund och botten kan datorer bara hantera siffror. All text som representeras på skärmen är uppbyggd av sifferkombinationer, och olika bokstäver har olika koder (Memon 2001, s. 229). På grund av att olika system har haft olika koder för text har det uppstått problem när utbyte av text ska ske och den ene datorn inte förstår den andres koder, dvs. inte visar rätt bokstäver eller tecken. Detta kan yttra sig i att fel tecken visas eller att inget tecken alls visas på skärmen.

Unicode är varken hårdvara eller mjukvara, utan en formell standard för flerspråkig kodning. Tanken bakom en teckenuppsättning som unicode är att, oberoende av plattform, program eller språk, så ska samma text visas likadant på två olika datorer. ”It [unicode] allows data to be transported through many different systems without corruption” (Chandrakar 2004, s. 422).

Målsättningen med unicode kan delas in i tre huvudteman: den ska vara *universell*, vilket innebär att den ska täcka alla skriftspråk, den ska vara *unik*, dvs. varje tecken kodas endast en gång även om samma tecken förekommer i flera olika språk, den ska vara *uniform*, dvs. alla tecken ska ha samma bit-längd (Risberg 2003, s. 28).

I den senaste versionen av unicode standarden, version 4.0, finns en kapacitet för 96,447 tecken (The Unicode Standard: a Technical Introduction, 2006). Detta inkluderar även matematiska, tekniska och geometriska former samt Braille (Risberg 2003, s. 29).

Från början baserades koderna på 16-bitar, vilket var tillräckligt för att koda alla moderna skrivsystem inom ramarna för den ursprungliga designprincipen, men utvecklingen medförde att en förlängning krävdes, och 32-bitars kod lades till

standarden (Chandrakar 2004, s. 422). Detta ledde alltså till att Unicode standarden kom att innehålla tre olika kodningsformat: 8-bitar (UTF-8), 16-bitar (UTF-16) och 32-bitar (UTF-32).

UTF-8 används mest inom HTML, det vill säga för webbsidor och liknande. Denna kodningen överensstämmer med teckenuppsättningen ASCII, och är den kodning som flest mjukvaror kan använda (The Unicode Standard: a Technical Introduction, 2006). De tecken som inte ryms inom 8-bitars kodningen kodas istället med hjälp av 16-bitar (Chandrakar 2004, s. 422). Tull ser en fördel med UTF-8 på grund av dess kompatibilitet med ASCII: "The big advantage of UTF-8 is that ASCII and its equivalent Unicode characters have the same value, making it more compatible with existing software" (Tull 2002, s. 181). Hon fortsätter:

Most Web browsers support UTF-8, and since online public access catalogs are displayed using Web browsers, it makes sense that library systems need to support that transformation format. However, as the software industry moves towards Unicode implementation around the world, support for more than just UTF-8 may be necessary in the future for easy exchange of records. For example, if you are trying to import records from a source that uses UTF-16 and your system supports UTF-8 only, then development will be required to make that exchange.

*Tull 2002, s. 183*

UTF-16 baseras på 2 byte till skillnad från UTF-8s 1 byte och kan ses som ett supplement till det mindre UTF-8 där inte alla tecken ryms (Chandrakar 2004, s. 422). UTF-16 används när ett större antal skrivsystem behövs, men datorutrymmet är begränsat, då det är ett mindre system än UTF-32 (The Unicode Standard: a Technical Introduction, 2006).

UTF-32 är det största av de tre kodningsformaten med 4-bytes koder och har kapacitet för miljontals tecken (Chandrakar 2004, s. 423). UTF-32 kan användas när datorutrymmet är obegränsat (The Unicode Standard: a Technical Introduction, 2006).

En av målsättningarna med unicode är att varje tecken ska ha unika koder vilket innebär att alla kombinationer av ett tecken får sina egna koder, exempelvis har bokstaven "a" kod nummer 0061 medan "á" har 00E1 och "à" har 00E0 (C1 Controls and Latin-1 Supplement, 2005). Detta gör att det inte kan bli någon konflikt inom systemet, eller som Adams skriver: "[...] Unicode uniquely associates codes with one and only one character. Consequently, given a particular encoding value, its identity as a Unicode character is never in question" (1994, s. 12).

#### *Skillnaden mellan textelement och kodelement*

I system anpassade för det engelska språket är skillnaden mellan textelement och kodelement oftast inte särskilt väldefinierad då de båda ses som direkt sammanlänkade. På grund av att olika skrivsystem har olika grundförutsättningar för basal textinformation blir definierandet av textelement och kodelement mer relevant för flerspråkiga system (Adams 1994, s. 59).

Vid representation och kodning av en bokstav ur det engelska alfabetet finns ett begränsat antal varianter på exempelvis bokstaven a som har en kod. Om man istället

vill representera en bokstav som inte finns i det engelska alfabetet, exempelvis bokstaven ä, krävs antingen att den har en kodning som består av koden för a samt koden för två prickar över a eller att den har en helt egen kodning oberoende av a. Unicode väljer att använda sig av den senare varianten.

Då de allra flesta systemen på marknaden idag utgår från det engelska alfabetet har inte behovet att närmare definiera skillnaden mellan ett textelement (en bokstav) och ett kodelement (koden bakom bokstaven) varit speciellt stort.

Multiple code elements may represent a composite formal unit, which, in turn, may represent, partially or wholly, a text element. It is up to the system, the application, or even the user to determine how to define text elements.

*Adams 1994, s. 59*

När man diskuterar teckensystem är det viktigt att hålla isär teckenuppsättning och typsnitt. Unicode är en teckenuppsättning, inte ett typsnitt, vilket innebär att unicode inte påverkar hur ett tecken eller en bokstav ser ut på skärmen, bara att det är rätt värde på tecknet eller bokstaven. Exempelvis är unicode-koden för värdet "a" 0061, men beroende på typsnittet kan detta värde visas som "a" eller "a" (Adams 1994). På grund av att unicode handlar om kodelementen och inte textelementen krävs alltså mer än bara ett unicodeanpassat system. Att ett system är anpassat för unicode betyder inte automatiskt att det har förmågan att visa alla teckensystem. För att visa ett tecken på en datorskärm krävs att det finns en font för detta, och vanligast är att systemet bara har fonter för ett fåtal skrivsystem (Risberg 2003, s. 33). Unicode ger basen för all världens teckensystem, men för de skrivsystem man vill kunna använda får man implementera aktuella fonter som ett komplement för att kunna representera tecknet i läsbar form på skärmen. I exempelvis operativsystemet för PC Windows XP finns ett mycket stort antal fonter vilket underlättar användandet av flera skrivsystem. Det är dock fortfarande inte att rekommendera att använda sig av icke-latinska tecken i exempelvis filnamn om inte operativsystemet i sig är gjort för detta.

### *Bibliotek och unicode*

I en internationell biblioteksmiljö är det önskvärt att kunna hantera ett flertal teckensystem och språk, oberoende av dess skrivsystem, vilket Tull illustrerar genom följande citat:

This standard [unicode] is important to libraries that collect materials in many different languages and want to be able to display the native script in their Web catalogs as well as allow users to search by typing in the native scripts. It is especially useful when trying to display a record that may have multiple scripts, such as a record that may contain both Arabic and Hebrew.

*Tull 2002, s. 181*

För att möjliggöra detta krävs dock att en del frågor undersöks och besvaras. Tull menar att en viktig fråga är hur övergången till unicode kommer att påverka utbytet av bibliografiska poster:

Record exchange between libraries as well as with all sorts of companies around the world is now quite common, and library systems will eventually need to be able to import and export records in



Unicode to do business. During the interim period when everyone is still implementing Unicode, companies will have to offer record exchange in both MARC21 character sets and Unicode until Unicode becomes the dominant character set used by most systems.

*Tull 2002, s. 181*

I sin artikel "Nonroman scripts in the bibliographic environment" beskriver Joan Aliprand övergången till en ny teckenuppsättning som introduktionen av nya katalogiseringsregler. Hon menar att det mest önskvärda skulle vara att göra om katalogiseringen men att detta inte alltid är möjligt. Detta ger två alternativ, antingen stängs katalogen och en ny påbörjas, eller så samkörs två versioner, den gamla och den nya:

The cleanest method of transition to a new encoding scheme is analogous to recataloging: to change over to a system designed for the global character set alone and to translate the existing database into the global character set during implementation of the new system. This might be possible for some individual systems, but some databases are simply too massive to convert. In this situation, the only solution is to superimpose the new character set on the old, mapping between the two sets.

*Aliprand 1992, s. 115*

Att byta teckenuppsättning kräver eftertanke och planering. Oftast finns inte möjligheten att göra om en katalog från början, och arbetet måste därför synkroniseras med den befintliga katalogen under en övergångsperiod. Aliprand menar dock att detta är något som redan pågår vad gäller informationsutbyte mellan system med olika skrift. När ett system baserat på enbart latinska bokstäver tar emot data från ett multiskriftssystem finns två alternativ. Antingen tas informationen som inte kan användas av systemet emot och lagras så att den i framtiden kan användas, eller så tas den omedelbart bort från systemet (Aliprand 1992, s. 115).

#### *Biblioteksystemsleverantörers inställning till unicode*

Laura Tull genomförde 2002 en undersökning bland elva företag som tillhandahåller bibliotekssystem. Undersökningen redovisades i artikeln "Library systems and Unicode: a review of the current state of development". Bland de elva företagen i hennes undersökning hade sex redan implementerat unicode i sina system, medan fem inte hade gjort det ännu. Av dessa fem hade fyra planer på att göra detta senast året därpå. Det enda företaget som inte hade eller planerade att införa unicode menade att för deras marknad var andra teckenuppsättningar mer passande (Tull 2002, s. 183).

Många av företagen hade arbetat med system för icke-latinsk skrift redan före deras införande av unicode, men det rörde sig då om specialterminaler eller mjukvara, inte teckenuppsättningar. Tull menar att detta inte längre är nödvändigt då man nu kan dra nytta av Input Method Editors (IMEs) som finns inbyggt i operativsystem som stöder unicode, exempelvis senare versioner av Windows (2002, s. 183).

I undersökningen efterfrågades även anledningen till att företaget valt att implementera eller inte implementera unicode i sina system. De flesta svaren åberopade efterfrågan och konkurrenskraft, antingen var implementeringen ett svar på kunders direkta efterfrågan eller ett led i expansion till länder med behov av mer omfattande teckenstöd. Det rörde sig även om förändrade behov på den befintliga marknaden då stora

folkförflyttningar äger rum över hela världen vilket får till följd att länders befolkning förändras och därmed även befolkningens språkliga bakgrund. ”Companies respond to growth of multiple languages within their own country to better serve their existing customers as well as market to libraries in other countries” (Tull 2002, s. 184).

Tull förutspår, baserat på sin undersökning, att allt fler företag kommer att anpassa sina produkter till unicode standarden, samt att detta kommer att pressa även de som inte gått över då utbyte av poster kommer att påverkas (Tull 2002, s.184). Hon avslutar med att konstatera unicode i sig själv inte är en komplett lösning för hantering av flera språk i bibliotekssystem, men att det är en stabil grund att bygga på som kommer att underlätta hanteringen av flerspråkigt material (Tull 2002, s.185).

### *Kritik mot unicode*

Det är oundvikligt att ett så stort och mångfacetterat projekt som Unicode Standard möter motstånd och kritik från olika håll. Här tar vi upp generell kritik men även lite av den kritik som kommer från de så kallade CJK-språken (kinesiska, japanska och koreanska), som är den grupp som varit mest missnöjd med hur unicode behandlar deras språk. Suzanne Topping har skrivit en artikel som tar upp den mest förekommande kritiken mot unicode: “The secret life of unicode: a peek at unicode’s soft underbelly”. Topping sammanställer och bemöter i artikeln den kritik som kommit fram genom en diskussionslista på Internet där personer som arbetar med teckenuppsättningar uppmanades att lista för- och nackdelar med unicode som teckenuppsättning. Topping själv är i grunden positiv till unicode, och bemöter därför kritiken i sin artikel. Trots omfattande informationssökning har vi inte kunnat hitta artiklar skrivna av uttalade kritiker av unicode.

### *Generell kritik*

Förenklat kan man säga att generell kritik mot unicode oftast handlar om tre huvudproblem, nämligen oro för att antalet kodkombinationer ska ta slut, att processen bakom införandet av nya tecken är osmidig och tar lång tid samt att unicode är ett komplext system som är svårt att överblicka. Kritiken mot att unicode i framtiden ska få slut på antalet koder är något som Topping tar med ro då hon menar att detta ligger så långt fram i tiden att när det blir aktuellt kommer teknologin att ha utvecklats bortom det unicode vi känner idag (2001).

Att introducera nya tecken till unicode standarden är en relativt osmidig process, vilket försvårar utvecklingen av språket och gör koden mindre dynamisk. En tillfällig lösning är att använda sig av ”private use area” (en lokal lösning som bara fungerar i begränsade sammanhang), men detta är inte en hållbar lösning i längden då det i dagsläget inte finns någon standard för hur man ska koda, avkoda eller översätta tecken så att de blir allmänt accepterade (Topping 2001). Private use area (PUA) är alltså en nödlösning som direkt går emot allt vad unicode står för då de tecken som kodas på detta sätt inte kan användas utanför systemet där de uppstått.

På grund av tidigare standarder har unicode tvingast till lösningar och kompromisser för att vara kompatibel med dessa, vilket har lett till att unicode-standarden är ett mycket komplext och tungrovt system (Topping 2001).

Vad gäller språk som arabiska och hebreiska uppstår problem när en text ska innehålla både tecken som läses/skrivs från höger till vänster och de som läses/skrivs från vänster till höger, exempelvis en text på arabiska med enstaka västerländska namn eller siffror. Här kan uppstå effekter där texten verkar ”dansa”, byta form, när de nya tecknen skrivs in. Problemet ligger dock snarare i text input än i koden, varför det är mer lämpligt att rätta till detta på typsnittsnivå än kodnivå (Topping 2001).

#### *CJK-kritiken*

Det största motståndet mot unicode kommer från de östasiatiska språken och deras ”han unification”<sup>7</sup>, det vill säga det faktum att de tecken som är gemensamma för kinesiskan, japanskan och koreanska har slagits ihop, vilket anses problematiskt då de olika språken har olika representationer av tecknen, även om tecknen är snarlika kan små skillnader förekomma. Ett problem består i att även om tecknen för de olika språken har samma grund, samma kodning, så ser de olika ut vad beträffar fonter. Exempelvis så skrivs tecknet för gräs med fyra streck på traditionell kinesiska men bara tre på förenklad kinesiska, japanska och koreanska. Men i unicode har båda varianterna samma kod (U+8349) oavsett vilken glyph, eller font, det rör sig om. Det är viktigt att skilja på kodningen och fonter eller glypher (den visuella representationen av en kod), då det skulle vara mycket svårt att göra en fungerande internationell standard för fonter, till skillnad från koder (Topping 2001).

Han-tecken är uppbyggda av enstaka fasta streck vilka utnyttjas för att sortera tecknen sinsemellan genom att räkna antalet streck i ett tecken. Det finns de som anser att det hade varit bättre att koda de olika radikaler<sup>8</sup> istället för de färdiga tecknen, då detta skulle innebära att hanteringen skulle gälla ett hundratal koder istället för de idag tiotusentals koder som krävs för att täcka in de flesta tecknen i de tre östasiatiska språken. Problemet med denna metod är dock att det i nuläget inte finns något effektivt sätt att arbeta med radikaler i en datormiljö (Topping 2001).

En annan del av problematiken är av mer kulturell art då många uppfattar det som att det är språken som blivit sammanslagna snarare än att koderna blivit det. Då man tidigare har haft teckensystem utvecklade utifrån det egna språket anser många att unicode har buntat ihop dem, något som blir problematiskt i en region som har haft, och till viss del fortfarande har, spända relationer mellan länderna (Topping 2001).

Även om antalet tecken i unicode är tillräckligt i de allra flesta vardagssammanhang, så finns det en svårighet att täcka in alla tecken som finns i kinesiskan och japanskan vad gäller mer specialiserade sammanhang som högre forskning inom smala discipliner samt alla tecken som används för exempelvis historiska japanska personnamn. Detta har man löst genom att använda sig av funktionen private use area (se ovan), som tillåter egna tecken, men som inte får officiella koder, vilket gör att de får begränsad användningsmöjlighet i andra system (Topping 2001).

<sup>7</sup> Han kommer av att tecken har sitt ursprung i Handynastins Kina, och fungerar som ett samlingsbegrepp för de ursprungligen kinesiska tecken som idag används, med viss modifiering, av kinesiskan, japanskan och koreanskan.

<sup>8</sup> En radikal är ett bastecken som kan ingå i ett mer komplext tecken. Exempelvis är 田 och 力 radikaler i tecknet för man: 男.

Ytterligare ett problem inom CJK-språken ligger inom indexeringen och det faktum att tecknen sorteras i nummerordning efter de kodnummer som tilldelats dem av unicode. Då del flesta användare av systemet inte vet vilka koder tecknen blivit tilldelade kommer dessa att sorteras i en för användaren fullständigt ologisk ordning (Della Porta 2005, s. 32).

På grund av CJK-kritiken florerar ett antal alternativa lösningar för teckenstöd för de aktuella språken. Framför allt i Japan har alternativa teckenstöd utvecklats vid sidan om unicode, men det rör sig om mindre aktörer som egentligen inte utgör något direkt hot mot unicode standarden, som trots allt är den mest använda (Topping 2001).

Då unicode är tänkt att innefatta alla världens skrivsystem är det oundvikligt att problem uppstår. Att i ett och samma system arbeta med problematik som japanskans oro över att inte få in alla sina mer ovanliga tecken till arabiskans problem med duala textriktningar är ett mycket komplext arbete med många faktorer att ta hänsyn till. Även om unicode har sina fel och brister, så är det utan tvekan den mest omfattande och utbredda standardens som finns att tillgå idag, vilket Suzanne Topping uttrycker i sin artikel:

Although Unicode might not be a perfect solution to the challenge of handling all the world's characters, it has moved us a long way toward being able to create systems that can deal with a wide range of languages. During its evolution, the original design goals have had to evolve, and in some ways degrade, to meet real-life challenges. As with any technology, Unicode will undoubtedly be replaced by something that works better. Eventually. But we aren't likely to see that happen any time soon. For now, it is a long-awaited and much-appreciated solution to the world's multilingual computing requirements.

