

# Datorstöd vid datainsamlingen

Carina Persson, U/STM

Marit Jorsäter, D/UTV



R&D Report  
Statistics Sweden  
Research-Methods-Development  
1990:9

## INLEDNING

### TILL

**R & D report : research, methods, development / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1988-2004. – Nr. 1988:1-2004:2.**

**Häri ingår Abstracts : sammanfattningar av metodrapporter från SCB med egen numrering.**

#### **Föregångare:**

Metodinformation : preliminär rapport från Statistiska centralbyrån. – Stockholm : Statistiska centralbyrån. – 1984-1986. – Nr 1984:1-1986:8.

U/ADB / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1986-1987. – Nr E24-E26

R & D report : research, methods, development, U/STM / Statistics Sweden. – Stockholm : Statistiska centralbyrån, 1987. – Nr 29-41.

#### **Efterföljare:**

Research and development : methodology reports from Statistics Sweden. – Stockholm : Statistiska centralbyrån. – 2006-. – Nr 2006:1-.

# Datorstöd vid datainsamlingen

Carina Persson, U/STM

Marit Jorsäter, D/UTV



R&D Report  
Statistics Sweden  
Research-Methods-Development  
1990:9

|                   |                                                  |
|-------------------|--------------------------------------------------|
| Från trycket      | September 1990                                   |
| Producent         | Statistiska centralbyrån, utvecklingsavdelningen |
| Ansvarig utgivare | Åke Lönnqvist                                    |
| Förfrågningar     | Håkan Lindström, tel. 019-17 66 94               |

© 1990, Statistiska centralbyrån  
ISSN 0283-8680  
Printed in Sweden

SCB-Tryck, Örebro. Miljövänligt papper

## FÖRORD

Rapporten DATORSTÖD VID DATAINSAMLINGEN, är resultatet av tre månaders praktikarbete vid U/STM och D/UTV under sommaren 1990, som utförts av Marit Jorsäter och Carina Persson, båda statistikstuderande vid högskolan i Örebro. I arbetet har flera SCBare, speciellt vid D/DC, I/UK, och U/ADB hjälpt till med synpunkter och information.

Rapporten består av två delar. Denna är en översiktsdel som vi härmed ger en mycket bred spridning. Den andra en fyllig sammanställning av fallstudier, som vi planerar att sprida främst till personer som skall arbeta aktivt inom området. **Rapporten är i första hand skriven för produktansvariga och användare och tar endast i begränsad utsträckning upp rena ADB-tekniska frågor.**

Syftet med denna rapport är att inom verket ge en bred spridning av information om några av de moderna tekniker för datainsamling, som får en allt större användning både vid centralbyråer och hos andra statistikproducenter. **Både teknikutvecklingen, ansträngningarna att implementera dess resultat och litteraturen om detta är mycket omfattande.** Hos oss finns än så länge enbart ett fåtal produkter, som infört eller försöker införa någon av de tekniker som beskrivs i rapporten. Det är viktigt att SCB inte hamnar på efterkälken.

De tekniker som diskuteras är i första hand sådana, som vi tror är så utprovade att de kan implementeras vid SCB med i stort sett förutsägbara effekter. Urvalet av artiklar har dock varit svårt med den korta tid som stått till förfogande men vi gör heller inte anspråk på fullständighet.

Örebro 1990-09-04

Håkan L Lindström U/STM



# INNEHÅLLSFÖRTECKNING

|                                                                   |    |
|-------------------------------------------------------------------|----|
| <b>INLEDNING</b> .....                                            | 1  |
| <b>BEGREPP</b> .....                                              | 2  |
| <b>UPPLÄGGNING</b> .....                                          | 3  |
| <br><b>1. DATORSTÖD VID REGISTRERINGSTILLFÄLLET</b> .....         | 4  |
| <b>1.1. OCR (Optical Character Recognition)</b> .....             | 4  |
| Inledning .....                                                   | 4  |
| Teknik .....                                                      | 5  |
| Vinstpotential .....                                              | 5  |
| Målpopulation och Informationsbehov .....                         | 6  |
| <b>1.2. OMR (Optical Mark Reading)</b> .....                      | 6  |
| Inledning .....                                                   | 6  |
| Teknik .....                                                      | 6  |
| Vinstpotential .....                                              | 7  |
| Målpopulation och Informationsbehov .....                         | 7  |
| <br><b>2. DATORSTÖD FÖR INTERVJUAREN</b> .....                    | 8  |
| <b>2.1. CATI (Computer Assisted Telephone Interviewing)</b> ..... | 8  |
| Inledning .....                                                   | 8  |
| Teknik .....                                                      | 9  |
| Hårdvara och programvara .....                                    | 9  |
| Programspråk .....                                                | 9  |
| Kontrollfunktioner hos programvaran .....                         | 9  |
| Organisationsstruktur .....                                       | 10 |
| Vinstpotential .....                                              | 11 |
| Målpopulation och Informationsbehov .....                         | 12 |
| <b>2.2. CAPI (Computer Assisted Personal Interviewing)</b> .....  | 13 |
| Inledning .....                                                   | 13 |
| Teknik .....                                                      | 13 |
| Bärbara och Handhållna datorer .....                              | 13 |
| HCR (Handwritten Character Recognition) .....                     | 14 |
| Programvara .....                                                 | 15 |
| Några Synpunkter .....                                            | 16 |
| Kommunikation .....                                               | 17 |
| Organisationsstruktur .....                                       | 17 |
| Vinstpotential .....                                              | 18 |
| Målpopulation och informationsbehov .....                         | 18 |
| <b>2.3. DATI (DATorstödda Insamlingsmetoder)</b> .....            | 19 |
| Inledning .....                                                   | 19 |
| Teknik .....                                                      | 19 |
| Organisationsstruktur .....                                       | 20 |
| Vinstpotential .....                                              | 21 |
| Målpopulation och Informationsbehov .....                         | 21 |

|                                                                                 |    |
|---------------------------------------------------------------------------------|----|
| <b>3. DATORSTÖD FÖR UPPGIFTSLÄMNAREN</b> . . . . .                              | 22 |
| <b>3.1. TDE (Touchtone Data Entry)</b> . . . . .                                | 22 |
| Inledning . . . . .                                                             | 22 |
| Teknik . . . . .                                                                | 23 |
| Organisationsstruktur . . . . .                                                 | 24 |
| Vinstpotential . . . . .                                                        | 24 |
| Målpopulation och Informationsbehov . . . . .                                   | 25 |
| <b>3.2. CASAQ (Computer Assisted Self-Administered Questionnaire)</b> . . . . . | 25 |
| Inledning . . . . .                                                             | 25 |
| Teknik . . . . .                                                                | 25 |
| Kommunikation . . . . .                                                         | 26 |
| Exempel på CASAQ-system . . . . .                                               | 26 |
| Vinstpotential . . . . .                                                        | 27 |
| Målpopulation och Informationsbehov . . . . .                                   | 28 |
| <br><b>4. DATORSTÖD I FRAMTIDEN?</b> . . . . .                                  | 29 |
| <b>4.1. Talteknologi (Röstsimulering och Talidentifiering)</b> . . . . .        | 29 |
| <br><b>SAMMANFATTNING</b> . . . . .                                             | 30 |
| Teknik . . . . .                                                                | 30 |
| Organisationsstruktur . . . . .                                                 | 32 |
| Vinstpotential . . . . .                                                        | 33 |
| Målpopulation och Informationsbehov . . . . .                                   | 35 |
| Några Synpunkter . . . . .                                                      | 36 |

## KÄLLFÖRTECKNING



Syftet med denna rapport är att ge läsaren en kort överblick över såväl de mest etablerade metoderna för datorstöd vid datainsamlandet, som de metoder där intensiv forskning och utveckling pågår.

Intresserade läsare kan även rekvidrera ytterligare en rapport i ämnet:<sup>1</sup>

- \* DATORSTÖD VID DATAINSAMLINGEN - Fallstudier och Artikel-förteckning

## **INLEDNING**

Den ökade utvecklingen av teknologin de senaste decennierna har ställt nya hjälpmedel till förfogande vid datainsamlingen. Det konventionella hjälpmedlet vid besöks- och telefonintervjuer har varit papper och penna. Detta kallas allmänt för PAPI (PAper and Pen Interviewing). Den nya tekniken har i många fall förändrat intervjusituationen.