*Topping 2001*

## Standardisering

Då vi i denna uppsats talar om icke-latinsk skrift i bibliotekskataloger är det nödvändigt att beskriva hur denna typ av skrift kan hanteras vid katalogisering. Vi talar i det här kapitlet om standardisering, vilket är en förutsättning för katalogisering och utbyte av bibliografiska poster, och med standardisering avser vi så väl teckenuppsättning som katalogiseringsregler. Kapitlet tar upp MARC-formatet och dess rekommendationer vad gäller icke-latinsk skrift samt användningen av unicode i MARC.

### Utbyte av bibliografiska poster

Utbyte av bibliografiska poster är ingen ny företeelse i biblioteksvärlden. Library of Congress har sålt färdtryckta katalogkort till bibliotek i närmare ett sekel (Understanding MARC bibliographic: parts 1 to 6, 2003). Genom att dra nytta av andras arbete kan biblioteken spara pengar då mindre tid måste läggas på att primärkatalogisera det egna materialet. Idealet är att en bok inte ska behöva primärkatalogiseras mer än en gång. Datoriseringen har aktualiserat denna princip då möjligheterna till utbyte över inte bara biblioteksgränser utan även landsgränser nu är fullt möjligt.

Once computers became available, it was no longer necessary for librarians everywhere to constantly "reinvent the wheel." Why should hundreds of catalogers each use valuable time to compose nearly identical cataloging records for the same item when one cataloger could do it and share the record that had been created?

*Understanding MARC bibliographic: parts 1 to 6, 2003*

För att en större mängd parter ska kunna dra nytta av en katalog krävs dock att samma system används, särskilt när informationen är datorbaserad, eftersom en dator kräver vissa standarder för att kunna tolka och visa den avsedda informationen. Genom internationellt accepterade standarder kan detta utbyte underlättas.

### MARC-formatet

För att en dator ska kunna tolka information krävs att informationen ges en beskrivning som datorn kan härleda till något den förstår. MARC är ett system för att markera information med den typen av beskrivning så att datorn vet vilken information som lagts in.

MARC (Machine-Readable Cataloging Record) character set utvecklades av Library of Congress i USA under 1960-talet. 1969 publicerades standarden under namnet USMARC character set. MARC-standarden används idag av bibliotek i hela världen och är den rådande standarden för hantering av bibliografisk data. Under utvecklandet i USA pågick ett liknande projekt i Storbritannien, men man beslutade sig 1968 för att gå

samman för att överbygga barriären mellan länderna. Trots att man arbetade tillsammans kom två versioner av MARC att utvecklas, USMARC och UKMARC, då de båda länderna hade ett behov av viss anpassning efter lokala förhållanden. Efterhand som fler länder anslutit sig till MARC standarden här även ytterligare lokala versioner utarbetats, och det finns nu ett 50-tal olika MARC-format. MARC-formatet i sig underlättar för utbyte och import av bibliografiska poster internationellt, men på grund av de olika lokala praxis som finns krävs oftast viss bearbetning av importerade poster (Risberg 2003, s. 21). Under slutet av 1990-talet utvecklades den senaste versionen, MARC21, som är det format som idag används av de allra flesta biblioteks som använder sig av MARC-formatet (USMARC and CAN/MARC become MARC 21, 1998).

#### *MARC-formatets teckenuppsättning MARC-8*

MARC-8, som bygger på ASCII och ANSEL<sup>9</sup>, är standardteckenuppsättningen för MARC-poster och utvecklades på 1960-talet. På 70- och 80-talet lades även en rad icke-latinska skrivtecken till, nämligen CJK-språken<sup>10</sup>, hebreiska, arabiska, kyrilliska och grekiska (McCallum 2005, s. 3). I början av 1990-talet beslutade administratörerna av MARC att inte lägga till fler teckenmöjligheter till MARC-8, då man väntade sig att en snar övergång till unicode, vilket dock skulle komma att dröja (McCallum 2005, s. 3). Anledningen till att man var så optimistisk till unicode i ett så tidigt stadium var att kostnaderna för att lägga till tecken i MARC-8 kod var höga då de var tvungna att anpassas till en rad olika bibliotekssystem med olika förutsättningar (McCallum 2005, s. 3).

MARC-8 är en 8-bitars kod och har en ”variable-byte set”, vilket gör att bokstäver med tillhörande diakriter kodas som två olika byte, en för det diakritiska tecknet som kommer före bokstaven och en för själva bokstaven. Det är vid behov möjligt att kombinera en bokstav med mer än en diakrit (Erickson 1997, s. 176). Teckenuppsättningen utvecklades för att tillgodose behovet av diakritiska tecken för att representera både de europeiska språken och romaniserade former av icke-latinsk skrift (Risberg 2003, s. 20). Med tiden har även kapacitet för andra alfabet och skrivtecken utvecklats inom MARC och man stöder idag följande språk och skrivsystem: latin, JACKPHY<sup>11</sup>, kyrilliska och grekiska (Della Porta 2005, s. 6).

#### *Bibliografiska poster i MARC*

Katalogisering i MARC-formatet kräver en relativt god inblick i detta omfångsrika system. I sin magisteruppsats *Metadata för det digitala biblioteket* beskriver Miriam Björkhem och Jessica Lindholm det så här: ”Det krävs specialistkunskap för att kunna katalogisera i MARC, och denna indexering utförs ofta centralt, vid ett huvudbibliotek eller motsvarande. Andra bibliotek nöjer sig med att kopiera posterna och lägga till sina lokala noteringar” (Björkhem & Lindholm 2000, s. 30). Att kunna knyta det egna beståndet till redan färdiga poster sparar bibliotekens tid och pengar, då

<sup>9</sup> ANSEL är namnet på en teckenuppsättning avsedd för att behandla ”extended latin”, dvs. en utbyggnad av det latinska alfabetet innefattande 36 specialtecken och 29 diakriter.

<sup>10</sup> Dessa är kodade i 24 bits istället för 8, vilket ger de 15,852 vanligaste tecknen

<sup>11</sup> JACKPHY är akronymen för följande språk: japanska, arabiska kinesiska, koreanska, persiska, hebreiska och yiddish.

primärkatalogisering i MARC-formatet kräver både utbildning och tid för den inblandade personalen.

MARC-formatet är uppbyggt kring fält som genom siffror definierar vilken funktion datan ska ha i posten. Exempelvis läggs titelinformation i ett fält med beteckningen 245, som förklarar för datorn vilken typ av information titeln är.

Formatet är uppbyggt kring en tresiffrig kod enligt följande mönster:

0XX	Kontrollinformation, nummer, koder
1XX	Huvuduppslag
2XX	Titel, utgåva, publikationsinformation
3XX	Fysisk beskrivning
4XX	Serieuppgifter hämtade från boken
5XX	Anmärkningar
6XX	Ämnesord
7XX	Biuppslag på exempelvis författare eller titel
8XX	Seriebiuppslag, beståndsuppgift, elektronisk adress

9XX-fälten har lämnats öppna för lokala anmärkningar som till exempel streckkoder (Understanding MARC bibliographic: parts 1 to 6, 2003). Idag används inte 9XX-fält för katalogisering i Libris, lokala tillägg görs istället i beståndsposten eller i bibliotekets egna lokala system.

För katalogisering av böcker i Libris i Sverige är de mest använda MARC-fälten följande:

020	International Standard Book Number (ISBN)
040	Det katalogiserande bibliotekets sigil
084	SAB-klassifikation
100	Huvuduppslag, författares namn
245	Titelinformation, deltitel och upphov
250	Upplaga
260	Publikationsinformation: stad, förlag och år
300	fysisk beskrivning, antal sidor, illustrationer
440	Serieuppgift
650	Ämnesord
700	Biuppslag för författare

Utöver dessa fält består MARC-formatet även av delfält, vars utformning varierar beroende på vilket fält de finns i. Exempelvis kan fält 245 (titel) innehålla tre delfält och ser då ut så här:

245 14 \$a The conversion of scripts : \$b it's nature, history, and utilization / \$c Hans H. Wellisch

Dollartecknet (\$) symboliserar den delfältsmarkör som markerar var ett delfält börjar. Ettan och fyran som står mellan fältnumret och delfält a kallas indikatorer, och i det här fallet står den första indikatorn, 1, för att titeln är ett sökelement medan den andra

indikatorn, 4, står för antalet tecken som ska räknas bort vid indexering, i detta fallet 4, eftersom titeln innehåller ”the” samt ett mellanslag innan första sökordet.

#### *Fält 880 – Alternate graphic information*

Som ett led i arbetet med att implementera icke-latinsk skrift i MARC-formatet utarbetades i början av 1980-talet ett särskilt fält som fick koden 880, Alternate graphic information. Fältet gjorde det möjligt för de bibliotek som önskade att lägga in icke-latinsk skrift, men även för de som inte har möjlighet att hantera skriften att välja bort det. Den information som läggs in i 880-fältet länkas till det av de reguljära fälten som det motsvarar genom ett delfält 6 följt av koden på det fält det motsvarar. Exemplet nedan är hämtade från Library of Congress hemsida om 880-fältet: Holdings, Location, Alternate Graphics, ect. Field (841-88X):

Exempel 1 **880 1#\$6100-01/(N \$a** [Heading in Cyrillic script]

Exempel 2 **880 10\$6245-02/\$1 \$a** [Chinese characters]

Exempel 3 **880 ## \$6 260-02/(2/r \$a** [Tel Aviv] :\$b) [Hebrew characters]\$c [1985]

I exempel 1 handlar det om en upphovsuppgift (fält 100) med kyrillisk skrift. Koden är uppbyggd av delfält 6 (för länkning), koden för fältet med upphovsuppgifter (100) samt en tvåsiffrig kod för att associera det länkade fältet med det reguljära fältet. Efter snedstreck (/) följer koden som specificerar vilken skrift som använts i fältet, i detta fallet (N för kyrillisk skrift. Efter skriftkoden följer den vanliga delfältsmarkören med tillhörande bokstav.

Även exempel 2 är uppbyggt efter samma princip, med skillnaden att det rör sig om ett 245 fält (titel) samt att texten i fältet är kinesiska med skriftkod \$1.

I exempel 3 visas ett 880-fält som är länkat till ett 260-fält (publikationsinformation). Koden följer samma mönster som de två andra fram till precis efter skriftkoden, som i detta fallet är (2 för hebreiska, som följs av ett r. Detta är indikerar att skriften ska följa höger till vänster riktning i stället för defaultvärdet som är vänster till höger.

I den senaste versionen av MARC-formatet, MARC21, är det möjligt att använda sig av icke-latinsk skrift i alla fält, inte bara i 880. Dock rekommenderas inte att använda annat än ASCII i fält 006, 007, 008 samt underfält som 533, 7XX och så vidare (se nedan) (McCallum 2005, s. 11). Hur man väljer att använda sig av 880-fältet beror på vilket utgångsläge man har. Ett bibliotek i Sverige skulle troligtvis välja att använda sig av svenska (latinsk skrift) i de reguljära fälten och ha icke-latinsk skrift i länkade 880-fält, medan ett bibliotek i Grekland troligtvis skulle föredra att använda sig av grekiska skrivtecken i de reguljära fälten och en eventuell romanisering med latinska bokstäver i de länkade 880-fälten. Oavsett vilket språk och vilka skrivtecken man använder sig av i de reguljära fälten är dock MARC utvecklat i en engelsktalande miljö och anpassat för dessa tecken i förstahand, vilket förklarar rekommendationerna att enbart använda sig av ASCII i vissa kritiska fält.

Trots att det är tekniskt möjligt att använda sig av icke-latinsk skrift i alla MARC-fält ser verkligheten annorlunda ut. Både RLG och OCLC kräver att icke-latinsk skrift ska



begränsas till 880-fält i sina poster. Detta har i sin tur lett till att mottagarna av poster förväntar sig att posten följer denna praxis, vilket gör att de aktuella biblioteken kan ta ställning till hur det vill göra med 880-fältet, dvs. om de vill visa dem i sin katalog eller undertrycka dem (McCallum 2005, s. 11). Om detta är ett krav som RLG och OCLC kommer stå fast vid även i fortsättningen har vi inte kunnat befästa i våra källor. Däremot har vi genom en mejllista för personer som arbetar med MARC och Unicode frågor kunnat bekräfta att Library of Congress så sent som augusti 2005 rekommenderar att inte använda andra fält för icke-latinsk skrift, trots att det är tekniskt möjligt enligt MARC-formatet (Unicode-MARC Archives August 2005 #54, #55 och #60).

## MARC och unicode

Även om MARC-8 i huvudsak är baserad på ASCII ligger en stor del av repertoaren utanför ASCII, vilket gör MARC-8 formatet unikt för biblioteksvärlden och dess hantering av data. På grund av att formatet är unikt finns mycket begränsat stöd för användning av MARC utanför biblioteksvärlden, vilket begränsar dess användning. Genom att övergå till unicode för lagring av text tar man ett steg närmare kompatibilitet med den övriga datorbaserade världen då man istället kommer använda sig av samma kodning som exempelvis Internet, nyare Microsoftprodukter samt programmeringsspråket Java (Cain 2004, s. 2).

Redan 1994 lades unicode till som en alternativ teckenuppsättning i MARC, och mellan 1994 och 2001 pågick arbetet med att mappa mellan unicodetecknen och MARC-8 tecknen för att underlätta övergången från MARC-8 till unicodeuppsättningen. Under detta arbete försökte man i så stor utsträckning som möjligt att undvika att använda sig av unicodes private area use för att minimera kommunikationsproblem mellan olika system (McCallum 2005, s. 3). McCallum bedömer att det kommer att ta flera år innan alla bibliotekssystem är kompatibla med unicode, varför det är nödvändigt att fortsätta exportera data även i MARC-8-formatet för att tillgodose allas behov under övergångsperioden. De bibliotek som har kapacitet att använda sig av fler tecken är givetvis angelägna om att göra detta, vilket kan bli problematiskt för de bibliotek som fortfarande har äldre system och ännu inte kan hantera tecken utanför MARC-8-repertoaren. För att dessa två system ska kunna samexistera i övergångsperioden krävs konventioner för den här typen av export av data (McCallum 2005). Som Cain skriver i sin artikel: "Character set is very fundamental and very unforgiving – if a system does not know about a character or code – then data corruption will result unless proper measures are taken" (2004, s. 4). För att kunna hantera övergången mellan de två olika systemen kommer krävas godtagbara lösningar som tillåter överföring både från UTF-8 till MARC-8 och tvärt om. Cain anser dock att eftersom trenden går mot att fler och fler system implementerar unicode, vilket även sätter press på de som inte gjort det ännu, bör inte tid eller pengar investeras i allt för komplexa lösningar (2004, s. 4).

Ett antal rekommendationer har föreslagits av de grupper som arbetat med frågor kring unicode och MARC, och man har kommit fram till några gemensamma beslut:

- Den version av unicode som ska användas i MARC-formatet är UTF-8
- I fält 006, 007 och 008 i den bibliografiska posten ska enbart ASCII-tecken användas för att förhindra problem för system som inte kan hantera andra teckenuppsättningar.

- Text i UTF-8 MARC-poster ska placera diakriterna efter, inte före som tidigare, de tecken som de modifierar.
- Att posten är i UTF-8 ska specificeras med hjälp av en kod i fältet ”leader” i den bibliografiska posten.
- Fält 066, som talar om vilken teckenuppsättning som använts i posten, ska inte användas i unicodeposter.
- En bibliografisk post får inte innehålla både MARC-8 och UTF-8 (McCallum 2005, s. 4).

Omvandlingen av tecken från MARC-8 och unicodes UTF-8 har utvecklats med hänsyn till de olika dataexportsystem som idag används av biblioteken. Omvandlingen ska kunna genomföras oavsett om exporten sker genom en batch import, många poster som förs över på samma gång, eller en on the fly import genom exempelvis z39.50<sup>12</sup> (McCallum 2005, s. 9).

Resultatet av en import av en bibliografisk post kan se lite olika ut beroende på det mottagande systemets begränsningar. Om en post innehåller data som det mottagande systemet inte kan visa kommer datan troligtvis att ignoreras av systemet. McCallum förklarar det så här:

[...] various systems handle MARC-8 according to their capabilities. For example, if an ASCII/ANSEL record came to an ASCII-only system, then a mapping may reduce the set to ASCII. Non-Latin characters entering systems that cannot support MARC-8 non-Latin may be ignored or discarded.

*McCallum 2005, s. 11*

Det mottagande systemet måste alltså innehålla samma teckenuppsättning som det exporterande för att posten ska få exakt samma innehåll. För att motverka att en post med originalskrift blir oanvändbar efter export till ett annat system bör posten innehålla romanisering av de obligatoriska fälten som titel, upphov och utgivning och ha originaltecken/skrift i 880-fält då det finns möjlighet att undertrycka de fält som innehåller mer kärvande text och istället bara visa de fält som innehåller romaniserad text (se stycket om 880-fält ovan).

#### *Library of Congress övergång till unicode*

Genom att övergå till unicode som teckenuppsättning blir UTF-8, som är den vanligaste versionen av unicode för databaser, standarden för lagring av text i stället för MARC-8, som är MARCs egen teckenuppsättning. Då Library of Congress under 2005 övergick till en unicodeversion av sin katalog bestämdes även att text skulle lagras i UTF-8, även om MARC-8 kommer att finnas som ett alternativ i en övergångsperiod, trots att övergångsperioden beräknas bli lång. Så här skriver man i en presentation av det nya systemet:

... any Unicode encoding can be used in a MARC record in UTF-8. Each library will have to make a decision about whether it should limit its data to characters that have MARC-8 mappings. What this means is if we input Thai characters in a MARC record in UTF-8, we will not be able to export a

<sup>12</sup> Z39.50 är ett standardiserat protokoll som möjliggör kommunikation mellan olika databaser. Protokollet används ofta för att importera enstaka poster åt gången från en central databas till det lokala systemet.

record in MARC-8 with those characters, nor would these characters be delivered via Z39.50 in MARC-8. For LC, we'll have to take into consideration what our customers need. Any library that has not yet converted to Unicode will still need to get MARC-8 records from the Cataloging Distribution Service. I think the real question will be, how long we'll have to provide MARC-8 records. I'll remind you that LC stopped selling printed cards just 8 years ago.