Användningen av datorstöd vid datainsamlingen förändrar arbetsuppgifterna så att behovet av kvalificerad personal ökar på bekostnad av mindre kvalificerad personal.

Ett användande av datorstöd vid datainsamlingen kan medföra:

- \* Integrering, olika funktioner i statistikproduktionen slås samman.<sup>2</sup>
- \* Rationaliseringsvinster, många monotona och rutinmässiga arbetsuppgifter blir överflödiga.
- \* Tidsvinster, integreringen av olika arbetsmoment och automatiseringen kan skynda på slutprodukten.
- \* Kvalitetshöjningar, bättre och tidigare kontroll erhålls över insamlingsprocessen.
- \* Standardisering, en gemensam standard för att lagra och dokumentera data underlättar datahanteringen.
- \* Omprioritering, övergång från en datainsamlingsmetod till en annan.

Integrationsprocessen kan försvåras om man vid de olika funktionerna i en undersökning har olika systemtänkande.

De flesta datorbaserade produktionssystem är utbyggda från ett av perspektiven:

- \* Datainsamling
- \* Databehandling
- \* Dataanalys

---

<sup>1</sup> Materialet kan beställas från: Marianne Rydberg, U/ADB, Statistiska Centralbyrån (SCB), Örebro.

<sup>2</sup> Den nya tekniken ledde på SCB till en omfördelning och integrering av arbetsuppgifterna, genom det sk ADA-projektet (Anskaffning av DATamaterial). Se Bergsten, O.; (1990)

Ett alternativ vore att göra ett helt nytt system som tar hänsyn till alla tre perspektiven. Inget system har, enligt Merrill Shanks<sup>3</sup> (1989), lyckats integrera alla tre perspektiven på ett optimalt sätt.

Inom detta område återstår mycket arbete. Shanks (1989) har identifierat tre strategier för en ökad integrering:

1. Välj ett system för varje steg och skapa förbindelser mellan dessa (en-till-en-strategin). Denna strategi kan bli mycket dyr att genomföra.
2. Skapa ett enda gemensamt system för alla stegen. Även om detta är den bästa strategin i längden är den mycket svår att genomföra. Ett system som inte blir bra kan innebära stora kostnader.
3. Skapa ett generellt "förbindelse system" för steg med samma data-definitioner (många-till-många-strategin). Om en gemensam standard för att lagra och dokumentera data skapas innebär detta att data fritt kan kopieras mellan de olika systemen.

## **BEGREPP**

Här presenteras de nyckelord (förkortningar) av datorstödda insamlingstekniker som används i rapporten.

|       |                                                     |
|-------|-----------------------------------------------------|
| CAPI  | - Computer Assisted Personal Interviewing           |
| CASAQ | - Computer Assisted Self-Administered Questionnaire |
| CATI  | - Computer Assisted Telephone Interviewing          |
| DATI  | - DATorstödda Insamlingsmetoder                     |
| DAVID | - DATainsamling Via Diskett                         |
| HCR   | - Handwritten Character Recognition                 |
| HHC   | - Hand-Held Computer                                |
| IFK   | - Interaktiv Fråge-Konstruktion                     |
| OCR   | - Optical Character Recognition                     |
| OMR   | - Optical Mark Reading                              |
| PAPI  | - PAPER and Pen Interviewing                        |
| TDE   | - Touchtone Data Entry                              |

I litteraturen förekommer även andra förkortningar för samma begrepp. Vi har i vår rapport valt att konsekvent använda oss av en av dessa förkortningar utan att för den skull lägga in några värderingar om vilken förkortning som är den bästa. För kännedom kan även nämnas CSAQ (Computerized Self-Administered Questionnaire) som är samma sak som CASAQ.

---

<sup>3</sup> Director, Computer-assisted Survey Methods, University of California, USA

## **UPPLÄGGNING**

I det följande ska några av de viktigaste datorbaserade metoderna kortfattat presenteras. De kommer att presenteras likartat och för de metoder det är relevant kommer vi att beröra följande avsnitt:

### **Inledning**

Här presenteras metoden kortfattat och eventuell historia och pågående forskning introduceras.

### **Teknik**

Detta avsnitt beskriver den teknik som en användare behöver känna till.

### **Organisationsstruktur**

Avsnittet beskriver hur den förändrade tekniken kan motivera organisationsförändringar.

### **Vinstpotential**

Här behandlas fördelar och nackdelar med avseende på bland annat kostnad och kvalitet samt eventuella tidsvinster. Det har visat sig i den litteratur vi studerat att dessa aspekter är relativt dåligt dokumenterade då de flesta undersökningar mest är inriktade på att testa hur tekniken fungerar. Avsnittet vinstpotential är därför genomgående ganska tunt och speglar behovet av vidare forskning på detta område.

### **Målpopulation och Informationsbehov**

De eventuella begränsningar för möjliga målpopulationer som metoden medför diskuteras. Avsnittet behandlar även det informationsbehov tekniken tillåter, dvs den sorts information som en undersökning kan inhämta med hjälp av en viss teknik. Här behandlas även personalens och uppgiftslämnarnas inställning till datorer.

## **1. DATORSTÖD VID REGISTRERINGSTILLFÄLLET**

### **1.1. OCR (Optical Character Recognition)<sup>4</sup>**

#### **Inledning**

Denna teknik har tidigare använts på SCB och innebär att blanketter läses av och dataregistreras automatiskt av en optisk läsare.

Den senaste varianten av OCR-läsare som användes på SCB anskaffades i början på 70-talet för att användas till Folk- och Bostadsräkningen 1975 (FoB 75). Den användes även vid FoB 80 och FoB 85.

Resultatet förbättrades kontinuerligt. Vid FoB 75 var andelen rejects, dvs oläsbara dokument, 21%. Andelen substitut, dvs andelen felaktigt avlästa siffror (tex en 1 uppfattad som en 7), uppskattades till ca 2%. Vid FoB 80 och FoB 85 hade andelen rejectmarkeringar sjunkit till 10,1% respektive 7,2%. Större delen av rejectmarkeringarna berodde på att granskningsorganen i kommunerna hade använt sig av olämpliga pennor vid kodningen.<sup>5</sup>

Från och med Folk- och Bostadsräkningen 1990 ska istället varje kommun själv dataregistrera informationen innan den sänds vidare till SCB. Anledningen till att ingen ny OCR-läsare har anskaffats är att SCB inte kommer att utnyttja en ny OCR-läsare tillräckligt för att ett inköp ska vara ekonomiskt lönsamt. Då en ny OCR-anläggning skulle kosta omkring fem miljoner kronor så måste den kunna utnyttjas fullt ut hela tiden. Detta innebär en praktisk omöjlighet då Folk- och Bostadsräkningarna nu registreras ute i respektive kommun.

På avdelningen för Areell statistik är man övervägande positiv till OCR-läsaren. Den har under drygt tio år använts för att genomföra ett externt uppdrag från lantbruksstyrelsen kallat "de objektiva skördeuppskattningarna". OCR-läsaren användes till detta fram till 1989. Numera ska en OMR-läsare (se sid 6 ff) användas istället. Nackdelar med detta teknikbyte är att hela blankettsystemet måste bytas ut. Det innebär att både ett nytt frågeformulär och en ny etikett som fästes vid de skördeprover som insamlas måste anpassas till den nya tekniken.

Även inom den interna administrationen användes förr OCR-maskinen för att registrera personalens arbetstid. Detta fungerade dock mindre bra då andelen rejects var ca 50%. Detta berodde till största delen på att OCR-läsaren var mycket känslig för hur siffror skrevs. Handskrivna siffror med "personlig touch" kunde inte läsas av automatiskt utan måste registreras manuellt.

---

<sup>4</sup> Information som gjort det möjligt att skriva detta avsnitt har vi fått från Lennart Österberg, D/DC, Systemsektionen, SCB, Örebro. Även personal på avdelningen för Areell statistik, avdelningen för Individstatistik (I/BE) och Personalenheten har hjälpt oss med information.