*Della Porta 2005, s. 24*

Byte av katalogiseringsrutiner eller teckenuppsättning i en databas är alltså ett omfattande arbete med lång startsträcka och inget som görs över en natt. Arbetet med att alfatesta MARC-konverteringen av poster skedde i samarbete med Library of Congress. En kopia gjordes på bibliotekets produktionsdatabas innehållande 13 miljoner bibliografiska poster, nästan 14 miljoner beståndsposter och 6 miljoner auktoritetsposter. Av dessa poster innehöll c:a en halv miljon data på de s k JACKPHY-språken, men konverteringen gjordes på alla MARC-posterna, oavsett språk. Man valde ut 150 poster per språk för konvertering och stickprov analyserades av Library of Congress tekniska experter genom att titta på fel-loggar samt import och export av data. Statistiken visade att antalet poster som lades i felläge vid konverteringen endast var 0.01%. Cirka 10 % av dessa berodde på att posterna innehöll tecken från CJK-språken som använt sig av tecken från private use area (se kapitlet om unicode). Library of Congress beslutade sig i ett tidigt skede för att inte använda sig av private use area tecken i sin databas då de ansåg att de inte kunde garantera dessas integritet i databasen, det går inte att garantera att tecknen representeras som de ska i posterna (Della Porta 2005, s. 15).

Efter den fullständiga övergången till unicode planerar Library of Congress att fortsätta använda sig av 880-fältet trots att deras mjukvara tillåter att icke-latinsk skrift läggs in var som helst i posten. Enligt Della Porta (2005, s. 21) ligger beslutet om att katalogisera enbart på originalspråk hos de berörda bibliotekarierna. Hon ifrågasätter om biblioteken kommer att ha råd att göra tvåspråkiga poster med dagens hårda ekonomiska klimat, men ställer sig även frågande till om det är möjligt att administrera en katalog utan romanisering. En lösning på detta kan, enligt Della Porta, vara automatisk translitterering (Della Porta 2005, s. 21). Dessa frågor saknar i skrivande stund svar eftersom de är så aktuella och fortfarande är under utredning.

## Empirisk studie

Redovisningen av vår empiriska studie är uppdelad i tre delar: presentation av informanterna och deras bibliotek samt aktuella bibliotekssystem, resultatredovisning av intervjuerna som i sin tur är uppdelad i tre olika teman och slutligen ett kort avsnitt som redovisar vår kontakt med Libris.

### Presentation av biblioteken, bibliotekssystemen och informanterna

#### **Biblioteken**

##### *Malmö stadsbibliotek*

Malmö stadsbibliotek är ett stort bibliotek med över 250 000 inskrivna användare, över 600 000 medier och sammanlagt nästan två och en halv miljon utlån på stadsbiblioteket tillsammans med de tio stadsdelsbiblioteken under 2004. Under samma period hade biblioteken 141 anställda.

Stadsbiblioteket ingår tillsammans med Malmö Museer, Malmö Konstmuseum, Malmö Kulturmiljö Malmö Konsthall, Malmö Kulturkompani och Malmö stadsarkiv i Malmö kulturförvaltning. Malmö stads kommunbidrag uppgick för år 2004 till 89 143 tkr varav 6 700 tkr utgjordes av anslag till förvärv av medier. Stadsbiblioteket leds av stadsbibliotekarien som under sig har enhetschefer för avdelningarna barn & unga, humaniora & samhälle, konst & musik, medicentral, natur & teknik, nyhetsmedia & torget, särskilda tjänster samt stabsfunktionerna systemavdelning med PR-frågor, ekonomi och personal. Under året genomfördes en omorganisation innebärande att en gemensam chef utsågs för avdelningarna barn & unga, konst & musik samt skönlitteratur & språk.

Under 2004 har ett intensivt arbete pågått för att ta fram en samlad medieplan i syfte att öka kunskapen om mediebeståndet, effektivisera mediehanteringen, profilera samlingarna samt att utgöra ett underlag vid fördelning av medieanslaget samt ligga till grund för långsiktiga analyser.

Stadsbiblioteket satsar mycket på verksamhet riktad mot invandrare och personer med ett annat modersmål än svenska. Under 2004 anlätade stadsbiblioteket gymnasieelever med persiska, arabiska och somaliska som modersmål för att tjänstgöra på Stadsbiblioteket två timmar per vecka. Efter att ha fått information om stadsbiblioteket samt katalogen och Internetresurser har de kunnat hjälpa personer som talar deras modersmål till ett mer aktivt utnyttjande av biblioteket. De tog också emot önskemål om inköp av böcker och prenumeration på tidningar och tidskrifter. Projektet har enligt

stadsbiblioteket fallit mycket väl ut och väckt intresse på andra bibliotek. Under hösten 2004 tillkom även guider på bosniska och albanska. Guiderna har också arbetat på bokbussen och varit till stor hjälp i områden med en mångkulturell befolkning.

Bibliotekens satsning på mångfald och tillgänglighet för alla berörs också i ett av målen i den biblioteksplan som antogs under 2005: ”Biblioteken ska vara tillgängliga för alla oavsett bakgrund och ge fri och jämlik tillgång till information och upplevelser samt väcka nyfikenhet och engagemang. Faktorer som boende, ålder, arbetstider, funktionshinder, språk och ekonomi ska inte utgöra något hinder” (Biblioteksplan Malmö stadsbibliotek 2005, s. 3). I samma biblioteksplan står också att ”Biblioteken ska fortsatt ligga i frontlinjen med hänsyn till utveckling och deltagande i virtuella/digitala tjänster både regionalt, nationellt och internationellt” (Biblioteksplan Malmö stadsbibliotek 2005, s. 5).

#### *Rosengårds stadsdelsbibliotek*

Rosengårds stadsdelsbibliotek är beläget i en invandrartät stadsdel i Malmö. Biblioteket besöks av ca 1200 personer per dag och har böcker på 50 olika språk. Man vänder sig till olika invandrargrupper och har bland annat sagostunder på bosniska, arabiska och albanska.

#### *Lunds universitets bibliotek*

Biblioteken vid Lund universitet skiljer sig från Malmös stadsbibliotek och stadsdelsbiblioteken i Malmö, både vad gäller organisation och vilka användargrupper som man primärt vänder sig till.

Biblioteken vid Lunds universitet bildar ett nätverk med de två gemensamma enheterna Biblioteksdirektionen och Universitetsbiblioteket samt ett trettio-tal institutions-, fakultets- och centrumbibliotek i Lund, Malmö och Helsingborg. (Lunds universitets bibliotek - organisation) Dessa c:a trettio bibliotek står för merparten av biblioteksservicen för studenter, lärare och forskare vid Lunds universitet och finansieras av respektive fakulteter och institutioner. Biblioteken har en del gemensamma policys och ett gemensamt bibliotekssystem, men är självständiga i de flesta frågor som rör den egna verksamheten. Biblioteksdirektionen verkar som ett samordnande organ för Lunds universitets biblioteks nätverk. Vid biblioteksdirektionen arbetar man i huvudsak med de för nätverket gemensamma elektroniska resurserna såsom bibliotekssystemet Lovisa (Virtua) samt databaser och elektroniska tidskrifter samt samordnar arbetet inom nätverket och utvecklar gemensamma policys (Lunds universitets bibliotek - organisation ).

Universitetsbibliotekets ansvarsområde ligger i hantering av det svenska trycket (universitetsbiblioteket är skyldigt att bevara och tillgängliggöra allt svenskt tryck) samt ett antal specialsamlingar, allmänt referensbibliotek, arkiv och depåer (Lunds universitetsbibliotek - organisation ). För vårt ändamål är i första hand två av nätverksbiblioteken intressanta, nämligen Språk- och litteraturcentrums bibliotek samt Asienbiblioteket. Både språk- och litteraturcentrums bibliotek och Asienbiblioteket bedriver verksamhet där hantering av material på språk med icke-latinska skrivsystem förekommer i stor utsträckning. Biblioteken serverar studenter, lärare och forskare som

dagligen använder sig av språk som omfattas av den för undersökningen aktuella problematiken

Även biblioteksdirektionen är av intresse för vår studie då det är där övergripande policy för nätverket utformas, inklusive policys för nätverkets katalog.

#### *Språk- och litteraturcentrums bibliotek*

Språk- och litteraturcentrums bibliotek, även kallat SOL-biblioteket, bildades efter en omorganisation hösten 2004, då språk- och litteraturinstitutionerna vid Lunds universitet fick en gemensam organisation. 27 olika ämnen samordnas under centrum, och av dessa ämnen är över hälften språkutbildningar, sex av språken har icke-latinska skrivsystem: arabiska, grekiska, hebreiska, japanska, kinesiska och ryska (Språk- och litteraturcentrum vid Lunds universitet - ämnen).

Arbetet med att flytta verksamheten från institutionerna till Språk- och litteraturcentrums bibliotek pågår fortfarande och först under sommaren 2006 kommer samlingarna på latin, grekiska, semitiska och slaviska språk att flyttas till de nya lokalerna (Språk- och litteraturcentrum vid Lunds universitet - organisation). Det är i dagsläget oklart om även de japanska och kinesiska samlingarna som i dag finns på Asienbiblioteket (se nedan) kommer att flyttas till Språk- och litteraturcentrums bibliotek.

#### *Asienbiblioteket*

Asienbiblioteket drivs i nuläget av Centrum för Öst- och Sydöstasienstudier samt avdelningen för kinesiska och avdelningen för japanska. Biblioteket används framför allt av studenter och forskare inom Asienstudier, både realia och språk. Biblioteket har en specialsamling innehållande c:a 10 000 böcker inom en rad olika ämnen på kinesiska och c:a 3000 på japanska, och anses vara det främsta av sitt slag i Norden (Centre for East- and Southeast Asian Studies).

### **Bibliotekssystemen**

För vår undersökning har vi valt att se närmare på tre olika bibliotekssystem. Dessa är för Malmö stadsbibliotek samt Rosengårds stadsdelsbibliotek Innovatives "Millennium", för biblioteken inom nätverket för Lunds universitets bibliotek VTLS' "Virtua" och Endeavors "Voyager" för den nationella bibliotekskatalogen Libris. De olika systemen liknar varandra så till vida att de alla är tillverkade av företag baserade i USA samt att alla har gjort någon form av unicodeanpassning efter marknadspåtryckningar. Virtua är det system som längst arbetat med unicode då VTLS tidigt marknadsförde sin produkt för en internationell marknad, vilket innebar att systemet designades för flerspråkig representation redan från början (Virtua ILS). Idag har dock även Millennium och Voyager möjligheten till unicode i sina system, vilket verkar spegla den utveckling som Laura Tull förutspådde i sin undersökning av företag som tillhandahåller bibliotekssystem, där hon konstaterade att globaliseringen och konkurrensen mellan företagen kommer tvinga fram en övergång till unicode som teckenuppsättning för majoriteten av bibliotekssystemen (se kapitlet om unicode för mer information om Tulls undersökning). Om denna utveckling kan man även läsa i en artikel på American Library Associations' hemsida:

Now that a many of the major vendors of automated library systems are selling systems throughout the world, it is important to them to support far more characters than the ubiquitous ASCII character set can accommodate. Academic research libraries that collect globally also must be able to include records for many different languages and scripts in their catalogs. The increasing diversity of North America's population also makes it increasingly important for its public libraries to acquire, catalog, and make available materials in many languages, and to store and display the records in the vernacular, rather than in transliterated form.

*Boss 2003*

### *Millennium (Malin)*

Millennium är en av bolaget Innovatives produkter och är det bibliotekssystem som används på Malmö stadsbibliotek. Systemet är baserat på moduler för bland annat cirkulation, förvärv och katalogisering, och används av olika typer av bibliotek i 42 länder (Innovative Interfaces 2004). Databasen är MARC-kompatibel och stödjer olika former av import och export av bibliografiska poster från andra system (Innovative Interfaces Millennium: Cataloging 2004, Innovative Interfaces Millennium: Technical information 2004).

Millennium stödjer en rad olika teckenuppsättningar, exempelvis ANSEL, EACC (CJK-språken), samt ISO för estniska, polska och turkiska (Innovative Interfaces Millennium: Cataloging 2004). På grund av en lansering av systemet i Östasien tvingades man att implementera även Unicode för att tillgodose behovet med flera parallella skrivsystem och teckenuppsättningar i samma bibliotekssystem (Boss 2003).

Innovative är medlemmar i Unicode Consortium och har möjligheten att erbjuda unicodeanpassning i sitt system, även om det är upp till de olika användarna av systemet att besluta om huruvida de vill använda sig av unicode som teckenuppsättning eller om de föredrar en annan.

### *Virtua (Lovisa)*

Virtua är utvecklat av det amerikanska företaget VTLS Inc. och riktar sig mot en global marknad med kunder i ett trettiotal länder (VTLS Inc. Corporate Information 2004). Företaget är aktiva medlemmar i föreningar som American Library Association, International Federation of Library Associations and Institutions, Unicode Consortium samt gruppen för utvecklandet av standarden z39.50 (VTLS: Visionary Technology in Library Solutions ). Liksom Innovatives Millennium är Virtua upplagt i moduler för de olika funktionerna i biblioteksverksamheten som cirkulation, förvärv, fjärrlån och katalogisering.

VTLS' Virtua var det första datoriserade bibliotekssystemet som var helt kompatibelt med unicode standarden. Redan i mars 1999 utannonserade VTLS att all data i alla poster skulle sparas i unicode, vilket tillät systemets användare att använda sig av sina lokala språk i samband med exempelvis katalogisering, men även för all form av input, indexering och representation av tecken från det egna språkets alfabet. Att denna utveckling skedde så snabbt inom VTLS kan bero på att nästan 60 procent av deras kundkrets befinner sig utanför Nordamerika (Boss 2003). I en broschyr om den egna verksamheten uttrycker man sig så här: "As Unicode (UTF-8) compliant software, **all of**

**our products** speak your language, making for truly global information sharing” (VTLS: Visionary Technology in Library Solutions ).

#### *Voyager (Libris)*

Företaget Endeavor står bakom bibliotekssystemet Voyager, vilket är det system som används för den nationella bibliotekskatalogen för högskolebibliotek i Sverige, Libris. Enligt den egna hemsidan används Voyager av mer än 1 300 forsknings- och akademiska bibliotek runt om i världen (Voyager). Voyager erbjuder funktioner för de delar av biblioteksverksamheten som normalt sett krävs av en bibliotekskatalog, till exempel katalogisering, fjärrlån, webbsökningsfunktioner med mera. Endeavor är, liksom de andra två företagen, medlemmar i Unicode Consortium och har även bildat en egen ”Unicode Task Force” grupp bestående av nio bibliotekarier från akademiska bibliotek med flerspråkiga samlingar för att utreda vad det innebär att implementera unicode och Voyager (Voyager with Unicode: Standards for Libraries 2003). På Endeavor är man noga med att ställa kravet att import och export av poster ska fortsätta vara möjligt med andra teckenuppsättningar än unicode (Voyager with Unicode: Standards for Libraries 2003).

#### **Informanterna**

Informant 1 arbetar vid intervjutillfället som kommunikationsbibliotekarie för institutionen för psykologi vid Lunds universitet, men är tillfälligt tjänstledig från sin tjänst på Asienbiblioteket, där hon normalt arbetar med selektion, förvärv och katalogisering. Asienbiblioteket har en stor samlig litteratur på asiatiska språk, främst kinesiska och japanska men även en del thailändska.

Informant 2 har arbetat med katalogisering vid Lunds universitetsbibliotek i många år. Han är vid intervjutillfället anställd som ansvarig för registervård vilket innefattar ansvar för den lokala katalogens bibliografiska kvalitet, utbildning samt viss omvärldsbevakning vad gäller katalogisering och bibliotekssystem. Informanten har god kännedom om problem och möjligheter i Lunds universitets bibliotekskatalog och även god kännedom om Libris. Även om informanten inte själv arbetar med icke-latinsk skrift så har han kommit i kontakt med problematiken genom samarbetet med de katalogisatörer inom Lunds universitets biblioteks nätverk som arbetar med katalogisering av material på olika språk.

Informant 3 är bibliotekarie på Språk- och litteraturcentrums bibliotek vid Lunds universitet. Hennes huvuduppgifter är katalogisering av bibliotekets material. Informanten talar finska och kan läsa estniska. Utöver katalogiseringen så tjänstgör informanten i lånedisken. Hon har även arbetat som kontaktbibliotekarie för slaviska institutionen vid Lunds universitet.

Informant 4 arbetar vid intervjutillfället som bibliotekarie på stadsdelsbiblioteket i Rosengård i Malmö. Informanten tjänstgör i disk och katalogiserar utländsk litteratur, främst arabiska och persiska men även övriga språk förutom polska, ryska och albanska, då det finns annan personal på biblioteken med relevant kompetens för de språken. Han möter dagligen användare med olika modersmål på ett bibliotek som har relativt stora samlingar av material med icke-latinsk skrift.



Informant 5 är chef för mediecentralen på Malmö stadsbibliotek och har ett övergripande ansvar för mediehantering, det vill säga inköp, katalogisering och bokvård samt planering av utveckling av mediehanteringen. Informanten har också erfarenhet av att katalogisera material på icke-latinska språk och att arbeta med hanteringen av tidningar och tidskrifter på flera olika språk.

Informant 6 är funktionsansvarig för katalogiseringsverksamheten på Malmö stadsbibliotek. En stor del av informantens arbetsuppgifter består av att handleda och hjälpa kollegor som arbetar med katalogisering. Informanten katalogiserar också själv och då mycket slavisk litteratur eftersom hon har läst ryska och kan ta sig fram på de slaviska språken.

## Resultat av intervjuerna

Vi har valt att presentera resultaten från intervjuerna under tre olika tema. Dessa är ”teknik”, ”policy, praxis och organisation” samt ”attityder och inställningar”. Under dessa huvudrubriker finns underrubriker baserade på de frågeområden vi använt oss av under intervjuerna (se intervjuguide i bilaga 2).

### **Teknik**

Under temarubriken teknik har vi samlat de svar som kom fram på frågor av mer teknisk art: huruvida informanterna haft vetskap om Libris unicodeanpassning, eventuell kännedom om förutsättningarna för icke-latinsk skrift i MARC-formatet samt hur bibliotekets lokala katalogs förutsättningar och möjligheter för unicode ser ut.