<sup>5</sup> se: "Datorstöd vid Datainsamlingen - Fallstudier" för en något utförligare presentation av OCR-användning i FoB

## **Teknik<sup>6</sup>**

Systemets basenhet är en minidator dit en OCR-läsare är ansluten.

I systemet kan även ingå:

- \* En till minidatorn kopplad skrivmaskin som styr och kontrollerar hela systemet.
- \* Radskrivare som kan användas till utskrift av lässtatistik och reject-filer.
- \* En korrigeringsenhet för manuell korrigering av läsfel som upptäcks i direkt anslutning till avläsningen.
- \* Bildskärmsterminaler för manuell dataregistrering.
- \* Utrustning för dataöverföring till andra datorsystem.

Avläsningen sker radvis uppifrån och ned. Dokumentet matas fram under läsaren en rad i taget. Överkanten och vänsterkanten används som utgångspunkter för att orientera data. Det är alltså viktigt att inte dessa kanter skadas genom tex slarvig avrivning. Avläsningen är även känslig för bla vikning, skrynkling, nedsmutsning och hålslagning av dokument.

De datamängder som inte kan avläsas, rejecten, och de som maskinen upptäcker innehåller logiska fel kan behandlas på tre olika sätt:

1. Korrigering i direkt anslutning till avläsningen med hjälp av korrigeringsenheten. Detta korrigeringssätt används främst vid radskrivarutskrift och vid mycket små datavolymer.
2. Blanketten går tillbaka till uppgiftslämnaren eller direkt till manuell dataregistrering. Detta används främst på blanketter med lite data.
3. Hela blanketten avläses och fält med reject eller logiska fel märks ut för att sedan korrigeras med hjälp av manuell dataregistrering. Denna metod används på blanketter innehållande mycket data.

## **Vinstpotential**

Denna teknik rationaliserar själva dataregistrerings- och databeredningsprocessen. Mycket monotont dataregistreringsarbete sköts av OCR-läsaren vilket snabbar upp processen och medför vinster ekonomiskt såväl som tidsmässigt. Då det manuella arbetsmomentet innebär risker för felregistrering innebär det att även datakvalitén kan förbättras då OCR-teknik används. Visst granskningsarbete kan även skötas av maskinen som kan sortera ut felaktiga dokument eller felaktiga delar av dokument.

---

<sup>6</sup> Otterberg, K.; (1980)

## **Målpopulation och Informationsbehov**

Siffrorna måste skrivas på ett visst sätt för att de ska kunna avläsas med OCR-teknik. De nyare OCR-läsarna är mindre känsliga för olika handstilar. Då dessa OCR-läsare inte ställer så hårda krav på att uppgiftslämnaren skriver korrekt så ökar naturligtvis OCR-teknikens användningsområde.

Vissa av de nyare OCR-läsarna kan även läsa av handskrivna bokstäver. Detta ger möjlighet till att registrera en större variation av svar med OCR-teknik. Registrerade sk öppna svar måste dock kodas separat.

## **1.2. OMR (Optical Mark Reading)<sup>7</sup>**

### **Inledning**

Denna teknik för optisk läsning, som innebär att en maskin registrerar blanketter, har i sin nuvarande form funnits tillgänglig i ca 5 år. På SCB har tekniken använts i ca 1 1/2 år.

För närvarande används på SCB OMR-teknik bland annat vid barnomsorgsundersökningen, vid registreringen av campingstatistik och vid genomförandet av de objektiva skördeuppskattningarna.

### **Teknik**

Blanketter och frågeformulär som ska läsas av med hjälp av OMR-teknik utrustas med ett id-nummer i form av en streckkod.

OMR-läsaren är utrustad med tre läshuvuden, ett som registrerar id-numret och två som läser av och registrerar övriga data på fram- och baksidan av dokumentet.

Om ett frågeformulär består av mer än ett blad rivs dessa isär innan avläsningen. Alla blad som hör ihop är förtryckta med samma id-nummer så att OMR-maskinen själv kan para ihop de dokument som hör ihop.

Dokumentet har förtryckta svarta markeringar på över/underkanten eller på sidorna, sk klockor, som talar om var det finns information att läsa av. Exempel på OMR-blanketter är olika spelblanketter tex V65- och lottokuponger.

OMR-läsaren kan inte läsa av röd text. Därför har de fält där markeringar, oftast ett lodrätt eller vågrätt streck, ska göras tryckts med röd text. Frågor och instruktioner kan däremot tryckas med svart text utan att OMR-läsaren uppfattar detta som en markering.

---

<sup>7</sup> Information som gjort det möjligt att skriva detta avsnitt har vi fått från Björn Örmeland, D/DB, SCB, Örebro.

## **Vinstpotential**

Många av OCR-läsarens fördelar i form av lönsamhets-, tids- och kvalitetsvinster erhålls även vid användning av OMR-tekniken. Förutom de ekonomiska vinster som rationaliseringen av datainsamlingen innebär så förbättras även arbetsklimatet för personalen då OMR-tekniken ersätter tråkigt och monotont dataregistreringsarbete. Datakvalitén kan också förbättras dels genom att OMR-läsaren kan sortera ut felaktiga dokument och dels genom att den manuella dataregistreringen som felkälla försvinner. Tekniken kan även medföra stora tidsvinster.

## **Målpopulation och Informationsbehov**

Frågeblankettens utseende blir givetvis något förändrad då OMR-tekniken ska användas. Detta borde dock inte innebära några restriktioner på lämpliga uppgiftslämnare. Det är mycket viktigt att uppgiftslämnaren har tillgång till klar och korrekt information om hur frågeformuläret ska fyllas i då det ska avläsas med OMR-teknik.

OMR-tekniken är mest lämplig då så få öppna svarsalternativ som möjligt är nödvändiga. Öppna svar måste kodas om manuellt om de ska registreras med OMR-teknik. Ett alternativ till detta är att manuellt dataregistrera de öppna svaren för att sedan samköra dessa med de OMR-registrerade svaren.

## **2. DATORSTÖD FÖR INTERVJUAREN**

I detta avsnitt presenteras datorstöd där uppgiftslämnaren är i direkt kontakt med en intervjuare. Denna kontakt kan ske genom telefon- eller besöksintervjuer. De gängse benämningarna på de två olika typerna av datorstöd vid intervjutillfället är:

- \* CATI (Computer Assisted Telephone Interviewing)
- \* CAPI (Computer Assisted Personal Interviewing)

CATI kan, liksom telefonintervjuer generellt, fungera i en decentraliserad såväl som i en centraliserad organisation. Den naturliga organisationen för CAPI är decentraliserad då den bygger på besöksintervjuer.

Vissa organisationer föredrar samma system för CATI och CAPI medan andra vill ha separata system då frågemetoderna kan vara olika. På SCB har ett integrerat system för DATorstödda Insamlingsmetoder, DATI, tagits fram (se sid 19 ff).

### **2.1. CATI (Computer Assisted Telephone Interviewing)**

#### **Inledning**

CATI-tekniken innebär att intervjuaren har en dator framför sig under arbetet med telefonintervjuerna. I datorn finns frågeformuläret inlagt så att intervjuaren kan läsa frågan på bildskärmen och knappa in svaret på tangentbordet. Datorn kan till viss del kontrollera svaret. Nästa fråga kommer sedan upp automatiskt på bildskärmen.

Den datorstödda datainsamlingens historia går tillbaka till 1971 då American Telegraph and Telephone (AT&T) sponsrade en CATI-undersökning, ledd av Chilton Research, USA. CATI utvecklades i början främst inom den privata sektorn vid marknadsundersökningar och var redan ett väl utvecklat system innan de statliga myndigheterna i USA började fundera på att använda CATI, William L. Nicholls II (1989)<sup>8</sup>.

I Sverige har den privata statistikproducenten SIFO använt sig av ett centraliserat CATI-system i ca 7 år.

Sedan 1971 har CATI utvecklats snabbt. Alois van Bastelaer, Frans Kerssemakers och Dirk Sikkel<sup>9</sup> (1988) menar dock att för lite forskning har gjorts kring metoder och kvalitén på informationen som erhålls med hjälp av CATI. Det främsta skälet till introduktionen av CATI tycks vara att skära ned kostnaderna jämfört med besöksintervjuer, vilket antyder att ett CATI-system både bör vara flexibelt och lätt att använda.

---

<sup>8</sup> Bureau of the Census (BC), USA

<sup>9</sup> Avdelningen för Arbets- och Lönestatistik, Centraal Bureau voor de Statistiek (CBS), Nederländerna



I USA har implementeringsperioden varit 20-25 år för de olika statliga institutionerna. Ofta saknades teknologin. Programvara måste utvecklas för att fylla de behov som fanns vid många statliga undersökningar. Den utvecklades i slutet av 70-talet. Efter 1980 då CATI började utvecklas internt startades också flera stora forskningsprogram för att undersöka CATI:s effekter på kostnaden och datakvalitén. Vid Bureau of the Census (BC) och Statistics Canada har först nu slutliga resultat producerats, Nicholls II (1989).

## **Teknik**

### Hårdvara och programvara

Ett CATI-system uppbyggt kring en stordator är ofta mycket resurskrävande om inte väntetider mellan frågorna ska bli oacceptabelt långa. Det är därför bättre att satsa på ett nätverk som länkar samman ett antal persondatorer. Då programvaran på CATI-området förbättras kontinuerligt så kan det vara mer ekonomiskt att köpa in programvara utifrån för den "lilla" statistikproducenten som inte har tillräckliga resurser för att hänga med i den teknologiska utvecklingen.

SIFO:s CATI-version var tidigare uppbyggt kring en stordator. Från och med 1 oktober 1990 kommer SIFO emellertid att gå över till att använda ett nyinköpt PC-nätverk med inköpt programvara. Enligt Ulf Ysander på SIFO är detta det mest lönsamma alternativet då "en programmerare kostar nästan lika mycket som hela programvaran inköpt utifrån". SIFO kan på detta sätt försäkra sig om att hela tiden ha tillgång till det senaste i fråga om programvara utan att behöva satsa egna resurser på forskning och utveckling.

### Programspråk

Ronald Biggar<sup>10</sup> (1988) upplevde programspråket som mycket tidsödande för användarna vid frågekonstruktionen av en National Health Interview Survey (NHIS). Då användes ett allmänt frågekonstruktionsprogram som inte täckte in alla de olika variabler som utredarna var intresserade av. Man upplevde även ett behov av att kunna samla all information i samma databas. Biggar har använt dBASE III för att lösa dessa problem.

För närvarande pågår ett forskningsprojekt i regi av Joint Centre for Survey Methods (London) där olika programspråk testas i sin lämplighet att användas vid konstruktion av frågeformulär, O'Muircheartaigh et al (1990). De programspråk som testas är CASES, QUANCEPT och BLAISE.

### Kontrollfunktioner hos programvaran

Datorn kan i de mest utvecklade systemen schemalägga påringningarna som ska göras och kontrollera intervjun. Den bestämmer vilka frågor som ska frågas och kontrollerar svaren så att de är konsistenta. Vid inkonsistenta svar uppmärksammas intervjuaren på detta och måste korrigera eller bekräfta svaret för att kunna gå vidare till nästa fråga.

---

<sup>10</sup> National Center for Health Statistics, USA

Konsistenskontroller har en stor betydelse i sammanhanget. De verkar i två riktningar, dels kontrollerar de att senare svar följer de tidigare uppgifterna, dels kontrollerar de att ett enskilt svar rör sig inom rimliga gränser.

Konsistenkontrollerna innebär att svaren kommer in korrekt under intervjun. De resulterande uppgifterna är alltså felfria och kan tom bli statistiskt behandlade direkt.

Konsistensen i data är nödvändig eftersom svaren leder till komplexa förgrenings- och hoppinstruktioner senare i intervjun. Intervjuarna förväntas utföra sitt arbete bättre då de själva slipper hoppa mellan frågorna och kan koncentrera sig på en fråga i taget.

Även andra funktioner kan läggas in i programmet, tex att alla relevanta frågor är besvarade, hur lång tid varje fråga tar att besvara och hur många korrigeringar varje fråga kräver.

Systemet måste även tillåta konflikt då "orimliga" resultat kan uppkomma. Att studera konflikterna och övriga registreringar av intervjuerna utvecklar metodologin

att utforma frågeformulär. Om tex inkonsistenser inträffar och många korrigeringar är nödvändiga kan detta indikera att ett begrepp är dåligt definierat.

CATI-insamling i tex Current Employment Statistics (CES) programmet<sup>11</sup> tillåter omedelbar redigering av rapporterade data. Olika kontroller är inbyggda i systemet. Dessa kontroller omfattar:

- \* Intervallkontroller som bestämmer om datauppgifterna är troliga om man jämför med andra uppgiftslämnare med liknande egenskaper.
- \* Konsistenskontroller som avgör om omöjliga kombinationer uppstår.
- \* Longitudinella kontroller som jämför rapporterad data med data från en föregående tidpunkt.

Dessa typer av kontroller är vanliga i panelundersökningar insamlade per post såväl som med CATI. Även med CATI fångas dock generellt bara stora fel upp då konsistenta men systematiska fel inte upptäcks, George Werking, Alan Tupek och Richard Clayton<sup>12</sup> (1988).

## **Organisationsstruktur**

Det största problemet med att implementera CATI vid statliga organ i USA tycks, enligt Nicholls II (1989), inte vara CATI-teknologin utan snarare organiseringen av intervjuerna geografiskt, dvs centralisering versus decentralisering.

Nicholls II anser vidare att i USA underlättas implementeringen av CATI om en övergång från en decentraliserad till en centraliserad organisation genomförs.

---

<sup>11</sup> se även "Datorstöd vid Datainsamlingen - Fallstudier"

<sup>12</sup> Bureau of Labor Statistics, USA

En övergång från en decentraliserad organisation till ett centraliserat CATI-system innebär dock vissa organisatoriska problem:

- \* Intervjuarkåren ute i landet måste minskas så att endast de intervjuare som ska sköta besöksintervjuerna är geografiskt utspridda. De får då ett större regionalt arbetsområde än tidigare, vilket innebär att mer resurser går åt till resor.
- \* Information om uppgiftslämnare måste kunna flyttas mellan det datoriserade CATI-systemet och intervjuerna ute i landet.

Fördelar med ett centraliserat CATI kan vara:

- \* Då varje intervjuare behöver en dator kan det bli problem vid sjukdom eller liknande om intervjuerna sker hemifrån.
- \* Intervjuerna delar samma hårdvara, program, urval och tekniska personal.
- \* Rekrytering, utbildning och handledning av intervjuerna samt vissa övriga procedurer kan lättare standardiseras.
- \* En centralisering isolerar inte intervjuerna i lika hög grad som om de vore geografiskt utspridda.

Fördelar med ett decentraliserat CATI kan vara:

- \* Intervjuaren delar ofta dialekt, omgivning och kulturell förståelse med uppgiftslämnaren.
- \* Variationen mellan telefon- och besöksintervjuer bibehålls. Detta kan dessutom förhindra att intervjuaren börjar uppfatta uppgiftslämnaren endast som en röst.

Vid SCB håller man på att införa ett decentraliserat CATI med en liten central telefongrupp i Örebro som främst sköter bortfallsuppföljning och dylikt.

## **Vinstpotential**

Datorstöd vid intervjutillfället har flera fördelar:

- \* Datainsamlingsprocessen kan göras mer centraliserad, kontrollerad och tidsmässigt mer effektiv.
- \* Samma databaserade filer kan användas till att definiera urvalet, fördela arbetet till intervjuare och föra statistik över intervjuer.
- \* Uppgifterna dataregistreras redan under intervjun vilket gör separat dataregistrering överflödigt.
- \* Samspelet mellan intervjuare och uppgiftslämnare kan bli bättre då intervjuaren inte behöver koncentrera sig på att ställa rätt fråga.
- \* Frågeformulären kan tillåtas att bli mer komplexa vilket kan föra med sig att ytterligare viktig information kan erhållas från uppgiftslämnarna.
- \* Möjligheten finns att kontrollera och korrigera felaktiga och inkonsistenta svar under intervjuns gång. Att korrigera fel är ofta omöjligt om det måste göras efteråt vid databehandlingen.