#### *Informanternas kännedom om Libris unicodeanpassning*

Alla tre informanterna från biblioteken inom nätverket för Lunds universitets bibliotek var väl medvetna om det faktum att Libris står inför en unicodeanpassning samt när detta skulle ske.

Två av de informanter vi talade med vid Malmö stadsbibliotek och Rosengårds stadsdelsbibliotek kände igen termen unicode, men var inte riktigt insatta i när denna förändring skulle komma. Den tredje kände inte till begreppet unicode men visste att en förändring av Libris skulle komma under våren.

#### *Kännedom om förutsättningarna för icke-latinsk skrift i MARC-formatet*

Informant 1 känner väl till MARC-formatets rekommendation att använda 880-fältet för inmatning av icke-latinsk skrift då experiment har gjorts på det bibliotek där informanten arbetat. Experimenten gick ut på att mata in kinesiska och japanska tecken för författare och titel i 880-fält i den lokala bibliotekskatalogen som, till skillnad från Libris, var unicodeanpassad och därmed hade möjligheten att ta emot dessa tecken. Anledningen till att detta experiment sedan lades ner var att Librisposten skrev över de ändringar som gjordes i den lokala katalogen när ett annat bibliotek uppdaterat Librisposten, vilket gjorde arbetet i den lokala katalogen förgäves.

Även informant 2 är mycket väl insatt i MARC-formatets förutsättningar och har arbetat aktivt med frågor kring 880-fältet. Han menar att det råder delade meningar inom branschen huruvida 880-fältet är det mest praktiska sättet att arbeta med icke-latinsk

skrift i MARC-poster eller om det finns alternativ, men att användningen av länkade 880-fält för icke-latinsk skrift är den absoluta huvudfåran. Under intervjun visar han utskrifter från Library of Congress webbkatalog för att illustrera hur de har valt att använda sig av fält 880 för att dubbla i stort sett allafälten i posten, vilket ger två titelfält, två huvuduppslag, två fält för publikationsinformation och så vidare, alla med ett romaniserat fält och ett fält med originalskrift. Informanten pratar själv en del om auktoritetsposter, och när vi frågar om icke-latinsk skrift i auktoritetsposter ber han att få återkomma då han inte alls tänkt på vad detta skulle innebära. Han är dock positiv till originalskrift i auktoritetsposter om det är tekniskt möjligt.

Vad beträffar informant 3 känner hon till att MARC-formatet har ett fält 880 för varianttitlar, men har inte alls arbetat med icke-latinsk skrift i MARC-formatet.

De två informanterna från Malmö stadsbibliotek känner inte alls till 880-fältet och dess användning, men efter att vi förklarat hur det fungerar är båda mycket positiva, och informant 6 kan tänka sig att ha dubbla fält för originalskrift på större delen av fälten.

#### *Lokala katalogens förutsättningar att använda unicode*

De informanter som arbetar på bibliotek inom Lunds universitets biblioteks nätverk har alla Virtua som lokal katalog, vilket innebär full unicodeanpassning. Informant 1 säger så här om den lokala katalogens unicodeanpassning:

Samtidigt är ju problemet att är det någon som gör en ändring i posten i LIBRIS slår den ut alla lokala ändringar som man gjort i den lokala databasen, så att då är ju det arbetet förgäves på nåt sätt och det är jättetråkigt. Så att vi är väldigt angelägna om att unicode kommer till användning i LIBRIS också.

Informant 3 anser att teckenhanteringen i den lokala katalogen är mycket bättre än den i Libris då diakriterna placeras bättre i det lokala systemet:

Informant 3: Teckenhanteringen är bättre i Lovisa, jag märker det nu att man kan, visserligen kan man gå in i en liten ruta och hämta tecknen, men så får man det då snyggt och ordentligt. I Libris blir det lite mellanslag och lite...

Intervjuare: Med de här diakriterna?

Informant 3: Ja, just det. Först kommer diakriten och sen kommer bokstaven, ibland blir det fel.

De informanter som arbetar på Malmö stadsbibliotek och Rosengårds stadsdelsbibliotek har Innovatives Millennium som lokalt system, och även om Millennium som system kan använda unicode som teckenuppsättning görs detta inte på dessa bibliotek. Informant 6 ställer sig frågande till hur det skulle gå till att implementera unicode i det lokala systemet:

Hittills är det ett par Millennium bibliotek, KTH och Chalmers, som har gjort det här och det har varit en hel del problem med det så, så vi ska nog välja att ligga lite lågt så här i början, men det är klart att det är intressant att göra det. Samtidigt är det ju väldigt komplicerat för att, inte minst när det gäller indexering då, hur hanterar man det? Kan man ha ett index med latinsk skrift, ett annat med [...] typ kyrilliska då? Det är väl en del frågor som är kvar att lösa där.

På frågan om det finns planer på att inom en överskådlig framtid övergå till unicode som teckenuppsättning svarar informant 6 så här:

Alltså det blir ju inte i sommar! (Skratt) Det blir det ju inte, men vi är ju kända i Malmö för att vara framstegsvänliga och har en väldigt duktig systemavdelning som sagt och så, så att, det, det, jag tror inte det skulle dröja så länge. Men som jag har förstått så är det någonting som fordrar en konvertering till unicode av hela katalogen, och det är ju ett stort arbete, som [...] ja, som kan påverka funktionen av katalogen under den tiden det pågår och, det är frågan om när man ska göra det. Om man ska göra det när det är stängt eller man kan göra det överhuvudtaget, eller allt på en gång eller så. Det är också sådana tekniska systemfrågor som jag har svårt att svara riktigt på.

I väntan på en lokal katalog som kan hantera icke-latinsk skrift på ett tillfredställande sätt experimenterar Malmö stadsbibliotek med att lägga in kyrillisk text i databasen, men detta sker efter en lokalt utarbetad lösning och kommer inte att stödja sökning med originalskrift, bara representation i OPAC:en (Informant 6).

Även informant 5 är inblandad i detta experiment och anser att detta första steg mot en flerspråkig katalog är ett bidrag till det rådande mångfaldsåret, men att man på sikt hoppas kunna göra alla språken sökbara. Hon menar dock att detta är en väldigt dyr och komplicerad process (Informant 5).

### **Policy, praxis och organisation**

Under denna rubrik har vi samlat resultat som berör rutiner och policys kring katalogisering och hantering av material med icke-latinsk skrift samt om bibliotekens organisation och samarbete mellan olika personalgrupper.

#### *Rutiner för katalogisering och övrig hantering*

På biblioteken vid Lunds universitet utnyttjas forskare och lärare i stor utsträckning för hjälp med selektion och förvärv vad gäller material på främmande språk då bibliotekarierna oftast inte har tillräckligt omfattande ämneskunskaper för att bedöma litteratur och tidskrifter. Kontaktbibliotekarier knutna till de olika institutionerna bevakar sitt ämnesområde i samråd med lärarna och forskarna via Internet, dagspress och information från förlag. Vem som helst kan föreslå böcker till inköp, men det mesta av tipsen och önskemålen kommer från lärare och forskare.

På Malmö stadsbibliotek ser det annorlunda ut. Det finns inga specialister i form av lärare eller forskare. Sambindningslistan är det traditionella hjälpmedlet för folkbiblioteken för att hålla sig uppdaterade vad gäller nya medier, men på Malmö stadsbibliotek försöker man gå ifrån denna och då speciellt när det gäller litteratur på främmande språk. Man samarbetar med internationella bokagenter som Aagot i Köpenhamn och utnyttjar informella kontakter i form av tips från användare, anställda på biblioteket med olika språkkunskaper samt kontakter med experter. En av dessa experter är en lärare i arabisk litteratur vid Lunds universitet som köper med sig böcker från resor i arabisktalande länder på uppdrag av Malmö stadsbibliotek och även hjälper till med transkriberingen av titlarna. Många av dessa kontakter är som nämnt informella och flera informanter påpekade under intervjuerna vikten av att kunna lita på den som har språkkunskaperna vad gäller urval av litteratur eftersom det inte är möjligt för dem att läsa materialet (Informant 5).

Både biblioteken vid Lunds universitet och biblioteken i Malmö använder sig av transkribering när de katalogiserar och hanterar material med icke-latinsk skrift, dels på grund av att bibliotekssystemen inte stöder andra teckenuppsättningar än den latinska och dels på grund av praktiska hanteringsskäl, till exempel etiketter med författare och titel transkriberat för att underlätta för dem som inte kan läsa på bokryggen. Majoriteten av de anställda på biblioteken har inte tillräcklig språkkompetens för att ställa upp böcker på alla språk utan att ha en transkriberad titel och författare utsatt på boken. På språk- och litteraturcentrums bibliotek har man dock valt att sluta med etiketter på utsidan av böckerna och nöjer sig med att skriva ut transkriberingen på titelsidan. Detta dels för att spara tid och arbete men också för att man inte anser att det behövs eftersom det är ett specialiserat språkbibliotek, enligt informanten. Varje skriftsystem har dessutom sitt eget sätt att rangordna tecken, men både i Lund och Malmö har man valt att konsekvent ställa upp materialet efter det latinska alfabetet på den transkriberade författaren alternativt titeln. På Malmö stadsbibliotek står all litteratur på främmande språk, inklusive facklitteratur, på skönlitteraturavdelningen. Det enda undantaget är kockböcker som står på en särskild avdelning som heter "all världens mat". Fackmaterial på skolspråken (engelska, franska, tyska och i viss mån spanska) står på fackavdelningarna.

På Asienbiblioteket har man valt att konsekvent följa SAB-systemet när det gäller uppställning av böcker, men har upprättat separata avdelningar för den kinesiska och den japanska samlingen.

På ett bibliotek med material på många olika språk är det inte möjligt för alla anställda att kunna alla språk. De flesta av informanterna tycker att transkriberingen fungerar bra för bibliotekens behov, men påpekar också problemen med transkribering vad gäller sökning av litteratur.

Så här långt har det fungerat ganska bra. Det som är problemet är ju att det inte är sökbart med tecken i databaserna. Det är ju det stora problemet, just för att transkriberingen aldrig är hundra procentig. Eller att, som med kinesiskan, att där är många tecken som skrivs med samma bokstäver eller transkriberas likartat kan man säga. Den enda skillnaden är då tonmarkeringarna, vilket man inte alltid skriver ut och inte ens dem skulle hjälpa en ändå utan det är enbart om man ser det specifika tecknet som man kan avgöra betydelsen. Samtidigt, om man ser transkriptionen i ett sammanhang så kan man ändå gissa sig till vilka, vilken betydelse, har jag förstått, jag kan ju inte det själv.

*Informant 1*

På Malmö stadsbibliotek försöker man komma runt problemet med sökbarhet genom att lägga upp listor på litteratur på främmande språk på respektive språks skrift som kan nås genom en länk på katalogsidan. I Millennium finns också möjligheten att lägga in bilder på tecken i katalogposter, dessa är dock inte sökbara. Detta har genomförts för material på kinesiska och ska även genomföras för material på ryska. Det är författare och titel som visas på detta sätt.

På biblioteken vid Lunds universitet och på stadsbiblioteket i Malmö använder man sig av transkriberingsscheman för att motverka att material inte kan återfinnas på grund av att två katalogisatörer har transkriberat t ex samma författares namn på olika sätt. Detta

har dock hänt i Malmö stadsbiblioteks katalog enligt en av respondenterna eftersom de som katalogiserar på stadsdelsbiblioteken inte använder sig av ett transkriberingsschema. Det har hänt att samma titel blivit transkriberad på olika sätt vilket gör att det blir väldigt svårt att hitta titeln i katalogen. Malmö stadsbibliotek har dessutom det ytterligare problemet att Libris använder ett annat transkriberingssystem än folkbiblioteken, vilket är problematiskt när katalogisatorerna importerar filer till det lokala systemet. Libris använder sig av den svenska forskningsbiblioteksmodellen för transkribering men trots det kan det ändå finnas fall när katalogisatorer tycker olika och det förekommer också en diskussion om huruvida forskningsbiblioteken ska ansluta sig till folkbibliotekens transkriberingsmodell eller inte.

När det gäller katalogisering, och då främst av material på främmande språk, uttryckte en av informanterna ett stort intresse för att Malmö stadsbibliotek skulle få möjlighet att katalogisera i Libris och inte bara knyta exemplarposter till redan befintliga poster. Samma informant ansåg också att detta kunde ge flera fördelar, till exempel att bättre utnyttja de katalogisatorer som har speciell språkkunskap genom att de katalogiserar material på ett visst språk åt alla bibliotek som är med i Libris. Ett samarbete mellan internationella biblioteket i Stockholm och Malmö stadsbibliotek med utbyte av poster har diskuterats men det förhindrades av att biblioteken hade olika bibliotekssystem och olika huvudmän. Det är däremot ett samarbete på gång med Göteborgs stadsbibliotek kring detta enligt en informant. Malmö stadsbibliotek importerar många poster från Libris, vilket kan leda till problem eftersom katalogisatorerna i Libris använder sig av ett annat transkriberingssystem. Posten måste då anpassas efter folkbiblioteksmodellen för att det inte ska uppstå problem i den lokala katalogen.

### *Språkkompetens hos personalen*

Alla informanter är överens om att språkkunskaper hos all personal på biblioteket är viktigt, men framförallt hos de i personalen som köper in och katalogiserar material på främmande språk.

Det tror jag är en önskan egentligen att man kan ha fler med specialkunskaper. Det underlättar så. Annars så finns, så kommer boksamlingarna att vara bara delvis tillgängliga. Och det finns ju andra bibliotek inom Lunds biblioteksnätverk där man har ämneskunniga och till och med disputerade bibliotekarier. Så det är en het önskan att, och det har vi ju diskuterat tidigare också, att när nästa bibliotekarie anställs så måste absolut vara en med språkkunskaper, för att det är viktigt. Det kan ju också avlasta förvärv och allting sånt, och avlastar ju då forskarna till viss del, med åtminstone det praktiska förvärvet att, för selektion kommer de då naturligtvis att jobba med ändå, men det praktiska förvärvet kan man ju bara göra om man faktiskt behärskar språket i fråga.

*Informant 1*

En annan sak som informanterna också är överens om är hur svårt det är att hitta utbildade bibliotekarier med en språkkompetens som går utanför skolspråken. I Lund tar man hjälp av forskare och lärare på språkinstitutionerna för att lösa problemet. ”Samtidigt så jobbar man ju i en miljö där man jobbar nära forskare och lärare, så har man ju ändå tillgång till deras kunskap och erfarenhet ändå. De är ju aldrig längre bort än en fråga egentligen. Det gör ju också att man så sakta tillägnar sig själv en hel del kunskaper” (Informant 1). I Lund är det heller inte ovanligt med bibliotekarier som är utbildade inom ämnet för det bibliotek där de arbetar.

På Malmö stadsbibliotek finns inga akademiska experter inkorporerade i verksamheten på samma sätt som vid Lunds universitets bibliotek. Därför blir det viktigt att i så stor utsträckning som möjligt ta till vara på den språkkunskap som personal och andra kontakter besitter. Dessa kontakter kan vara anställda tolkar, bibliotekarier från andra bibliotek, användare, lärare eller forskare som bibliotekarierna har kontakt med på en mer informell basis och inte genom till exempel kontaktbibliotekarier som på Lunds universitet. På stadsdelsbiblioteket i Rosengård engageras till exempel användare med språkkunskaper i katalogiseringsarbetet och med urvalet av litteratur.

[...] när det gäller arabiska så har jag tur, jag har en låntagare som praktiserat här lite grann tidigare, han är doktorand i historia från Irak och kan mycket litteratur, så han hjälper mig att transkribera böckerna. Han hjälper mig med persiska också, så att det..., när det gäller hanteringen av arabiskan så litat jag på honom då, han hjälper mig [...]

*Informant 4*

På stadsbiblioteket arbetar en assistent som kan arabiska som ibland får göra inköp av litteratur till biblioteket och katalogiserar litteratur på arabiska. Ibland har det också kommit personal till stadsbiblioteket genom arbetsmarknadsåtgärder som kunnat bland annat kinesiska och thailändska. Som tidigare nämnt köper även en professor i arabisk litteratur från Lund in litteratur på uppdrag av stadsbiblioteket.

På frågan om de ser en utveckling där specialister inom olika språk ersätter generalister får vi svaret att det är blandat. De är överens om att på mindre bibliotek är det inte praktiskt med alltför specialiserade bibliotekarier, utan där är det viktigt att ha personal som är allsidig.

Det beror ju också på, är man två personer på ett bibliotek så får man ju faktiskt vara generalist, det är ju inte mycket att be för. Sitter man på ett bibliotek med hundra personer som UB [Universitetsbiblioteket] så kan man ju kosta på sig att vara mer specialiserad.

*Informant 3*

### *Inställning till och policy för romanisering vid katalogisering*

Vad gäller inställningen till romanisering vid katalogisering är informanterna överens om att det fungerar för bibliotekens behov, men att det också innebär en del komplikationer. Ett stort problem är graden av standardisering vid romaniseringen, vilka transkriberingsscheman som används. På Malmö stadsbibliotek använder man sig av folkbibliotekens transkriberingsschema, men på stadsdelsbiblioteket i Rosengård använder man inte något transkriberingsschema överhuvudtaget utan det är upp till varje katalogisatör att själv romanisera som de anser lämpligt. Detta har lett till att samma titel eller författare har fått olika transkriberingar, vilket gör det svårt att hitta alla exemplar av en titel i katalogen. I Lund finns det transkriberingsscheman som är gemensamma, men kunskapen i språk är utspridda till de olika institutionerna och det finns ingen central styrning eller kontroll av hur de används i praktiken. Informanten från biblioteksdirektionen menade att transkriberingsschemana och praxisen kring romanisering befinner sig på ett primitivt stadium. ”[...] så är det de här tabellerna, om ni tittar på dem här så är det ju, det är gamla scannade sidor. Och där har katalogisatörerna suttit och mödosamt arbetat sig igenom. [...] Så det är ju på ett

primitivt stadium förvisso. [...] De är det enda hjälpmedel som de har haft hittills” (Informant 2).

Och även om katalogisatorerna på ett bibliotek använder sig av samma transkriberingschema är det ändå problematiskt menar en del av informanterna. En transkribering är aldrig hundra procentig.

Det är ju det stora problemet, just för att transkriberingen aldrig är hundra procentig. Eller att, som med kinesiskan, att där är många tecken som skrivs med samma bokstäver eller transkriberas likartat kan man säga. Den enda skillnaden är då tonmarkeringarna, vilket man inte alltid skriver ut och inte ens dom skulle hjälpa en ändå utan det är enbart om man ser det specifika tecknet som man kan avgöra betydelsen.