- \* Antalet korrigeringar och svarstider för varje fråga och intervjuare kan registreras.
- \* En kontroll att alla relevanta frågor är besvarade kan läggas in i programmet.

Mick Couper, Robert Groves<sup>13</sup> och Curtis Jacobs<sup>14</sup> (1990) hänvisar till studier som har visat att konsistensfelen minskar vid användandet av CATI men att det generellt tar längre tid med en CATI-intervju eftersom intervjuaren inte ser nästa fråga förrän han/hon först knappat in svaret<sup>15</sup>.

Intervjuaren förlorar till viss del översikten jämfört med PAPI då bara en bildskärm kan ses åt gången. Att gå bakåt i "datasidorna" kan dock ge intervjuaren en viss översikt samtidigt som svaren kan korrigeras.

För intervjuaren kan den ökade datoranvändningen dock innebära hälsorisker. De allergier och övriga problem som kan uppstå vid mycket terminalarbete är ännu ett relativt utforskat område.

Företag som vill investera i CATI-teknologi måste även räkna med kostsamma investeringar i hårdvara, programvara och utbildning som först betalar sig på sikt.

### **Målpopulation och Informationsbehov**

Datorstöd vid telefonintervjuer påverkar inte nämnvärt uppgiftslämnarens situation utan målpopulationen har här i stort sett samma begränsningar som vid telefonintervjuer utan datorstöd.

Genom att tekniken gör det möjligt att använda sig av mer komplexa frågeformulär kan ett större informationsbehov tillgodoses.

Intervjuarens inställning till datorer är avgörande för en lyckad implementering av tekniken. En intervjuares attityd kan även ha stor betydelse för uppgiftslämnarens reaktioner. Uppgiftslämnarens inställning till datorer är inte en fullt så avgörande faktor för undersökningens svarsfrekvens då intervjun sker per telefon och uppgiftslämnaren inte ser datorn.

---

<sup>13</sup> båda från University of Michigan, USA

<sup>14</sup> Bureau of Labor Statistics, USA

<sup>15</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; House, C.C.; (1985);

## **2.2. CAPI (Computer Assisted Personal Interviewing)**

### **Inledning**

CAPI innebär att intervjuaren vid besöksintervjuer och observationsundersökningar direkt knappar in data i en bärbar eller handhållen dator. Frågeformuläret är, liksom vid CATI-tekniken, inprogrammerat i en dator.

CAPI-teknologin har varit känd i över ett årtionde. Dess praktiska nytta har dock varit begränsad tills små men kraftfulla handhållna och bärbara datorer blev allmänt tillgängliga. Det var först 1987 som CAPI blev en livskraftig metod vid riksomfattande undersökningar<sup>16</sup>, Nicholls II (1989).

I motsats till vid CATI-teknikens införande var de statliga institutionerna i USA, enligt Nicholls II (1989), tidigt ute både som sponsorer av CAPI-test och som utvecklare av CAPI-system för egen användning. En annan skillnad mellan CATI och CAPI var att införandet av CAPI skedde mycket snabbare vid de statliga undersökningarna utan att föregås av lika mycket utvärderande forskning.

CAPI påverkar hela undersökningsprocessen från utformningen av intervjuformuläret till intervjuarens tillvägagångssätt. Frågeformulärets utseende vid CATI och CAPI liknar varandra. Utvärderingar och erfarenheter från CATI:s frågeformulär har därför ofta legat till grund för utformningen av frågeformuläret till CAPI, Nicholls II (1989).

I övrigt kan resultat vid CATI-studier inte ge mycket hjälp inför CAPI-användningen. Genom CATI-systemet kan datorn tala om i vilken ordning uppgiftslämnarna ska kontaktas osv. CAPI-systemet kräver en helt annan frihet för intervjuarnas kontakter med sina uppgiftslämnare då CAPI-metoden används vid besöksintervjuer. Större krav ställs även på datorerna, som bör vara lätta, då intervjuarna själva måste bära runt dem.

### **Teknik**

#### **Bärbara och Handhållna datorer**

I litteraturen förekommer i allmänhet två begrepp för hårdvaran vid datorstödda besöksintervjuer:

- \* "Laptops" vilket vi översätter med "bärbara datorer" som man bär i en väska med gängse tillbehör.
- \* "Hand-held computers" (HHC) vilket vi översätter med "handhållna datorer" som alltså kan bäras i den ena handen medan andra handen används till dataregistrering.

---

<sup>16</sup> se även "Datorstöd vid Datainsamlingen - Fallstudier"

Dessa två begrepp hör samman med två olika typer av besöksintervjuer:

- \* Situationer då intervjuerna är längre och intervjuarna sitter ned hemma hos uppgiftslämnaren och kan placera datorn i knät eller på ett bord. I denna situation används bärbara datorer.
- \* Korta intervjuer som går på ca 5-10 minuter och sker stående vid dörren eller i en affär för att kontrollera priser. I denna situation används handhållna datorer.

Undersökningar som har gjorts visar upp inkonsistenta slutsatser på hur tunga datorer intervjuarna orkar bära. I dessa tester har ofta återkommande klagomål från intervjuarna varit att maskinerna är tunga, skrymmande och att de stör den personliga kontakten med uppgiftslämnaren.<sup>17</sup>

De datorer som finns tillgängliga är generellt ganska tunga för att kunna användas stående och med en hand. Bärbara datorer är dessutom känsliga för vätska, mycket damm eller fukt och extrema temperaturer. Intervjuare, eller andra familjemedlemmar, kan också vara frestade att använda dessa datorer för andra ändamål. Detta kan innebära att uppgiftslämnarens konfidentialitet äventyras eller att program eller filer raderas. Bärbara datorer är dessutom ett åtråvärt byte att stjäla och risken ökar kanske då intervjuer sker vid områden med hög kriminalitet. Slutligen är inte bärbara datorer passande vid dörrintervjuer eftersom dessa datorer ska vila i ett horisontellt läge och måste ha stöd underifrån då de ska användas, Nicholls II (1989).

### HCR (Handwritten Character Recognition)

Nicholls II (1989) anser att HCR (ung "elektroniska anteckningsblock") kan vara ett alternativ till bärbara och handhållna datorer då intervjuerna sker stående. Deras plasthölje tål hårdare behandling och deras speciella ändamålsdesign torde motverka både stöld och otillåten användning. Bärbara HCR väger 1.35-1.80 kg och kan hållas med en hand medan text skrivs in med den andra. Batterier och andra tyngre delar kan bäras i en axelväska med sladdar till datorn. Eftersom HCR-tekniken försöker efterlikna PAPI så är utbildningen av intervjuarna enkel.

Data förs in genom kryss i rutor eller genom att skriva bokstäver i ett definierat område med en penna eller ett skrivstift som systemet sedan omvandlar till ASCII-koder. För bekräftande visas sedan svaret upp i vanlig dataskrift. Senare överförs färdiga intervjuer via telenätet till en central dator.

---

<sup>17</sup> För vidare diskussion se "Datorstöd vid Datainsamlingen - Fallstudier och Artikelförteckning"; Fallstudier om CAPI; Nicholls II, W.L.; (1988) och Sebestik, J. et al; (1988)

Nicholls II (1989) beskriver två typer av detta system:

1. DATAKECH Scriptwriter där intervjuaren använder en vanlig kulspetspenna. Intervjuaren skriver ner svar på ett papper som fästs över en tryckkänslig yta så att datorn kan registrera uppgifterna. Bara siffror, versaler och standardiserad kommatering kan användas. Dessa måste skrivas ner var och en inom en begränsad yta och på ett särskilt sätt.
2. LINUS Write-Top där en speciell stiftpenna används för att skriva på en glasskiva. Denna täcker i sin tur en LCD<sup>18</sup>-skärm där frågeformuläret eller blanketten syns. När stiftpennan berör ytan mörknar den så att det ser ut som om intervjuaren skrivit med en penna. Det skrivna omformas sedan till datorskrift. Denna teknologi tillåter en större variation i "skrivstil". Intervjuaren måste dock "lära" sin bärbara HCR hur han/hon skriver sina tecken. Detta kan ta upp till 2-2.5 timmar.

Dessa bärbara HCR verkar i nuvarande läge inte vara ett så passande alternativ till bärbara datorer i ett CAPI-system vid större undersökningar:

- \* Bildskärmen är alltför liten eller för svår att läsa från.
- \* HCR är ej utformade för att rymma register eller komplexa flersidesformulär.
- \* HCR innehåller inga sådana funktioner som automatiska förgrenings- och hoppinstruktioner eller konsistenskontroller.

HCR passar främst för enkel datainsamling med tex ett frågeformulär på en sida.

Kanske kan en dator utvecklas som har fördelarna från både den bärbara datorn och HCR genom att tex förenkla tangentbordet så att det fungerar bra att skriva på med en hand<sup>19</sup>. Kanske behövs det en mängd olika alternativ av hårdvara för flera olika situationer.

### Programvara

Enligt Nicholls II (1989) har programvaran hittills vid CAPI-användning i de flesta fall varit utformad för att passa CATI-teknologin eller besöksintervjuer vid företagsundersökningar. Vad som passar bra under dessa mer kontrollerade former kanske inte passar fullt så bra för användning vid den mer krävande och distraherande omgivningen i uppgiftslämnarens hem eller för intervjuer stående vid dörren. Generella riktlinjer för datorstödda frågeformulär, särskilt utformade för CAPI, har dock tagits fram men inte använts i stor utsträckning<sup>20</sup>.

---

<sup>18</sup> Liquid Crystal Display

<sup>19</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; Pallett, D.S., (red); (1987); Artikel skriven av Bryant

<sup>20</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; House, C.; (1985); House, C. et al; (1988) och Nicholls II, W.L. et al; (1987)



### Några Synpunkter

I det följande presenteras några artikelförfattares synpunkter på CAPI-teknologin.

Enligt Couper et al (1990) behövs för ett väl fungerande CAPI-system flera tekniska egenskaper såsom:

- \* Hårdvara och programvara som gör telekommunikation möjligt, tex modem, telefontillgång, kodningsprogram och telekommunikationsprogram.
- \* Förmåga hos centralkontoret att under telekommunikationen med intervjuarens dator kunna ställa problemdiagnos.
- \* Möjligheter till att skicka arbetsuppgifter från centralkontoret eller de sk "field supervisors" till intervjuarna och skicka färdiga intervjuer tillbaka från intervjuaren.
- \* Möjligheter till att kunna övervaka undersökningens framsteg centralt genom att data överförs från intervjuarnas datorer kontinuerligt.
- \* Ett meddelandesystem, dvs möjligheter för intervjuaren att skicka meddelanden om problem eller framsteg på fältet till "supervisors" eller centralkontor.
- \* En urvalsadministration där fall kan listas och sorteras, uppgifterna schemaläggas och undersökningens framsteg visas.
- \* Möjligheter att fördela om uppdrag från en intervjuare till en annan och ett system för protokollföring som förhindrar dubbelarbete eller försvunna fall.
- \* Funktioner som liknar CATI-systemet med presentation av frågor, kontroller, hopp, kommentarer och hjälpfunktioner.
- \* Möjligheter till sammanslagning av intervjuresultat för ev kodning, bearbetning och analys.

Enligt Biggar (1988) finns det 7 viktiga problem med programvaran som måste lösas innan CAPI fungerar som ett effektivt hjälpmedel vid insamlingen av information:

- \* Ett bra dataprogram för att konstruera frågeformulär av skiftande karaktär.
- \* Möjligheter till intervallredigeringar för all väsentlig data.
- \* Möjligheter till att fritt gå fram och tillbaka i frågeformuläret.
- \* Användarvänliga menyer och svarsredovisningar.
- \* Möjligheter till sk "pop up windows" för att kunna assistera intervjuaren med hjälpinformation.
- \* Rutiner för inträde och avslutning av intervjun, så att informationen blir ordentligt sparad.
- \* Snabbhet av programmet vid användningen.

Nicholls II (1989) menar att flera viktiga teknologiska frågor fortfarande är obesvarade angående CAPI:

- \* Vilka är de bästa metoderna för att överföra uppgifter och frågeformulär till CAPI-intervjuare och den kompletta intervjun tillbaka?
- \* Hur kan CAPI-programvaran göras mer bekväm för de speciella kraven som finns vid individ- och hushållsundersökningar?
- \* Med vilken typ av hårdvara borde CAPI-intervjuare utrustas?

### Kommunikation

National Analysts i USA har använt sig av en metod för att skicka uppgifter mellan intervjuare <-> lokala kontor <-> centrala kontor. Denna metod innebär att disketter skickas per post. Telekommunikation ansågs som alltför riskabelt då undersökningen var rikstäckande.<sup>21</sup>

Van Bastelaer et al (1988) redogör dock för hur CBS framgångsrikt har använt telekommunikation<sup>22</sup>. Otvivelaktigt är telenätet ett effektivt och snabbt sätt att överföra data. Nicholls II (1989) påpekar dock att telekommunikation kan utgöra ett hot mot konfidentialiteten.

En typisk telekommunikation kan se ut på följande sätt: Efter arbetets slut går intervjuaren hem och laddar batterierna till datorn. Den bärbara datorn kopplas också till en telefon och ett modem. På natten överförs insamlade data automatiskt till centralkontoret. Nästa morgon har intervjuaren en laddad dator med rent arbetsutrymme.

### **Organisationsstruktur**

CAPI-teknologin har också stor inverkan på organisationsstrukturen. Nicholls II (1989) anser att dessa problem måste lösas innan CAPI kan fungera i USA:

- \* "Supervisors" måste inte bara kunna hantera människor utan också ha förmågan att kunna lösa ev tekniska problem med CAPI.
- \* Utbildningen av intervjuare måste ta upp nya ämnen som tex användningen av CAPI-datorer och intervjuer med CAPI. (För detta kan man skicka instruktionspaket på diskett till intervjuarna.)
- \* Sättet att kontrollera kvalitén på intervjuarnas arbete måste förändras. "The supervisors" kanske måste få fram nya sätt att identifiera de intervjuare som behöver mer övning i CAPI hantering. (De genomförda intervjuernas summeringsresultat kontrolleras och kommer att ersätta granskning av enskilda intervjuer. Automatiska hopp i intervjun kan eliminera många vanliga typer av fel.)

---

<sup>21</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; Rothschild, B.B. et al; (1988)

<sup>22</sup> se även "Datorstöd vid Datainsamlingen - Fallstudier"; "Data Collection with Hand Held Computers: Contributions to Questionnaire Design"

Nicholls II (1989) pekar vidare på följder som kan uppstå i och med ett införande av ett CAPI-system:

- \* Behovet av regional personal som tar hand om intervjuerna minskar medan behovet av tekniskt kunnig personal ökar.
- \* Relationerna mellan de olika kontoren på olika geografisk nivå förändras i och med att nya kommunikationssätt införs.

## **Vinstpotential**

CAPI kombinerar fördelarna med besöksintervjuer och datorstödda intervjuer. Många av CATI-teknologins fördelar i form av konsistenskontroller, rationalisering av dataregistreringen, registrering av intervjuförloppet och möjligheter till komplexitetsökningar kan även återfinnas vid CAPI-användningen.

Nackdelarna med CATI-teknologin, som tex stora investeringskostnader, gäller även vid CAPI. Ytterligare en nackdel med CAPI-teknologin är datorernas vikt som ännu inte är tillräckligt låg. Detta kan på sikt ge upphov till fler arbetsskador.

## **Målpopulation och informationsbehov**

En begränsning av målpopulationen kan vara uppgiftslämnarnas eventuella obehag vid att deras svar registreras direkt i en dator. Många har tagit intryck av debatten kring möjligheterna till samkörning av dataregister och vad detta kan få för konsekvenser för den personliga integriteten. Visst stöd finns även i litteraturen att äldre tenderar att känna större oro inför den nya tekniken<sup>23</sup>. Ytterligare en begränsning är utformningen av hårdvaran. Datorns vikt, storlek och hållbarhet utgör idag begränsningar på CAPI:s användningsområde. I övrigt borde inte datorstöd innebära ytterligare restriktioner på möjliga målpopulationer än de begränsningar som redan finns vid besöksintervjuer.