*Informant 1*

Ett problem för användaren som söker i katalogen är att de inte alltid vet vilken praxis som ett bibliotek har när det gäller romanisering av olika språk och de vet inte vilken stavning som biblioteket valt för en specifik författare eller titel. Ett sätt att komma runt detta är hänvisningar i den bibliografiska posten, där den som söker efter till exempel en viss författare med en annan stavning än som är standard i bibliotekssystemet får hänvisningar till den stavning som är ”rätt”. På språk- och litteraturcentrums bibliotek ser inte informant 3 detta som ett stort problem eftersom de flesta av dem som använder litteraturen på främmande språk är vana vid transkriberingen eller använder sig inte av katalogen för att hitta de böcker de söker. Dock menar informanten att transkribering är ett hinder när det gäller import av katalogposter från andra länder eftersom de har helt olika transkriberingssystem. Ett liknande problem finns också på Malmö stadsbibliotek när det gäller import av katalogposter från Libris.

Det kan bli lite problem när man tar poster från Libris så tillvida att vi, inte har samma translittereringspraxis på folkbiblioteken som de har. Och det kan klydda till det ganska mycket, i synnerhet då när det är någon katalogisator som normalt inte hanterar, material på icke-kyrilliskt alfabet, som får en sån här post och tror att det är bara att köra på. Då blir det ju inte så bra i fråga om index och sådana saker så...

*Informant 6*

När det gäller användare som har ett annat modersmål och som söker material i katalogen tror informanten från språk- och litteraturcentrums bibliotek att det är en mer komplicerad situation för denne användare än för en användare med svenska som modersmål.

Det vet jag om man ser sitt eget språk i transkriberat. Jag ser ibland, man har japanska böcker som är översatta, alltså svenska författare översatta till japanska, så är de ju oigenkännliga. Astrid Lindgren heter då Rinduguren eller någonting sånt, och vem tänker på att söka på det när man är svensk? Jag misstänker att det är samma sak. Rinduguren... hur ska man komma på det? När man söker efter Lindgren så... (skratt)

*Informant 3*

Informanten från stadsdelsbiblioteket i Rosengård anser att detta är ett väldigt stort problem som kommer upp väldigt ofta i arbetet på biblioteket och leder till frustration för både bibliotekarierna och användarna. Användare som har ett annat modersmål har ofta svårt att transkribera till exempel arabiska namn och titlar. Ofta ber de då bibliotekarierna om hjälp att transkribera och om varken bibliotekarien eller användaren lyckas komma fram till "rätt" transkribering så tillgriper användarna istället en browsing-strategi där de helt enkelt söker igenom hyllorna i biblioteket för att hitta det de söker. Informanten ansåg utifrån detta att transkribering inte är en riktigt bra lösning, men sade sig inte veta om något annat sätt att hantera materialet på. Informanten nämnde dock att möjligheten att söka med icke-latinska tecken i katalogen är något som låntagare efterfrågar. Även en informant från Malmö stadsbibliotek uttryckte en liknande åsikt som informanten från Rosengårds stadsdelsbibliotek vad gäller romanisering:

Ja, [...] för våra biblioteksbehov går det väl bra att göra så, men frågan är hur användarna upplever det, och det är ju det som är problemet [...] med det. Det är, vi kan ju följa de regler som BTJ, eller deras standard, och köra den linjen och skapa överensstämmelse mellan katalog och objekt så att säga. Men det säger ju inte en användare så mycket. Jag tänker mig att, det kan man ju tänka sig själv att det skulle vara väldigt frustrerande att se svensk text transkriberad till arabiska eller så va, om man, så länge man inte har ganska goda kunskaper i det nya språket också, så det... Men, som sagt, problemet vi har mött det är väl då när vi har importerat poster som inte har följt det system som vi tidigare har tillämpat.

*Informant 6*

På frågan om de skulle kunna tänka sig att helt gå ifrån romanisering i biblioteket svarar alla informanter att det nog inte är möjligt med tanke på att materialet måste kunna hanteras av personal som inte behärskar språket. Däremot är informanterna väldigt positivt inställda till möjligheten att ha poster med både originaltecken och romaniserad form, framförallt om båda varianterna är sökbara. Däremot är informanterna mer skeptiska till om detta skulle kunna genomföras i verkligheten på grund av begränsningar i ekonomi och det arbete det skulle krävas för att genomföras fullt ut i katalogen.

Ja, rent tekniskt är det möjligt, men det är väl omöjligt i praktiken, anta jag. Därför att vi har ju så himla mycket gamla böcker som är transkriberade och katalogiserade så det skulle bli, då fick man [...] Jag ser det er som ett mervärde att man plockar ut de särskilt intressanta, kanske aktuella böckerna, och ser till så att de blir snygga. Det är ett gigantiskt projekt det här att, om man ska gå bakåt och göra det retrospektivt. Jag tror knappt det. Men det är klart, man vet aldrig, hur tekniken utvecklas.

*Informant 2*

### *Kommunikation mellan personalgrupper*

Kunskapen om unicode och dess användning i bibliotekssystem är väldigt varierande mellan de olika informanterna. Vissa är väldigt medvetna om möjligheterna som unicode innebär, andra har inte ens hört talas om det. De informanter som har positioner som ansvariga för bibliotekssystemet eller som arbetat mycket med denna typ av frågor är bättre informerade och informanterna från Lund var överlag mer insatta i fråga om unicode. Enligt informanterna är de flesta anställda på biblioteken i Lunds



universitetsnätverk informerade om unicodeanpassningar i Lovisa och Libris, och vad det kan innebära. Man informerar främst genom e-postlistor eftersom det är ca åttio personer som katalogiserar och de sitter utspridda på olika bibliotek i nätverket. Detta innebär att det är svårt att organisera möten med alla katalogisatorerna och få ut information på det sättet. Frågor kring unicode verkar främst ha diskuterats bland personal som på något sätt arbetar med material på främmande språk.

Institutionsbiblioteken är relativt självständiga, vilket kan bli ett problem i dessa sammanhang eftersom det kan leda till att det utvecklas olika rutiner på olika bibliotek.

Ett bibliotekssystem mår bäst av att det finns få människor som registrerar. Det ska finnas så många användare som möjligt men så få människor som registrerar. Då är det enkelt att sköta. Det blir jobbigare ju fler det är som registrerar, och här i Lund är det ju nästan extremt där biblioteken alltid gjort precis som de velat. [...] Jag kan tänka mig att det är mycket enklare på mindre universitetsorter, i Kristianstad och Växjö och sådana ställen, därför att där har man gått ut med en stark centralmakt, så att säga, från början. Men här är det mer kaotiskt i Lund, det är det verkligen.

*Informant 2*

När så många katalogisatorer katalogiserar i samma system blir det viktigt att informationsutbytet mellan katalogisatorerna och mellan katalogisatorer och tekniskt ansvariga för katalogen fungerar på ett bra sätt. Men detta hindras enligt en informant dels av att allt är utspritt på olika institutioner och på grund av bibliotekariers rädsla att avslöja eventuell teknisk okunskap.

Det är så svårt det här, så någon, det hade ju varit underbart om någon kliva fram och göra en enkel, begriplig sammanfattning av hela situationen. Både rent tekniskt men även med andra aspekter. Som det är idag så går alla omkring och är positiva, men ingen vill säga något, som jag sa, av rädsla för att göra bort sig, och visa sin egen okunskap. En del, de bryr sig då inte. Men, alltså det är ju också som jag har sagt, dem som inte kommer i kontakt med de här språken de har naturligtvis väldigt begränsat intresse.

*Informant 2*

Informant 3 uttrycker sig så här:

Nja, det beror nog på, det beror på i vilken mån och vilka eldsjälur det blir som arbetar med det. Det hänger ofta på det, att någon tar i det och gör någonting av det. Det vore väl konstigt om inte... om någon börjar så. [...] Egentligen är unicode början på någonting större än man kanske föreställer sig.

*Informant 3*

Även i Malmö är kommunikationen mellan olika grupper i personalen oklar framförallt mellan stadsbiblioteket och stadsdelsbiblioteken. Stadsdelsbiblioteken katalogiserar inte material på svenska, men däremot material på främmande språk. När informanten från Rosengårds stadsdelsbibliotek tillfrågades om han kände till den kommande uppgraderingen av Libris svarade han att han läst ett e-post om det, men inte fått någon mer information.

## Attityder & inställningar

Temat attityder och inställningar sammanfattar hur informanterna ser på införande av originalskrift i bibliotekskatalogen, vad en unicodeanpassning skulle innebära för det aktuella biblioteket, attityder till standardisering i allmänhet samt åsikter kring Libris kommande implementering av unicode.

### *Inställning till införande av icke-latinsk skrift i katalogen*

Informant 1 har stora förhoppningar om att en fullständig unicodeanpassning av Libris skulle möjliggöra användning av originalspråk i bibliotekskatalogen:

[...] jag tycker det är fantastiskt, jättebra att det kommer till stånd äntligen. Att vi kan använda unicode. Vi var väldigt glada när vi här konverterade till en nyare version av vår lokala databas och det visade sig att vi kunde använda unicode och det var verkligen underbart och sen så tog det stopp i Libris och sen fick vi ju packa ihop alla de drömmarna och tankarna då, men är det härligt, nu är det verkligen på gång.

Informant 1 anser att stödet för att möjliggöra katalogisering med icke-latinsk skrift är mycket stort från katalogansvarig vid Lunds universitets bibliotek, och möjligheterna har diskuterats i stor omfattning på biblioteket:

Jag tror det kommer bli jättebra när väl allt detta kommer igång och fungerar så kommer det att göra materialet mer hanterbart och också synligt ju, såklart. Det kommer ju då att synas, framförallt bättre för dom som kan språken. Just för att det är svårt att söka på bara transkribering. Och sen är det väl möjligt att också andra, jag menar tredje man där kan vara många som är intresserade utav att låna och läsa material på biblioteket, men som kanske då inte är så bevandrade i transkribering och då kan de använda tecknen. Det är ju så att på Asienbiblioteket finns en hel del material, i alla fall när det gäller Kina, som inte finns i Kina längre, som är bortplockat och censurerat sedan länge. Det finns ju på det biblioteket och det finns ju på likartade bibliotek överallt i världen i övrigt utanför Kina, men det är väldigt intressant för många kineser då som kommer ut och kan hitta det här som de aldrig kan hitta hemma.

Hon anser sig se en klar påtryckning inte bara från bibliotekarier utan även användare som vill kunna söka med olika specialtecken och språk i databasen. På frågan om informanten anser att sökmöjlighet på originalskrift skulle behövas på biblioteket där hon jobbar svarar informant 1 så här:

Ja, det tycker jag absolut är, med tanke på att det är en så stark önskan att fortsätta att bygga upp samlingarna på Asienbiblioteket, och framförallt då i kinesiska och japanska så är det ju mycket önskvärt, och anledningen är ju det som jag har nämnt innan att man får en mer precis sökning gjord. [...] Jag tror att träffarna blir mer precisa än om man söker i transkriberad form. Man kan uttrycka sig mer exakt.

Informanten anser att inläggning av originalskrift i katalogen skulle vara värt den extra ekonomiska belastningen på grund av den respons hon fått från forskare och kollegor på olika europeiska bibliotek som experimenterat med olika metoder för att lägga in originalskrift i bibliotekskataloger. Hon berättar dock att många av dessa experiment har varit lokala lösningar då de aktuella biblioteken, exempelvis Cambridge, inte har varit unicodekompatibla, och därför istället arbetat med att klistra in bilder i de bibliografiska

posterna eller gjort en länk i posten som leder till en post från CHINA-MARC<sup>13</sup>, och på så sett ger användarna tillgång till originalschrift (Informant 1).

Informant 2 är hoppfull inför framtida möjligheter med originalschrift i Libris och tror att de berörda biblioteken skulle kunna dra stor nytta av detta. Han tror dock att det inte är alla som kommer vara intresserade av en sådan här utveckling, då det kommer innebära en del kostnader och merarbete, vilket inte kommer tilltala de bibliotek inom Lunds universitets biblioteks nätverk som inte arbetar med material på icke-latinska språk.

På frågan om informanten anser att sökmöjlighet med originalschrift behövs på biblioteken svarade han:

Det är ju ett stort universitet, man har ju väldigt många gäststudenter. Och det är ett stort universitet som syns, katalogen besöks av besökare från hela världen, jag tror det. [...] Det borde väl rimligt sätt betyda mycket även för undervisningen och universitetet, inte bara för biblioteket.

Informant 3 är lite osäker på hur en unicodeanpassning av Libris skulle påverka hur hennes bibliotek arbetar med material på icke-latinska språk, men tror att det skulle underlätta arbetet i längden:

Ja, det kommer vi också få se när det händer. Det är alltid lite jobbigt när man får nya versioner innan datorn så att säga är omställd och allting fungerar. Det är många förhoppningar knutna till det, och under diskussionerna om transkribering under det senaste året har man hört ”ja men sen kommer unicode”, ”ja men då är det snart unicode, det kan ju vänta till dess”. Så jag antar, jag antar att det blir lättare. [...] Och hur det blir i unicode har jag inte fått någon kläm på. Jag har tittat i den där Karlsruhe virtual catalog, och sett att där dyker upp kyrillisk stil lite här och där nu. Jag har tänkt ”jaha, de har unicode, de skriver med kyrilliska alfabetet”. Det vore ju faktiskt det allra bästa om man tänker på det.

På frågan om ifall hon anser att sökmöjlighet på originalspråk behövs på hennes bibliotek svarar hon att det inte är nödvändigt men att det låter bra och antagligen skulle ge mer precisa träffar med färre tveksamheter (Informant 3).

På frågan om hon anser att katalogisering med originalschrift skulle vara värt den ekonomiska satsningen svarar informant 3 så här:

Ja, jag hoppas ju det. Det är alltid en avvägning. Men om det sparar tid i framtiden, och det gör det väl vid katalogisering, och det är arbetskostnader som är dyra, dyrare i längden än ren installering, jag tror det.

För informant 4 råder ingen tvekan om att sökmöjligheter på originalschrift skulle vara till stor hjälp för användarna på hans bibliotek. Han anser att det absolut skulle vara värt en ekonomisk satsning samt merarbete vid katalogisering för att möjliggöra sökning på originalspråk. Han menar att användarna sällan hittar vad de söker efter med hjälp av bibliotekskatalogen då transkriberingen är ett hinder, eftersom de inte vet hur det aktuella språket transkriberas i katalogen. Sökning direkt på originalspråk hade underlättat processen avsevärt (Informant 4).

---

<sup>13</sup> CHINA-MARC är en kinesisk version av föregångaren USMARC, som idag är MARC21, det mest använda MARC-formatet, även om det fortfarande finns en rad nationella versioner, till exempel i Kina.

Informant 5 är positiv till originalskrift i katalogen, men ser ekonomin som ett stort hinder:

Ja, det är det, för att dels så kostar det att anställa folk som ska göra det och dels så kostar det att tekniskt göra det också, det är system som skall utvecklas och så och det är samarbete mellan dom som är systemkunniga, katalogiseringskunniga, inköpskunniga, ja det är många om man skall göra det bra tror jag.

Hon menar att arbetet med att göra den lokala katalogen unicodekompatibel ekonomiskt sett ligger långt bak i kön för tillfället, men anser dock att om det funnits ekonomiska möjligheter hade arbetet redan varit igång (Informant 5). På frågan om hon tror att en fullständig unicodeanpassning av Libris skulle påverka hur biblioteket arbetar med material med icke-latinsk skrift svarar hon så här:

Definitivt, definitivt. Vi skulle ju kunna använda oss av Libris på ett helt annat sätt och kanske också, som jag säger, komma in och bli, på vissa språk, kanske kunna få katalogisera då<sup>14</sup>. Det hade ju kunnat gå då, om vi har den kompetensen här.

Informant 5 anser att sökmöjlighet med originalskrift behövs på biblioteket, och tror på att samarbete mellan olika institutioner och bibliotek skulle underlätta införandet:

Ja, klart att det vore ju väldigt önskvärt. Det vore det ju. Och jag tror ju att utvecklingen går därhän. Det tycker jag är positivt. Men jag tror att det kommer att ta tid, det är ju en stor process. Det är förenat med stora kostnader och så. Men som sagt, jag tror på samarbete där, om man satsat tillräckligt med resurser hos något eller några som kan göra det, grundarbetet, som sen andra kan tillgodogöra sig också.

Hon anser att en fullständig unicodeanpassning skulle förändra hanteringen av material med icke-latinsk skrift, men är osäker på hur, då en unicodeanpassning av den lokala katalogen ligger i framtiden och man inte diskuterat frågan i detalj ännu. Hon tror dock att man skulle försöka ta tillvara de möjligheter sökning med originalskrift skulle innebära för verksamheten (Informant 6).

### *Inställning till standardisering*

Informant 1 är mycket positivt inställd till standardisering som exempelvis MARC21-formatet:

Det finns ju ingen anledning att uppfinna hjulet när de redan gjort det minst en gång innan. Dubbelarbete är ju någonting man ska försöka undvika tycker jag. Och det är ju just möjligheten idag med Internet att man kan gå in i varandras kataloger och titta, så kan man hitta informationen där. Man behöver inte göra allt precis från början. Naturligtvis får man ofta anpassa posterna till sina egna standarder, lokala regler och så, som det ju ändå alltid finns lite anvisningar för, men det mesta jobbet är ju ändå gjort. Och det underlättar ju arbetet och kanske snabbheten.

---

<sup>14</sup> Malmö stadsbibliotek får inte katalogisera eller städa i katalogposter i Libris, utan bara lägga in sitt bestånd. Detta är något som man skulle vilja ändra på för att få tillstånd att även göra katalogposter i Libris.

Informanten själv är positiv till standardisering men berättar dock om andra bibliotek hon känner till där man arbetat med egna lösningar som exempel vis att ha två olika kataloger för de olika språken samt att använda sig av bilder istället för text i katalogen.