Nicholls II (1989) anser att förhållanden i USA är CAPI-teknologin mest lämplig vid individ- och hushållsintervjuer. I USA är det mycket vanligt med klusterurval i denna typ av undersökningar vilket gör att uppgiftslämnarna finns inom ett begränsat geografiskt område. Vid företags- och organisationsundersökningar är inte alltid detta fallet. CAPI:s främsta användningsområde är vid besöksintervjuer. CAPI-tekniken kan i framtiden även komma till användning vid tex prisundersökningar. Detta förutsätter att mindre och lättare datorer kan utvecklas.

---

<sup>23</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; Gilroy, F.D., Desai, H.B.; (1986) och Czaja, S.J.; (1988)

Ett större informationsbehov kan tillgodoses då tekniken gör det möjligt att använda sig av mer komplexa frågeformulär. Tester har visat att nästan alla uppgiftslämnare accepterar datainsamling via CAPI och att de flesta erfarna intervjuare kan utbildas till att använda en dator<sup>24</sup>.

Enligt en fallstudie gjord i Nederländerna har det inte varit några problem för uppgiftslämnarna att acceptera CAPI, Van Bastelaer et al (1988). Intervjuarnas attityd till datorer är också positiv; majoriteten ser handhållna datorer som mer bekvämt än PAPI. Resultaten antyder att felkontroller i CAPI är nödvändigt. Dessa är ett steg i att utvärdera frågeformulärets kvalitet.<sup>25</sup>

### **2.3. DATI (DATorstödda Insamlingsmetoder)**

#### **Inledning**

DATI är ett system utvecklat på SCB för datorstödda besöks- och telefonintervjuer. Datorstödet fungerar på samma sätt som vid CATI- och CAPI-teknikerna. DATI har börjat användas på SCB vid den stora arbetskraftsundersökningen (AKU) som görs varje månad.

1984 gjordes de första proven med datorstödda intervjuer i SCB:s regi. Undersökningar där test med DATI pågår är Hushållens Inköpsplaner (HIP), Konsumentprisindex (KPI) och Partisynptatiundersökningen (PSU), SCB (1990).

#### **Teknik**

Om telefon- och besöksintervjuer ska kunna utföras i en decentraliserad miljö så behövs:

- \* En bärbar datorutrustning.
- \* Ett fungerande datakommunikationssystem.
- \* Hjälpmedel och övervakning som är anpassade till de krav som en decentraliserad organisation ställer.

DATI erbjuder möjligheter att integrera de olika momenten som ingår från frågekonstruktion till rapport:

- \* Möjligheter finns att använda ett speciellt utarbetat program för Interaktiv Fråge-Konstruktion (IFK). Även mycket komplicerade frågeformulär med många hoppinstruktioner och logiska och aritmetiska beräkningar kan lätt skapas då IFK-programmet är mycket användarvänligt.

---

<sup>24</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; Birkett, N.J.; (1988); CBS; (1987); Nicholls II, W.L.; (1988); Rice Jr et al; (1988); Rothschild, B.B. et al; (1988) samt Sebestik et al; (1988)

<sup>25</sup> se även "Datorstöd vid Datainsamlingen - Fallstudier"; "Data Collection With Hand Held Computers: Contributions to Questionnaire Design"

- \* Konsistenskontroller och automatiska hopp- och förgreningsinstruktioner kan ingå i systemet. Även registrering av intervjuförloppet, textantalet korrigeringar och svarstid för varje fråga, kan läggas in i programmet.
- \* Intervjuarna utrustas med fältdatorer som nattetid står i förbindelse med en centraldator via telenätet. Intervjuaren tömmer då sin dator på dagens information samtidigt som centraldatorn skickar iväg morgondagens uppgifter.
- \* Utredaren kan via centraldatorn få tillgång till färsk information och kan under utredningens gång lätt ta fram tabeller för att se på de delresultat som kan behövas. Slutrapporten skrivs i samma system som även omfattar en ordbehandlingsfunktion.

Fältdatorn som används i DATI bör ha följande egenskaper, SCB (1987):

- \* Vara bärbar med stor bildskärm och fullständigt tangentbord.
- \* Inte väga mer (högst 2,5 kg) än att tangentbordsdelen med bärremmar ska kunna användas under huvuddelen av en arbetsdag.
- \* Remmar med tryckavlastning och beslag för att kunna bära tangentbordsdelen under längre arbetspass (mer än ca 10 minuter).
- \* Bildskärmsdelen ska kunna stå stabilt för sig själv men även lätt kunna sammanfogas med tangentbordsdelen till en stabil enhet.
- \* Lätt kunna omställas så att endast tangentbordsdelen med centralenhet och mindre bildskärm kan användas och transporteras separat utan att ta skada.
- \* Tåla att användas och transporteras mellan många skiftande miljöer.
- \* God slitstyrka som medger användning under 2000 timmar per år.
- \* Inbyggt modem för datakommunikations- och telefonanslutningar.
- \* Batterier som garanterar oavbruten drift i minst 8 timmar och återuppladdas efter högst en natts uppladdning.
- \* Bildskärm som lätt kan läsas av även i svåra ljusförhållanden.
- \* Tangentbord som medger enkel alfanumerisk inmatning med endast en hand.
- \* Bra sekretesskydd.

## **Organisationsstruktur**

DATI används för närvarande på SCB som en variant på ett decentraliserat CATI. Intervjuare ute i landet har en dator som de kopplar in på telenätet nattetid. Även en liten central stödjande telefongrupp finns på SCB i Örebro som använder DATI vid sina intervjuer.

Meningen är dock att DATI i framtiden även ska användas som stöd vid besöksintervjuer (en variant av CAPI).

## **Vinstpotential**

Många av CATI-teknologins fördelar i form av konsistenskontroller, rationalisering av dataregistreringen, registrering av intervjuförloppet och möjligheter till komplexitetsökningar kan även återfinnas vid DATI-användningen. Datorstöd kan på sikt minska kostnaderna då själva informationsprocessen rationaliseras och mycket dubbelarbete försvinner. Kostnaden vid införandet av ett sådant omfattande system som DATI är dock betydande, vilket naturligtvis även innebär stora risker.

## **Målpopulation och Informationsbehov**

I fallstudien om DATI användning vid AKU<sup>26</sup> har inget större motstånd från uppgiftslämnarna kunnat upptäckas. Detta "bekräftar" påståendet att datorstöd vid telefonintervjuer inte borde innebära någon ytterligare begränsning, än telefonintervjuer generellt, av målpopulationen. Då visst stöd finns i litteraturen<sup>27</sup> för att äldre tenderar att ha en mer misstänksam attityd till datorer och AKU endast omfattar personer under 64 år så kan detta innebära att motviljan generellt sett är något större. När DATI används vid besöksintervjuer kommer uppgiftslämnarens attityd till datorer att bli mer betydelsefull.

Vid observationsundersökningar (tex KPI-undersökningen) utgör även hårdvaran en begränsning (se CAPI sid 18).

Då automatiska hopp- och förgreningsinstruktioner underlättar användningen av mer komplexa frågeformulär kan ett större informationsbehov tillgodoses.

---

<sup>26</sup> se "Datorstöd vid Datainsamlingen - Fallstudier"; DATI

<sup>27</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"; Gilroy, F.D., Desai, H.B.; (1986) och Czaja, S.J.;(1988)

### **3. DATORSTÖD FÖR UPPGIFTSLÄMNAREN**

Datorstöd vid uppgiftslämnandet har förändrat förutsättningarna för och utseendet på de intervjutekniker som traditionellt har använts. I vissa fall har automatiken ersatt intervjuarens direkta kontakt med uppgiftslämnaren. Uppgiftslämnaren för in data direkt utan att intervjuaren fungerar som mellanhand. De datorstöd vi kommer att behandla är:

- \* TDE (Touchtone Data Entry)
- \* CASAQ (Computer Assisted Self-Administered Questionnaire)

#### **3.1. TDE (Touchtone Data Entry)**

##### **Inledning**

För att upprätthålla det som uppnåts under CATI (bättre datakvalité och ökad punktlighet), men komma ifrån CATI:s högre kostnader, initierades forskning kring TDE för att utveckla en alternativ telefoninsamlingsmetod med en lägre enhetskostnad, Werking et al (1988).