Informant 1: Jag har ju sett hur otroligt mycket arbete de har lagt ner i att försöka hantera, framförallt kinesiskt tryck är det, att få det hanterbart i databasen, och sökbart och har hittat många olika lösningar, bra eller dåliga kan ju inte jag uttala mig om men, där det har kostat väldigt mycket arbete och energi att få lösningar till det, så att kan man lösa det på nåt sånt sätt som med unicode så känns det ändå som överkomligt vad det bär med sig i kostnader jämfört med vad de redan har lagt ner mycket arbete.

Intervjuare: Så de har inte använt sig av unicode i sina system?

Informant 1: Nej, nej. De har gjort egna databaser för att deras, hur ska jag säga, normala bibliotekskatalog har man inte kunnat ha kompatibel, för dom har inte haft tillgång till unicode än, utan man har gjort parallella lösningar och försökt vidarekopplingar. I Cambridge till exempel har man också Voyager, som ju är Librissystem och där vet jag att man har jobbat med det, i posterna att man gör bilder på tecknen för att kunna visa det både i transkriberad form och med tecken friständigt, och sen det man har gjort för att ge låntagarna, användarna en bättre bild utav det hela eller visa informationen på ett bättre sätt, att man då har köpt poster från Kina som man visar, man har alltså en länk till CHINA-MARC där man då visar den fullständiga posten.

*Informant 1*

Även informant 2 är mycket positiv till internationella standarder då det medför utbyte av poster mellan biblioteken. Han anser att innan de flesta biblioteken gick över till MARC21 var det väldigt rörigt med många olika MARC-format, vilka hindrade utbytet av poster. Att man idag kan hämta poster från exempelvis Library of Congress och British Library underlättar arbetet. Han tar dock upp att det finns en del kritik mot MARC inom branschen. Det sägs vara för utförligt och svårarbetat:

MARC är ju tungrott och det kräver rätt omfattande intellektuell insats av den som ägnar sig åt det, alltså kostar det för mycket pengar. Därför ser vi den här utvecklingen, med databaser, med tidskriftslistor och liknande. Men där vill man säkert inte börja peta i olika tecken och så vidare därför att där ska det gå fort och där ska man tjäna pengar, eller spara pengar rättare sagt.

*Informant 2*

Men han säger även att han själv inte ser några egentliga konkurrenter till MARC-formatet då han anser det vara överlägset de andra systemen vad gäller exempelvis teckenuppsättning. Informanten tror inte att bibliotekets användare i första hand har påverkat att biblioteken eftersträvar standardisering, utan att det är bibliotekarierna som tryckt på i den frågan. När vi undrar om detta även gäller standardisering genom unicode för representation av tecken svarar han att påtryckningar om detta skulle kunna komma även från användarna.

Där kan jag ju gott tänka mig att användarna kan uppfattas som en påtryckningsgrupp. Det här med att kunna visa tecknen i original så att säga det hade ju varit, ja, det hade ju verkligen befrämjat internationella kontakter, det kan ju komma kinesiska studenter hit och söka i katalogen med kinesiska tecken.

*Informant 2*

Informant 3 anser inte att internationell standardisering är livsnödvändigt på hennes bibliotek, men att det kan vara praktiskt inom vissa områden. Hon anser att datoriseringen har påverkat i vilken utsträckning man måste hålla sig till standarder, då man tidigare arbetade med katalogkort som bara berörde det egna biblioteket. ”Då hade man ju ganska fria händer. Man kunde tycka, nu kan man inte tycka lika mycket, nu bör man veta. Och man får resten av Sverige på sig om man avviker från normen.”

Informant 3 tar även upp problemet med att det finns olika transkriptionsstandarder, vilket påverkar hur en post ser ut i Libris. Ett bibliotek kan ha bestämt sig för en transkribering medan ett annat vill ha en annan transkribering, folkbiblioteken har en modell och forskningsbiblioteken har en annan. Hon menar att hon ibland föredrar när en bok inte redan finns i Libris eftersom hon då slipper ta kontakt med det bibliotek som redan lagt in den, men använt ett annat system än det hon själv vill använda. Hon anser dock trots allt att standardisering i stort sett underlättar arbetet då det finns en norm att följa.

Informant 4 anser att standardisering utan tvekan är det bästa både för bibliotekarierna själva och bibliotekets användare då det förenklar arbetet. Även informant 5 är mycket positiv till MARC-formatet:

Det är bara positivt. Det är absolut, vi var bland de första som också som hoppade på det här MARC 21 när det kom. [...] Det är ju väldigt viktigt inom den här branschen att man faktiskt försöker hålla sig till standards. Det är ett alldeles för litet land för att hålla på och peta med våra olika utvecklingar på olika nivåer så att säga. Vi måste bli mycket bättre på att plocka av varann, dela med oss av vår katalogkunskap. [...] Egentligen får man inte ägna sig åt sånt [egna lösningar] för att det blir bara besvärligt, du skall ju föra kunskapen vidare, du ska utbilda folk i massa konstigheter, och du kan så att säga inte direkt börja samarbeta med någon för ni pratar inte samma språk, då blir det ju som att transkribera, även på den biten.

*Informant 5*

Hon har dock inte känt att bibliotekets användare har kommit med påtryckningar för att biblioteket ska anpassas till mer internationella standarder: ”Det tror jag inte de är medvetna om”.

Informant 6 tycker att internationella standarder är att föredra framför lokala lösningar vad gäller bibliotekskatalogen. Hon anser att hennes bibliotek haft mycket stor glädje av att byta format till MARC21 från UKMARC som man tidigare använde. Hon menar att den absoluta majoriteten av de poster som importeras på hennes bibliotek är MARC-poster. Vid enstaka tillfällen importeras poster från ett danskt bibliotekssystem som inte använder MARC-formatet, vilket kan ställa till en del problem i katalogen om katalogisatören inte uppmärksammar skillnaden och redigerar posten. Det ligger exempelvis skillnader i indikatorer och delfält i titelfältet 245.

#### *Funderingar kring varför Libris beslutat att göra en unicodeanpassning*

Informant 1 är osäker på vad som gjort att man i detta skede fattat beslutet men hoppas att det är så att Libris äntligen lyssnat på de bibliotekarier som arbetar med den här typen av material som har drivit frågan. Informant 2 tror att katalogisatörerna runt om i landet

tröttnat på den mycket klumpiga hanteringen av diakriter i Librisklienten och klagat på detta till Libris.

Informant 3 tror att det kan ha skett påtryckningar på bibliotekarierna från forskare och professorer inom exempelvis östasiatiska och slaviska språk som ansett att språken ser konstiga ut i katalogen, och att bibliotekarierna sedan i sin tur fört fram kritiken till Libris. Informant 4 kan tänka sig att internationella biblioteket i Stockholm kan ha tryckt på för en förändring, och att de i sin tur fått påtryckningar från användare. Mångfaldsåret är en teori som informant 5 lägger fram och informant 6 tror att det är något som ligger i tiden nu med internationalisering och globalisering på många områden samt att den tekniska möjligheten är något som kommit först nu.

### Libris' uppgradering till unicode

Under vecka 21 i år kommer Libris att byta teckenuppsättning till unicode. Enligt Libris' hemsida kommer inte uppgraderingen att förändra exporten av poster till lokala system (Libris uppgradering – katalogiseringsklient). Förändringen kommer i huvudsak att underlätta arbetet med registrering av specialtecken och diakriter då flertalet tecken som finns på tangentbordet nu kommer kunna användas i katalogiseringsklienten samt att tangentbordets konstruktion av diakriter kommer att fungera. Export av poster till lokala system kommer att ske med teckenuppsättningen unicodes UTF-8, men även med Latin-1 (ISO 8859-1) (Libris katalogiseringsklient – nyheter). Dessvärre för våra syften med den här uppsatsen har man dock på Libris beslutat att inte uppmuntra till användandet av icke-latinska skrivtecken för katalogisering trots unicodeanpassningen:

Kyrillisk, arabisk och annan icke-latinsk originalskrift skall enligt MARC21-standard registreras i länkfält 880, med transkriberad/translittererad uppgift i de sammanlänkade normalfälten (245, 260 ...). Länkfält 880 **kan** användas aktivt, men rekommenderas **inte** eftersom LIBRIS tills vidare inte kommer att erbjuda något stöd för denna metodik.

*Libris katalogiseringsklient – nyheter*

Det finns dock ett kryphål då grekiska tecken kommer att kunna användas samt att poster som importerats med ett 880-fält innehållande icke-latinsk skrift från exempelvis Library of Congress kommer att kunna användas: ”Grekiska tecken kan dock registreras i den mån de förekommer i matematiska eller kemiska formler. Man kan också låta de icke-latinska tecken som förekommer i remote-importerade posters fält 880 ligga kvar oförändrade” (Libris katalogiseringsklient – nyheter).

Vi har haft kontakt med ansvarig för bibliografiska frågor på Libris via e-post och ställde då frågan varför man valt att inte rekommendera användandet av icke-latinsk skrift. Han svarade:

De tecken som läggs in måste kunna hanteras även för t.ex. sökning, sortering och visning. De måste också fungera i alla de kringrutiner och kringssystem som byggts upp (LIBRIS webbsök, fjärrlån, nyförvärvslistor ...), och inte minst vid export till de lokala LIBRIS-bibliotekens egna katalogsystem. Vad gäller postexporten, så kommer vi även i fortsättningen att erbjuda MARC-poster i Latin-1 (ISO 8859-1) till de bibliotek som så önskar.<sup>15</sup>

<sup>15</sup> E-post från ansvarig för bibliografiska frågor på Libris

Vi frågade även om det finns några planer på att möjliggöra användandet av icke-latinsk skrift i Libris, varpå vi fick svaret att det inte finns några tidsplaner, men att han ser det som en självklar inriktning. Han menar dock att fokus nu ligger på övergången till ett nytt teckensystem med allt vad det innebär med omfattande arbete med anpassning till system och rutiner.<sup>16</sup>

Slutligen undrade vi varför man valt att anpassa sitt system till unicode, varpå vi fick svaret: ”Det enklaste svaret är kanske att vi eftersträvar tillämpning av standarder och interoperabilitet. Unicode är en generell teckenstandard.”<sup>17</sup>

---

<sup>16</sup> Ibid.

<sup>17</sup> Ibid.



## Analys

Vår analys grundar sig på resultatredovisningen i den empiriska undersökningen i samband med vår bakgrunds- och teorienomgång.

### Implementering av unicode

Enligt den undersökning som Laura Tull gjorde 2002 hade 6 av 11 bibliotekssystem implementerat unicode och fyra av de övriga planerade att göra detta inom kort. Detta visar på att en klar majoritet av företagen bakom dessa bibliotekssystem stödjer unicode som den nya internationella standarden för teckenuppsättning. Vad beträffar de tre bibliotekssystemen i vår undersökning har alla tre möjligheten att använda unicode som teckenuppsättning, men det är bara Virtua (Lovisa), som används vid Lunds universitets bibliotek, som i skrivande stund har fullt stöd även på lokal nivå, även om Voyager, Libris, är nu på gång med en uppgradering under våren.

De informanter i vår undersökning som arbetar på bibliotek inom Lunds universitets biblioteks nätverk har följaktligen tillgång till ett lokalt bibliotekssystem med full unicodeanpassning, vilket har utnyttjats av Informant 1 genom experiment med kinesiska och japanska tecken i den lokala katalogen. Även Informant 2 var delaktig i dessa experiment. Det som hindrat att experimenten fortsatt är det faktum att alla ändringar i den lokala katalogen skrivs över av eventuella uppdateringar av posterna i Libris. Hade det inte varit för att Libris skriver över de lokala posterna hade arbetet med att lägga in icke-latinsk skrift i Virtua på Asienbiblioteket troligtvis redan varit igång enligt Informant 1, något som vi tolkar som ett stort intresse för att lägga in originalspråk i katalogposterna då man aktivt arbetat med detta.

I Laura Tulls undersökning efterfrågades motivet till bytet av teckenuppsättning och som främsta orsak nämndes efterfrågan och konkurrenskraft. Antingen var implementeringen ett svar på kunders direkta efterfrågan eller ett led i expansion till länder med behov av mer omfattande teckenstöd. Även förändrade behov på den befintliga marknaden togs upp då stora folkförflyttningar äger rum över hela världen vilket fått till följd att länders befolkning förändras och därmed även befolkningens språkliga bakgrund. Vi frågade våra informanter vad de trodde var anledningen till att Libris nu valt att byta teckenuppsättning till unicode och svaren vi fick varierar från att bibliotekarierna tröttnat på begränsningar och klumpiga hanteringar vad gäller specialtecken i systemet till att internationalisering och globalisering i kombination med ny teknik ger nya möjligheter och krav på systemen. Även det faktum att det är mångkulturår togs upp som anledning.

Tull drar slutsatsen att allt fler företag kommer anpassa sina produkter till unicode standarden och detta kommer pressa de som inte gått över eftersom utbyte av poster kommer påverkas. Vi tycker oss se samma tendenser och håller med om att detta verkar vara trenden idag, något som vi anser att vi finner stöd för i de svar våra informanter gav.

Vad beträffar kännedom om dels unicode som system och Libris byte av teckenuppsättning var denna märkbart högre bland informanterna vid Lunds universitets bibliotek, vilket kan bero på att man där katalogiserar i Libris, vilket inte Malmö stadsbibliotek gör. På Malmö stadsbibliotek samt Rosengårds stadsdelsbibliotek kände två av informanterna till termen unicode men var inte insatta i när Libris uppgradering egentligen skulle ske. Den tredje visste att en förändring var på gång under våren men kände inte till begreppet unicode.

Att byta teckenuppsättning kräver eftertanke och planering. Oftast finns inte möjligheten att göra om en katalog från början, och arbetet måste därför synkroniseras med den befintliga katalogen under en övergångsperiod. Aliprand menar dock att detta är något som redan pågår vad gäller informationsutbyte mellan system med olika skrift. När ett system baserat på enbart latinska bokstäver tar emot data från ett multiskriftssystem finns två alternativ. Antingen tas informationen som inte kan användas av systemet emot och lagras så att den i framtiden kan användas, eller så tas den omedelbart bort från systemet (Aliprand 1992, sid. 115). Aliprand menar även att det kan bli svårt att konvertera mer massiva databaser, och att lösningen på detta är att mappa mellan två teckenuppsättningar i övergångsperioden (Aliprand 1992, s. 115).

Informant 6 uttrycker en osäkerhet angående alla tekniska detaljer kring implementering av unicode i katalogen. Hon ifrågasätter hantering med index samt hur den praktiska övergången till en annan teckenuppsättning ska gå till. ” Det är också sådana tekniska systemfrågor som jag har svårt att svara riktigt på” (Informant 6). Även Informant 3 är lite mer försiktig i sina lovord när det gäller införandet av unicode än våra andra informanter. Hon är skeptisk till att det kommer att göras i någon större utsträckning och är inte övertygad om att det är nödvändigt för bibliotekets verksamhet, även om han kan se fördelarna tvivlar hon på att det kommer satsas på en konvertering ur ekonomisk synpunkt. Även flera av de andra informanterna har uttryckt tveksamhet inför den ekonomiska aspekten av hantering av icke-latinsk skrift i bibliotekskatalogen.

### Standardisering – MARC-formatet och dess förutsättningar

Som ett led i arbetet med att implementera icke-latinsk skrift i MARC-formatet utarbetades i början av 1980-talet ett särskilt fält som fick koden 880, Alternate graphic information. Vi ville undersöka hur väl våra informanter kände till denna möjlighet och kom fram till att kännedom om 880-fältet var stor bland informanterna på Asienbiblioteket och Biblioteksdirektionen, medan de från Rosengårds stadsdelsbibliotek, Språk- och Litteraturcentrums bibliotek samt Malmö Stadsbibliotek hade begränsade om ens någon kännedom om fältet, trots stora samlingar på icke-latinska språk. Informanterna från Asienbiblioteket och Biblioteksdirektionen har aktivt arbetat med att på prov lägga in tecken i den lokala databasen, och är på grund av detta väl insatta i hur rekommendationerna i MARC-formatet ser ut när det gäller icke-latinsk

skrift. Detta tyder på ett mycket stort intresse av möjligheten att katalogisera med originalskrift oavsett språk. Trots bristande kännedom om förutsättningarna i MARC-formatet har även informant 5 och 6 visat förståelse för mervärdet av att använda originalskrift i katalogen, men då deras lokala system idag inte stödjer katalogisering med icke-latinska skrivsystem tittar man på andra möjligheter som att använda sig av bilder av text istället för riktig, sökbar text.

Genom att dra nytta av andras arbete kan biblioteken spara pengar då mindre tid måste läggas på att primärkatalogisera det egna materialet. Idealet är att en bok inte ska behöva primärkatalogiseras mer än en gång. Datoriseringen har aktualiserat denna princip då möjligheterna till utbyte över inte bara biblioteksgränser utan även landsgränser nu är fullt möjligt. Informant 1 är mycket positivt inställd till standardisering som exempelvis MARC21-formatet då hon anser att det inte finns någon anledning att ”uppfinna hjulet när det redan gjorts minst en gång innan”. Även informant 2 är mycket positiv till internationella standarder då det medför utbyte av poster mellan biblioteken. Informant 3 anser inte att internationell standardisering är livsnödvändigt på hennes bibliotek, men att det kan vara praktiskt inom vissa områden. Hon anser att datoriseringen har påverkat i vilken utsträckning man måste hålla sig till standarder, då man tidigare arbetade med katalogkort som bara berörde det egna biblioteket. ”Då hade man ju ganska fria händer. Man kunde tycka, nu kan man inte tycka lika mycket, nu bör man veta. Och man får resten av Sverige på sig om man avviker från normen.”

Informant 4 anser att standardisering utan tvekan är det bästa både för bibliotekarierna själva och bibliotekets användare då det förenklar arbetet och även informant 5 är mycket positiv till MARC-formatet. Informant 6 tycker att internationella standarder är att föredra framför lokala lösningar vad gäller bibliotekskatalogen. Hon anser att hennes bibliotek haft mycket stor glädje av att byta format till MARC21 från UKMARC som man tidigare använde. Hon menar att den absoluta majoriteten av de poster som importerar på hennes bibliotek är MARC-poster. Vid enstaka tillfällen importerar poster från ett danskt bibliotekssystem som inte använder MARC-formatet, vilket kan ställa till en del problem i katalogen om katalogisatören inte uppmärksammar skillnaden och redigerar posten. Detta förstärker ytterligare vikten av att utarbeta och följa fasta standarder för att på så sätt kunna dra nytta av det arbete som görs på olika bibliotek runt om i världen. Idealet skulle vara att kunna importera en post med japansk skrift från den japanska nationella bibliotekskatalogen och bara lägga minimalt med tid på att anpassa posten till de lokala förutsättningarna med exempelvis svenska ämnesord och klassifikation.

På grund av svårigheterna att utveckla teckenuppsättningar som fungerar för många språk samtidigt har en del andra lösningar uppstått. En är att använda sig av bilder av text, för att på så sätt komma runt teckenhantering och teckenstöd, då en bild för en dator alltid är en bild, oavsett vad den föreställer. Denna metod har dock en stor brist, nämligen att text i bildform inte är sökbar. Erickson tar upp detta i sin artikel ”Options for presentation of multilingual text: use of the unicode standard”:

On first glance, images seem the ideal solution: most if not all browsers understand their content and they are fairly simple to produce. However, these images are not searchable, and what is seen on the screen is what is copied to a local drive when the image is “saved”. A person saving an image-filled

page receives only a reference to the image, not the image itself. [...] while a particular application may solve the problems of display in this way, images in and of themselves do not provide a satisfactory general solution.

*Erickson 1997, sid. 175*

Informant 5 säger att standardisering är bra eftersom Sverige är för litet för egna lösningar, men samtidigt experimenterar de på biblioteket med egna lösningar i form av bilder i katalogen. Även informant 1 berättar om experiment som gjorts på andra bibliotek där bilder och liknande lösningar används, men menar att med tanke på alla tid och energi som gått åt till detta verkar unicode vara en enklare lösning.

Även om en lokal lösning hade kunnat underlätta sökandet på icke-latinska språk har samtliga bibliotekarier förespråkat internationella standarder i första hand. Förteckningar över vad som finns på de olika språken är ett önskemål som återkommer i undersökningen från Göteborg, och detta är på gång på MSB genom att man markerar det material som finns på olika språk för att kunna generera alfabetiska listor. På Malmö stadsbibliotek försöker man komma runt problemet med sökbarhet genom att lägga upp listor på litteratur på främmande språk på respektive språks skrift som kan nå genom en länk på katalogsidan. I Millennium finns också möjligheten att lägga in bilder på tecken i katalogposter, dessa är dock som vi påpekar ovan inte sökbara. Detta har genomförts för material på kinesiska och ska även genomföras för material på ryska. Det är författare och titel som visas på detta sätt.

Implementeringen av unicode handlar i första hand om standardisering. Det finns en uppsjö av lokala varianter som fungerar bra för just det landets eller bibliotekets behov, men dessa finns bara där, vilket omöjliggör kommunikation mellan datorer med andra system. För att fullt ut kunna utnyttja export och import av bibliografiska poster måste en enhetlig standard användas. Idag är det bara unicode som har kapacitet att koda alla världens språk enligt samma system. Unicode är den mest utbredda teckenuppsättningen och detta bidrar till att det även är det system som ligger bäst till för att bli internationell standard. En standard som inte används fyller inte sin funktion som standard.

### Transkribering och translitterering

En viktig faktor för den här undersökningen är det faktum att vi anser att transkribering och translitterering inte erbjuder ett tillräckligt bra alternativ till originalskrift. Detta är en åsikt som får stöd av den litteratur vi använt oss av i form av Wellisch och Risberg, men även något som bekräftades av våra informanter.

Språkstudier är mer vana vid att arbeta med transkribering och problemet är troligtvis inte lika stort för dem som för invandrare som ska söka material på sitt eget språk genom transkribering. Utbytesstudenter kan, beroende på förkunskaper både vad gäller språk och biblioteksvana, hamna mitt emellan. Både biblioteken vid Lunds universitet och biblioteken i Malmö använder sig av transkribering när de katalogiserar och hanterar material med icke-latinsk skrift, dels på grund av att bibliotekssystemen inte stöder andra teckenuppsättningar än den latinska och dels på grund av praktiska hanteringskäl. Flera av informanterna uttrycker dock ett visst missnöje över detta. Informant 3 illustrerar problemet mycket väl när hon talar om transkribering av svenska till japanska och använder sig av Astrid Lindgren som exempel (se sidan 53).

Enligt uppgifter från våra informanter har det hänt att material transkriberats enligt olika transkriptionsmodeller i samma katalog, vilket tydliggör problematiken med att inte ha enhetliga standarder. Hade materialet haft sin originalskrift i posten hade användaren lättare kunnat hitta denna även om transkriberingen var otydlig. Originalskrift i katalogen ökar alltså träffsäkerheten vid sökning. På Library of Congress diskuteras huruvida transkribering ska överges till förmån för originalskrift i den egna katalogen, detta på grund av den merkostnad som kommer av att göra det dubbla arbetet vid katalogisering då alla fält ska skrivas två gången, ett för transkription och ett för originalspråk (se kapitlet om standardisering). Della Porta ifrågasätter dock om en katalog med bara originalskrift skulle gå att administrera. Vi frågade våra informanter om deras åsikter kring detta, och de var alla överens om att det inte skulle vara praktiskt möjligt att överge transkribering då ingen kan behärska alla språken som biblioteket kan tänkas ha i sin samling. Däremot är informanterna väldigt positivt inställda till möjligheten att ha poster med både originaltecken och romaniserad form, framförallt om båda varianterna är sökbara, trots det merarbete det skulle innebära.

Vi håller med om att katalogen bör innehålla båda varianterna, trots de stora bristerna med romanisering, då det trots allt är det enda sättet för de som inte talar språket i fråga att kunna söka efter materialet. Det bästa vore dock om möjligheten att söka med originalspråk finns för de som behärskar detta, då det skulle vara överlägset romaniseringen vad gäller träffsäkerheten i sökningen.

### Språkkompetens hos personalen

I undersökningen "Mer nytta än nöje – en undersökning av invandrade göteborgares syn på biblioteksverksamhet" framgår att många av invandrarna uppfattade att träffa någon som pratar hans eller hennes egna språk som oerhört värdefullt. På de flesta biblioteken i vår undersökning fanns anställda med olika etnicitet och språklig bakgrund alternativt personal som hade för bibliotekets verksamhet lämpliga språkkunskaper. Ofta var dessa inte utbildade bibliotekarier och/eller assistenter, något som även Du Mont tar upp som ett vanligt problem i sin bok *Multiculturalism in libraries*. Alla våra informanter är överens om att språkkunskaper hos all personal på biblioteket är viktigt, men framförallt hos de i personalen som köper in och katalogiserar material på främmande språk. Informant 1 betonar att det skulle vara önskvärt att nästa bibliotekarie som anställs vid Asienbiblioteket har språkkunskaper för de aktuella språken, då det idag inte finns någon bibliotekarie med önskvärda språkkunskaper. Informanterna var även eniga om svårigheterna med att hitta utbildade bibliotekarier med en språkkompetens som går utanför skolspråken. På Malmö stadsbibliotek använder man sig mycket av informella kontakter med språkkunniga användare. Man försöker även knyta personal med speciell språkkompetens till biblioteket, ofta i form av projektanställningar, biblioteksassistenter eller språkvårdar (ungdomar med de efterfrågade språken som modersmål). Då denna bibliotekspersonal inte har den utbildning som övrig personal kan detta medföra att servicen inte blir den samma som för personer som ingår i den etniska och språkliga majoriteten. På biblioteken inom Lunds universitet utnyttjas närheten av lärare och forskare i de aktuella språken vid urval och bearbetning av material på främmande språk. Kontaktbibliotekarier knutna till de olika institutionerna bevakar sitt

ämnesområde i samråd med lärarna och forskarna via Internet, dagspress och information från förlag. Då Malmö stadsbibliotek inte har samma naturliga kontakt med språkexperter får de ta till andra metoder, så som samarbete med internationella bokagenter som Aagot, som specialiserat sig på utländsk litteratur samt utnyttja informella kontakter, som vi nämner ovan.

Då vi frågade informanterna om de såg en utveckling där specialister kommer att ersätta generalister på biblioteken med stora specialsamlingar som exempelvis samlingar på främmande språk ansåg de flesta att det på ett mindre bibliotek inte är praktiskt utan att alla måste vara generalister.

I undersökningen från Göteborg säger flera av deltagarna sig aldrig ha vänt sig till informationsdisken och kontakterna mellan personal och låntagare verkar fungera bättre på de mindre biblioteken eftersom personalen tycker sig känna låntagarna bättre och stämningen är mer avslappnad. Personalen anser också att det är lättare att hitta någon i närheten som kan fungera som tolk, något som vi kunnat se tydligt på Rosengårds stadsdelsbibliotek, där man engagerad användare som tolkar. Även på Malmö stadsbibliotek förekom detta i viss utsträckning. Enligt undersökningen från Göteborg använder de flesta invandrare inte bibliotekets katalog, utan använder en form av browsingmetod där de istället tittar igenom hyllorna för att se vilket material som finns på det aktuella språket. Detta beteende bekräftades från de flesta av informanterna i vår undersökning. Informant 4 säger sig ha fått förfrågningar från användare som efterlyser sökmöjligheter på originalspråk i katalogen medan Informant 5 inte alls stött på efterfrågan. Detta kan bero på mängden tid som spenderas i yttre tjänst, då Informant 4 verkar tillbringa mer tid i yttre tjänst än informant 5.

### Brister i kommunikationen mellan avdelningar

De informanter som har positioner som ansvariga för bibliotekssystemet eller som arbetat mycket med denna typ av frågor är bättre informerade och informanterna från Lund var överlag mer insatta i fråga om unicode. Det finns informationskanaler för katalogisatorer och systemavdelningen inom Lunds universitets bibliotek genom e-postlistor, men på grund av den stora mängden anställda som arbetar med katalogisering (c:a 80 personer) är det svårt att hålla organiserade möten för alla. Den decentraliserade organisationen ökar svårigheterna att utarbeta gemensamma standarder, och, som informant 2 påpekar, fungerar ett bibliotekssystem bäst med så få administratörer som möjligt. Det system som finns idag för katalogisering inom Lunds universitet ställer stora krav på en väl fungerande kommunikation.

Även i Malmö är kommunikationen mellan olika grupper i personalen oklar framförallt mellan stadsbiblioteket och stadsdelsbiblioteken. Stadsdelsbiblioteken katalogiserar inte material på svenska, men de katalogiserar material på främmande språk. När informanten från Rosengårds stadsdelsbibliotek tillfrågades om han kände till den kommande uppgraderingen av Libris svarade han att han läst ett mejl om det, men inte fått någon mer information. Det framgick även under intervjuerna att informanten på Rosengårds stadsdelsbibliotek och informanterna på Malmö stadsbibliotek hade olika uppfattning om regler och standarder vad gäller transkription, trots att de katalogiserar i samma katalog.

### Kritik mot samt alternativ till unicode för teckenhantering

Den kritik som uttryckts mot unicode faller sig vara ganska naturligt då ett så omfattande system som tar sig an alla världens språk kommer ha svårigheter att vara alla språken till lags ur alla tänkbara aspekter. Suzanne Topping konstaterar dock i sin artikel "The secret life of unicode: a peek at unicode's soft underbelly" att unicode trots kritiken är den mest tillfredställande lösningen som det ser ut idag. Även om vi inte frågade våra informanter rakt ut om huruvida de såg några problem med unicode eller kunde föreslå alternativa teckenuppsättningar fick vi intrycket att alla såg införandet av möjligheten med icke-latinsk skrift som positiv, men inte närmare satt sig in i de alternativ som finns förutom unicode och därför inte heller funderat över eventuell kritik. Inte ens informanten på Asienbiblioteket verkade kritisk trots att den största kritiken mot unicode kommer från CJK-språken. Ingen av informanterna verkade vara medvetna om alternativ till unicode.

## Slutsatser och slutdiskussion

I den här uppsatsen har vi undersökt så väl de tekniska förutsättningarna för hantering av icke-latinsk skrift som problematiken kring flerspråkig biblioteksverksamhet. Det finns en stor mängd material och undersökningar som behandlar de kulturella aspekterna vad beträffar invandrare och utbytesstudenter och bibliotek, men mindre om de språkliga aspekterna. Även då det gäller de mer tekniska aspekterna kring unicode och bibliotekssystem finns en uppsjö av rapporter och skrifter, medan de delar som fokuserar på användarnas behov och önskemål är färre. Genom att sammanfoga dessa två delar samt en kvalitativ undersökning ute på bibliotek anser vi att vi har bidragit men ett nytt perspektiv.

För första gången sedan datoriseringen av bibliotekskatalogerna finns nu en möjlighet att använda originalskrift i den bibliografiska posten. Detta innebär att material på icke-latinsk skrift kan hanteras på ett helt nytt sätt. Om bibliotek över hela världen håller sig till en specifik standard när det gäller teckenuppsättning underlättar det avsevärt både för katalogisatörer och slutanvändare, då det underlättar både import och export av bibliografiska poster samt förbättrar användarvänligheten för både de som söker och hanterar material med icke-latinsk skrift. Trots att MARC-8 har stöd för latin, JACKPHY, kyrilliska och grekiska är detta system inte lika bra som unicode, då det är en teckenuppsättning som bara används av bibliotek samt att möjligheterna att använda mer än ett av dessa språks skrivsystem i samma katalog är starkt begränsade.

En övergång till unicode är inte något som bara sker inom biblioteksvärlden, utan en stor fråga inom hela IT-sektorn. Allt fler system och applikationer görs unicodekompatibla för att stå sig väl i den globala konkurrensen, hela datorvärlden har tvingats tänka globalt istället för lokalt. Att unicode är framtiden verkar vara ett säkert antagande, och nästa steg blir att påvisa behovet av att kunna hantera fler skrivsystem i bibliotekskataloger för att underlätta informationsåtervinning oavsett källans språk.

Även om unicode verkar vara framtiden för bibliotekssystemen kommer det att ta tid innan alla bibliotek kommer att ha denna möjlighet, om de ens kommer att bli så att alla verkligen byter teckenuppsättning. Ekonomin på biblioteken verkar vara en viktig faktor som påverkar, men även den fördröjning som ofta uppstår när nya system och standarder ska införas och spridas. Detta är något som exporterande bibliotek bör ta hänsyn till, vilket till exempel Library of Congress också verkar göra. Under en övergångsperiod måste det finnas en möjlighet att välja bort användande av icke-latinsk skrift i lokala kataloger, för att bibliotek med system som inte stöder unicode ska kunna importera poster.



Det går alltså att dra slutsatsen att det ställs allt större krav på försäljarna av bibliotekssystem idag då de måste kunna tillgodose en flerspråkig marknad. Det är inte längre acceptabelt för ett bibliotekssystem att inte kunna hantera mer än ASCII eller ANSEL. Viktiga faktorer som lett till denna utveckling är till exempel att länder som har språk med icke-latinsk skrift blir en allt viktigare marknad för dator och mjukvaruföretagen och den ökade internationaliseringen där allt fler bibliotek importerar och exporterar poster över nationella gränser. Import och export av poster kräver gemensamma regler och standardisering av system. Det är mer ekonomiskt för biblioteken att anpassa sig till allmänna standarder vilket krävs för utbyte av data mellan biblioteken. Bibliotekssystem som inte följer utvecklingen med unicode riskerar att stå utanför i konkurrensen. Fram tills idag har dock bara ett av bibliotekssystemen på de bibliotek vi undersökt haft unicodeanpassning (Virtua), vid Lunds universitet. Detta trots att lokala lösningar isolerar. För att verkligen kunna dra nytta av andras arbete krävs standardisering.

Genom välfungerande internationell standardisering underlättas kommunikation mellan datorsystem, vilket i sin tur underlättar arbetet för biblioteken ur många aspekter. Genom att använda sig av samma katalogiseringssystem och samma teckenuppsättning som ett bibliotek i Amerika eller Asien kan svenska bibliotek dra nytta av det arbete som gjorts där. Datorsystem mår bäst av total likriktning genom gemensamma regler, då det förhindrar att data blir korrupt eller inte kan mottas av olika system.

Även om några av informanterna experimenterar med egna lokala lösningar i sina bibliotekssystem vad gäller icke-latinsk skrift var alla överens om att internationella standarder var det bästa.

Beslut som tas om katalogisering påverkar stora delar av bibliotekets verksamhet och kan ha stora konsekvenser för användarnas informationssökning. Katalogen kan ses som en central spindel i nätet som påverkar informationsåtervinningen i hög grad. Till exempel kan avsaknaden av ett gemensamt transkriberingsschema innebära att en sökt bok inte återfinns.

Vanliga bibliotekarier och katalogisatörer behöver inte känna till alla detaljerna kring unicode eller teckenuppsättning, det räcker med att systemavdelningen jobbar med det tekniska. Bibliotekarierna behöver bara veta hur de rent praktiskt ska använda sig av tecken i katalogen. Vi tror att många katalogisatörer idag inte är så insatta i tekniken bakom de teckenuppsättningar som används, och därför behöver de inte heller veta allt om unicode för att använda det. Däremot är en bra kommunikation mellan katalogisatörer, systemansvariga tekniker och bibliotekarierna ute "på golvet" viktig. Tyvärr verkar kommunikationen mellan olika personalgrupper, till exempel mellan systemavdelning, katalogisatörer och andra delar av verksamheten, brista på de bibliotek vi undersökt. Detta kanske delvis beror på bibliotekets organisation. Det är till exempel mycket svårt att hålla god kommunikation mellan 80 katalogisatörer som vid Lunds universitets bibliotek. En annan faktor kan vara rädsla bland bibliotekarier för att visa sig okunniga, vilket leder till att kommunikationen blir sämre mellan systemansvariga och katalogisatörer, en teori som en av våra informanter lade fram.

Personal och dess språkkompetens är ett annat viktigt problemområde. Oftast är den språkkompetenta personalen inte utbildade bibliotekarier, vilket begränsar servicen till användarna vilket gör att det lätt blir en snedvriden service till de olika användargrupperna. Samtidigt som det inte är realistisk att kräva att alla språk ska vara representerade i den utbildade personalstyrkan, bör bibliotekspersonalen i så stor utsträckning som möjligt spegla sammansättningen i samhället. Vi är dock medvetna om svårigheterna att rekrytera utbildade bibliotekarier med invandrabakgrund eller kunskaper i språk utom för skolspråken då dessa är få. Speciell språkkompetens är önskvärd, men inte alltid möjligt. Ju färre språk biblioteket hanterar desto större sannolikhet att de kan täcka in kompetensen, men på exempelvis Malmö stadsbibliotek är det omöjligt att ha kompetens inom alla språk i personalen.

Det finns inom den svenska bibliotekariekåren redan en diskussion om att dra nytta av utbildade bibliotekarier bland invandrare genom mentorskap och överbyggande utbildning. Avsaknad av kunskaper i svenska är antagligen det största hindret för invandrade bibliotekarier för att få jobb. Det enda alternativet till språkkompetent personal har på flera bibliotek varit att utnyttja de informella kontakterna med experter som lärare, invandrare och liknande. Här kan uppstå problem då bibliotekarierna inte har möjlighet att kvalitetsgranska materialet, utan måste lita på sina källor. De användare som inte har samma modersmål som majoriteten riskerar då att inte få likvärdig service eftersom de som köper in och katalogiserar material på dessa språk inte har likvärdig utbildning som de bibliotekarier som har hand om material på majoritetsspråket.

Genom att ha en katalog med större språkmöjligheter underlättar man för bibliotek som har användare med olika språkbehov då dessa kan söka fram vilket material som finns på andra bibliotek, vilket skulle underlätta vid fjärrlån. Användare med ett annat modersmål är av olika skäl obenäga att använda katalogen, vilket starkt begränsar deras möjligheter att tillgå bibliotekets fulla resurser som till exempel fjärrlån från andra bibliotek. Möjligheten att söka i och använda katalogen på användarnas modersmål skulle troligtvis öka deras användning av denna.

På alla biblioteken i vår undersökning har man idag en hantering som enligt informanterna fungerar för den dagliga verksamheten, men det finns dock önskemål om att hitta effektivare lösningar. Exempelvis ansåg alla vid första anblick att transkriberingen fungerade bra men vid närmare eftertanke tyckte de flesta att det fanns brister. Många var överens om att transkriberingen ofta är en barriär för användarna och hindrar åtkomst i katalogen snarare än underlättar den. Idealet för både personal och användare är att posten är tvåspråkig, det vill säga innehåller både sökbara originalskriftsfält och transkriberade fält.

Vi tror att ett hinder för utvecklingen kan vara att man inte anser sig ha tid till att titta upp från det dagliga arbetet för att se vilka alternativ och möjligheter som finns att lösa dessa problem. Vi tror dock att detta snarare beror på stor arbetsbelastning än ointresse. Vår undersökning visar att det finns ett intresse bland många av de som arbetar med den här typen av frågor, särskilt hos de som arbetar mycket nära med användarna. Alla våra informanter har varit entusiastiska och positiva till att lägga in originalskrift i katalogen,

men vår tolkning av intervjuerna är att många är tveksamma inför vad det skulle innebära tekniskt och ekonomiskt.

Vi tycker oss i denna undersökning ha sett en bristande kommunikation mellan de olika avdelningarna på biblioteken, där närheten till användarna spelat en stor roll för lyhördhet mot användarnas reella behov. Ju närmare användarna informanterna har varit desto bättre verkar förståelsen för värdet med originalspråk i katalogen vara, detta var särskilt tydligt på Rosengårds stadsdelsbibliotek jämfört med Malmö stadsbibliotek.

Med Malmö stadsbiblioteks tydliga engagemang för att ge lika service till alla, inte minst invandrare med ett annat modersmål än svenska samt viljan att "ligga i frontlinjen" när det gäller digitala tjänster både nationellt och internationellt borde det finnas ett stort intresse för unicodeanpassning både på stadsbiblioteket och de stadsdelsbibliotek som är belägna i invandrartäta stadsdelar och de möjligheter som en sådan anpassning skulle kunna ge till hantering av medier med icke-latinsk skrift i katalogen, framförallt för användare som skulle kunna få möjlighet att söka litteratur och andra medier på sitt modersmål. Vi upplevde att viljan fanns men att det inte var prioriterat i dagsläget.

Det finns en skillnad mellan Malmö stadsbibliotek och Lunds universitets bibliotek då det på Malmö stadsbibliotek oftast är invandrare som efterfrågar material på främmande språk, medan det på Lunds universitets bibliotek oftast är forskare och utbytesstudenter. Invandrare är en heterogen grupp som är svår att definiera, det kan finnas alla tänkbara bakgrunder, utbildning, biblioteksvana osv. Språkstudenter har oftast ett annat modersmål än det språk de studerar och är en relativt homogen grupp i jämförelse med invandrare. De har oftast också tidigare studie-, dator- och biblioteksvana, faktorer som kan variera kraftigt i invandragrupper. Utbytesstudenter hamnar i ett mellanläge. De har oftast ett annat modersmål men liknar mer studenterna än invandrarna på grund av en liknande studiebakgrund. Utbytesstudenter är dock ofta helt obekanta med svenska transkriptionssystem.

Trots att man serverar mycket olika användargrupper på folk- och högskolebibliotek verkar det dock röra sig om samma problematik då originalskrift hade tjänat båda användargrupperna bättre. En skillnad är att studenter är mer vana vid att transkribering än invandrarna, varför transkribering torde vara mer problematiskt för folkbiblioteken, särskilt i invandrartäta områden.

Nu när vi etablerat att det finns ett klart uttalat behov och intresse av hantering av icke-latinsk skrift i bibliotekskataloger skulle det vara intressant att följa upp detta genom att tala med personer i chefspositioner på biblioteken för att klargöra eventuella strategier och prioriteringar inför framtiden vad gäller den här frågan. Personerna vi talat med är visserligen de som arbetar med dessa frågor i sin vardag, men saknar den överblick och det strategiska tänkandet som informanter i chefspositioner skulle kunna erbjuda.

## Litteraturförteckning

Adams, Glen (1994) *Introduction to Unicode*. Cambridge, Institute for Advanced Professional Studies, Unicode Implementers Workshop, Tokyo, 11-12 April.

Aliprand, Joan M. (1992) Nonroman scripts in the bibliographic environment. *Information Technology and Libraries*, vol. 11.

Arsenault, Clément (2002) Pinyin romanization for OPAC retrieval – is everyone being served? *Information Technology and Libraries*, vol. 21:2.

Arsky, J.M. & Cherny, A.I. (1997) The Ethno-Cultural, Linguistic and Ethical Problems of the ‘infosphere’. *The International Information & Library Review*, vol. 29:2.

*Biblioteksplan Malmö stadsbibliotek* (2005) (Elektronisk) Tillgänglig:  
<<http://www2.malmo.stadsbibliotek.org/biblioteksplan.pdf>> (2006-05-02)

Björkhem, Miriam & Lindholm, Jessica (2000) *Metadata för det digitala biblioteket – objektbeskrivning av digitala resurser*. Magisteruppsats, Lunds universitet.

Brunnström, Ann-Christine (2005) Mer nytta än nöje – en undersökning av invandrade göteborgares syn på biblioteksverksamhet. Göteborgs Stad. Tillgänglig:  
<[http://www.stadsbiblioteket.goteborg.se/prod/kultur/stadsbibliotek/dalis2.nsf/0/0414aa4388af2105c1256cbf004f1fe0/\\$FILE/rapport.pdf](http://www.stadsbiblioteket.goteborg.se/prod/kultur/stadsbibliotek/dalis2.nsf/0/0414aa4388af2105c1256cbf004f1fe0/$FILE/rapport.pdf)> (2006-04-26)

Boss, Richard W. (2003) Unicode. (Elektronisk) American Library Association. Tillgänglig: <<http://www.ala.org/ala/pla/plapubs/technotes/unicode.htm>> (2006-04-27)

C1 Controls and Latin-1 Supplement (2005) (Elektronisk) Tillgänglig:  
<<http://www.unicode.org/charts/PDF/U0080.pdf>> (2006-04-26)

Cain, Jack (2004) Assessment of options for handling full unicode character encodings in MARC21 – a study for the Library of Congress Part 1: New scripts. Tillgänglig:  
<<http://www.loc.gov/marbi/2004/2004-report01.pdf>> (2006-04-26)

Centre for East- and Southeast Asian Studies (Elektronisk) Tillgänglig:  
<<http://www.ace.lu.se>> (2006-06-05)

Chandrakar, Rajesh (2004) Unicode as a multilingual standard with reference to Indian languages. *The Electronic Library*, vol. 22:5.

- Della Porta, Ann (2005) Planning for Unicode in Libraries: the LC perspective. (Elektronisk) SLA Annual Conference 2005. Tillgänglig: <<http://www.loc.gov/ils/UNICODE/Unicode@SLA.pdf>> (2006-04-26)
- Du Mont, Rosemary Ruhig et al. (1994) *Multiculturalism in libraries*. Westport, Greenwood Press
- Erickson, Janet C. (1997) Options for presentation of multilingual text: use of the unicode standard. *Library Hi Tech*, vol. 15:3.
- Holdings, Location, Alternate Graphics, ect. Field (841-88X)(Elektronisk) Tillgänglig: <<http://www.loc.gov/marc/bibliographic/ecbdhold.html#mrcb880>> (2006-04-27)
- Innovative Interfaces (2003) (Elektronisk) Tillgänglig: <<http://www.iii.com>> (2006-04-27)
- Innovative Interfaces Millennium: Cataloging (2004) (Elektronisk) Tillgänglig: <<http://www.iii.com/mill/catalog.shtml>> (2006-04-27)
- Innovative Interfaces Millennium: Technical information (2004) (Elektronisk) Tillgänglig: <<http://www.iii.com/mill/tech.shtml>> (2006-04-27)
- Jeong, Wooseob (2004) Unbreakable ethnic bonds: Information-seeking behaviour of Korean graduate students in the United States. *Library and information science research*, vol. 26:3.
- Kockum, Keiko (2001) *Japanska för svensktalande*. Del 1. Institutionen för östasiatiska språk, Lunds universitet.
- Kvale, Steinar (1997) *Den kvalitativa forskningsintervjun*. Lund, Studentlitteratur
- Lantz, Annika (1993) *Intervjumetodik - den professionellt genomförda intervjun*. Lund, Studentlitteratur
- Library of Congress pinyin conversion project: coordinated timeline (Elektronisk) Tillgänglig: <<http://www.loc.gov/catdir/pinyin/timeline.html>> (2006-04-30)
- Libris katalogiseringsklient – nyheter (Elektronisk) Tillgänglig: <[http://www.libris.kb.se/tjanster/katalogisering/voyager\\_nyheter.jsp](http://www.libris.kb.se/tjanster/katalogisering/voyager_nyheter.jsp)> (2006-05-01)
- Libris uppgradering – katalogiseringsklient (Elektroniskt) Tillgänglig: <<http://www.libris.kb.se/tjanster/katalogisering/uppgradering06.jsp>> (2006-05-01)
- Lunds universitets bibliotek – organisation (Elektronisk) Tillgänglig: <[http://www.lub.lu.se/new\\_top/omlub/om\\_lub.shtml](http://www.lub.lu.se/new_top/omlub/om_lub.shtml)> (2006-05-01)

Malmöbor (2005) (Elektronisk) Tillgänglig:

<<http://www.malmo.se/download/18.fe3a2d310a090ba6c380003835/Sammanfattn..doc>> (2006-05-02)

Malmö stad – 02. (2006) Utländsk bakgrund (Elektronisk) Tillgänglig:

<<http://www.malmo.se/faktaommalmopolitik/statistik/02utlandskbakgrund.4.33aee30d103b8f15916800028396.html>> (2006-02-06)

Mandernak, Scott B., Lin, Poping & Hovde, David M. (1994) Cultural awareness and bibliographic instruction in academic libraries. I Hoover Hill, Katherine (red.) *Diversity and multiculturalism in libraries*. Greenwich, JAI Press

McCallum, Sally (2005) Assessment of options for handling full unicode character encodings in MARC21 Part 2: Issues. (Elektronisk) Tillgänglig:

<<http://www.loc.gov/marc/marbi/2005/2005-report01.pdf>> (2006-04-26)

Memon, Aslam Parvez (2001) Study of Unicode specifications and their implementation in Arabic script languages by designing a multilingual unicode editor. *Multi Topic Conference, 2001. IEEE INMIC 2001. Technology for the 21st Century. Proceedings. IEEE International*.

Risberg, Elisabeth (2003) *Originalskrift i flerspråkiga bibliotekskataloger*. Magisteruppsats, Bibliotekshögskolan i Borås.

Spolsky, John (2003) The Absolute Minimum Every Software Developer Absolutely, Positively Must Know About Unicode and Character Sets (No Excuses!). (Elektronisk) Tillgänglig: <<http://www.joelonsoftware.com/articles/Unicode.html>> (2006-04-26)

Språk- och litteraturcentrum vid Lunds universitet – ämnen (Elektronisk) Tillgänglig: <<http://www.sol.lu.se/amnen/index.html?topMenuIndex=4>> (2006-05-01)

Språk- och litteraturcentrum vid Lunds universitet – organisation (Elektronisk) Tillgänglig: <[http://www.sol.lu.se/bibliotek/om/index.html?expand\\_menu=1](http://www.sol.lu.se/bibliotek/om/index.html?expand_menu=1)> (2006-05-01)

Topping, Suzanne (2001) ”The secret life of unicode: a peek at unicode’s soft underbelly” (Elektronisk) Tillgänglig: <<http://www-128.ibm.com/developerworks/library/u-secret.html>> (2006-05-28)

Tull, Laura (2002) Library systems and Unicode: a review of the current state of development. *Information Technology and Libraries*, vol. 21:4.

Understanding MARC bibliographic: parts 1 to 6 (2003) (Elektronisk) Tillgänglig: <<http://www.loc.gov/marc/umb/um01to06.html>> (2006-04-27)

Unicode-MARC Archives August 2005 (#54) (Elektronisk) Tillgänglig: <<http://listserv.loc.gov/cgi-bin/wa?A2=ind0508&L=unicode-marc&T=0&P=6505>> (2006-04-27)

Unicode-MARC Archives August 2005 (#55) (Elektronisk) Tillgänglig:  
<<http://listserv.loc.gov/cgi-bin/wa?A2=ind0508&L=unicode-marc&T=0&P=6621>>  
(2006-04-27)

Unicode-MARC Archives August 2005 (#60) (Elektronisk) Tillgänglig:  
<<http://listserv.loc.gov/cgi-bin/wa?A2=ind0508&L=unicode-marc&T=0&P=7199>>  
(2006-04-27)

Unicode Standard: a Technical Introduction, The (2006) (Elektronisk) Tillgänglig:  
<<http://www.unicode.org/standard/principles.html>> (2006-04-26)

USMARC and CAN/MARC become MARC 21 (1998) (Elektronisk) Tillgänglig:  
<<http://www.loc.gov/marc/annmarc21.html>> (2006-04-27)

Virtua ILS (Elektronisk) Tillgänglig: <<http://www.vtls.com/virtua/components.shtml>>  
(2006-04-27)

Voyager (Elektronisk) Tillgänglig:  
<[http://www.endinfosys.com/prods/brochures/RB/full\\_voyager\\_FINAL.pdf](http://www.endinfosys.com/prods/brochures/RB/full_voyager_FINAL.pdf)> (2006-04-27)

Voyager with Unicode: Standards for Libraries (2003) (Elektronisk) Tillgänglig:  
<<http://www.endinfosys.com/prods/unicode3.pdf>> (2006-04-27)

VTLS Inc. Corporate Information (2004) (Elektronisk) Tillgänglig:  
<<http://www.vtls.com/Corporate/>> (2006-04-27)

VTLS: Visionary Technology in Library Solutions (Elektronisk) Tillgänglig:  
<<http://www.vtls.com/brochures/vtls.pdf>> (2006-04-27)

Wellisch, Hans H. (1976) *The Conversion of Scripts – its Nature, History, and Utilization*. New York, John Wiley & Sons.

Zoe, L.R & DiMartino, D. (2000) Cultural diversity and end-user searching - An analysis by gender and language background. *Research Strategies*, vol. 17:4.

## Bilagor

### Bilaga 1 – Ordlista

ANSEL	ANSEL är namnet på en teckenuppsättning avsedd för att behandla ”extended latin”, dvs. en utbyggnad av det latinska alfabetet innefattande 36 specialtecken och 29 diakriter
ASCII	American Standard Code for Information Interchange. Den första teckenuppsättningen i datormiljö, anpassad för det engelska språket och dess skrivsystem. ASCII är byggt på 7 eller 8 bitar, vilket som mest ger 128 alternativt 256 tecken.
CJK-språken	CJK-språken är en förkortning för de östasiatiska språken kinesiska, japanska och koreanska.
Diakrit	Tecken som skiljer snarlika bokstäver åt, exempelvis accenter (ó), tilde (ö) och cirkumflex (ô).
Font	En font eller typsnitt är den representation som text har på exempelvis en datorskärm. Exempel på fonter är Arial, Times new Roman eller Courier.
Glyph	Den visuella representationen av en kod, exempelvis en bokstav.
HTML	Hyper text markup language. Det språk som oftast används vid kodning av hemsidor på Internet.
IME	Input method editor. Ett program som eller system som ger möjligheten att skriva komplexa tecken som kinesiska och japanska med ett vanligt västerländskt tangentbord.
ISO	International Standards Organisation är en organisation som definierar och verifierar internationella standarder.



JACKPHY-språken	Akronym för de sju språken: japanska, arabiska, kinesiska, persiska, hebreiska och yiddish, alla med skrivsystem som inte utgår från det latinska.
MARC	Machine Readable Cataloging Record. Ett format uppbyggt kring fält och delfält som används av bibliotekssystem över hela världen.
OCLC	Online Computer Library Center. En stor amerikansk agent inom biblioteks- och datorutveckling.
PUA	Private use area. En rad koder inom unicode standarden som inte har formell status, utan kan användas till valfria symboler, till exempel varumärkessymboler. Då de saknar formell status finns ingen garanti att de representeras på rätt sätt i ett system utanför det egna.
Radikaler	De grundkombinationer av streck som ett kinesiskt eller japansk tecken är uppbyggt av.
RLG	Research Libraries Group. Amerikanskt ickekommersiellt företag som arbetar med informationssystem.
Typsnitt	se font.
Unicode	En standard för att hantera tecken i datormiljö. Gör anspråk på att ha kapacitet att koda alla världens skrivspråk inklusive utdöda språk.
z39.50	z39.50 är ett standardiserat protokoll som möjliggör kommunikation mellan olika databaser. Protokollet används ofta för att importera enstaka poster åt gången från en central databas till det lokala systemet.