Genom TDE kan uppgiftslämnaren själv ringa in sina svar genom en tonvals-telefon. En dator registrerar de siffror som trycks in efter det att en datasimulerad röst ställt en fråga och gett instruktioner om svaret.

TDE-teknologin har hittills främst implementerats i privata organisationer, tex hos banker och bokklubbar, men är ännu en ganska ovanlig teknik vid statistikproduktion.

Några exempel finns dock på TDE-användning vid statistikproduktion:

- \* I Sverige har Svenska Potatisodlarnas Riksförbund använt sig av TDE-teknik för att snabbt kunna lämna prisbesked till sina medlemmar.
- \* I USA har Bureau of Labor Statistics (BLS) i sitt Current Employment Statistics (CES) program testat ett integrerat CATI/TDE-system<sup>28</sup>.

---

<sup>28</sup> se "Datorstöd vid Datainsamlingen - Fallstudier", CES-Programmet

## **Teknik**

Chester Ponikowski, Kennon Copeland och Sue Meily<sup>29</sup> (1989) beskriver hur insamlingsprocessen till BLS gick till vid CES-programmet med det integrerade CATI/TDE-systemet:

1. Uppgiftslämnarna erhåller ett frågeformulär som sänds ut per post där de ombeds svara via sin tonvalstelefon.
2. Uppgiftslämnarna tar fram sina uppgifter och ringer upp TDE-systemet på ett avgiftsfritt telefonnummer. Frågorna läses upp av en datasimulerad mänsklig röst varefter uppgiftslämnaren knappar in sina data.
3. Uppgiftslämnare som ej har ringt in sina data identifieras och rings upp. De kan alternativt påminnas att ringa in sina data eller också samlas uppgifterna in direkt av en intervjuare per telefon (här kan då CATI användas som ett stöd).
4. Data som rings in i TDE-systemet bearbetas, misstag identifieras och eventuella kompletteringssamtal görs.

TDE-systemet innebär:

- \* Sökande av giltiga uppgiftslämnare baserat på en fil med alla uppgiftslämnares id-nummer (telefonnummer, personnummer eller dylikt).
- \* Spårandet av uppgiftslämnare som har ringt in sina data mer än en gång om han/hon tex har upptäckt ett fel efter första samtalet.
- \* Variation i frågesetet beroende på vilken delgrupp uppgiftslämnaren tillhör.
- \* En datasimulerad röst som läser tillbaka alla svar för verifikation till uppgiftslämnaren.
- \* Erbjuder en begränsad intervjuöversikt (tex uppgiftslämnaren har tre tillfällen att rapportera en datauppgift korrekt och krävs på att bekräfta alla svar så att misstag kan korrigeras direkt).
- \* Lagring av data, datum, starttid, och sluttid för varje samtal.

Ett alternativ till att använda sig av datasimulerade röster är att sända ut frågeformulär på vanligt sätt och sedan be uppgiftslämnaren att själv knappa in svaren på sin tonvalstelefon.

Systemet kräver en kraftfull mikrodator och växlar som tillåter flera telefonlinjer.

Ett avgörande hinder för storskalig användning av TDE är kravet på att uppgiftslämnaren har tillgång till en tonvalstelefon. Tillgången på tonvalstelefoner ökar dock samtidigt som teknologin för talidentifiering förbättras snabbt.

Enligt Shanks (1989) så pågår just nu även en intensiv forskning för att avancerad talteknologi och även grafik ska ingå i dataprogrammen för frågor direkt till uppgiftslämnarna (se även avsnitt 4.1. sid 29).

---

<sup>29</sup> Bureau of Labor Statistics, USA



## **Organisationsstruktur**

Organisationen vid TDE-användning liknar i sin uppbyggnad ett centraliserat CATI. Skillnaderna är dels att uppgiftslämnaren själv väljer att ringa upp systemet och dels att samtalen tas emot av en dator.

## **Vinstpotential**

TDE minskar undersökningskostnaderna genom att eliminera arbetsintensiva aktiviteter som brevhantering och knapptryckning. Kostnaderna minimeras också då systemet fungerar 24 timmar varje dag utan att en intervjuare behövs.

Ponikowski et al (1989) skriver att användningen av TDE-systemet kan spara in på flera ställen; lägre kostnader, förbättrad punktlighet och förbättrad kontroll. TDE är också ett bekvämt system för uppgiftslämnaren.

TDE ger utrymme till kontroller i frågeformuläret à la CATI vilket kan förbättra datakvalitén.

Ponikowski et al (1989) nämner vidare fördelar och vissa begränsningar med TDE på flera olika avseenden.

Fördelarna med TDE jämfört med insamling genom postenkäter är:

- \* Reducerade kostnader för frankering och kuvert.
- \* Eliminerar posthantering och avprickningsrutiner.
- \* Eliminerar tiden som förloras vid posthanteringen.
- \* Eliminerar huvudsakligen manuell dataregistrering.

Fördelarna med TDE jämfört med CATI-insamling är:

- \* Reducerade personalkostnader (jämfört med att hålla en stor CATI-arbetsstyrka).
- \* Eliminerar schemaläggning och upprepade kontakter med uppgiftslämnaren.

Fördelarna med TDE-insamling för uppgiftslämnarna är:

- \* Det är bekvämt, de kan ringa in data när som helst.
- \* Det är en snabb (tar ofta mindre än två minuter) och lätt metod för uppgiftslämnaren att rapportera in sina data.

Begränsningar associerade med TDE är:

- \* Kravet att uppgiftslämnaren har en tonvalstelefon.
- \* Bortfallsfrekvensen kan öka samtidigt som anledningen till bortfall blir okänd då uppgiftslämnaren själv bestämmer om kontakt skall inledas (samma problem som med insamling genom postenkäter).
- \* TDE-system har inte använts i någon större utsträckning av statistikproducenter. Behov finns därför av vidareutveckling av teknologin.

## **Målpopulation och Informationsbehov**

TDE passar bäst vid undersökningar där företag, institutioner och organisationer utgör uppgiftslämnarna. Dessa kräver ofta bara korta numeriska svar från samma uppgiftslämnare med återkommande intervall. För dessa uppgiftslämnare kan de nya teknikerna kanske spara mycket tid. Metoderna är däremot svårare att tillämpa vid individ- och hushållsintervjuer.

Werking et al (1988) påpekar att nuvarande forskningsresultat antyder att användandet av TDE och datasimulerade röster kan användas i en mängd undersöknings-situationer. Särskilt vid panelundersökningar där uppgiftslämnaren rapporterar samma uppgifter regelbundet kan en tonvalstelefon vara bekvämt för uppgiftslämnaren. De kan ringa in sina data när som helst. TDE kan även vara lämplig för sk "snabbvarsundersökningar" där möjlighet ges till svar genom tonvalstelefon. Det kan också vara ett effektivt informationsredskap för ledningen i en decentraliserad organisation. Intervjuare och "field supervisors" kan rapportera färdiga fall dagligen.

## **3.2. CASAQ (Computer Assisted Self-Administered Questionnaire)**

### **Inledning**

CASAQ-tekniken innebär att uppgiftslämnaren svarar på ett frågeformulär vid sin egen dator. De färdiga svaren sänds sedan tillbaka till undersökningsledningen på en diskett eller genom telekommunikation.

Denna teknik har ännu inte börjat användas av det stora flertalet statistikproducenter. Metoden ligger ännu på utvecklingsstadiet. Ett exempel där CASAQ används i produktionen är dock CBS-undersökningen Survey on Fire Reports, Wouter Keller, Kees-Jan Metz och Jelke Bethlehem<sup>30</sup> (1989).

CASAQ har även testats av tex:<sup>31</sup>

- \* Energy Information Agency (EIA) undersöker användningen av CASAQ i sitt Petroleum Supply Reporting System.
- \* Carneige-Mellon universitetet har genomfört en studie där CASAQ och postenkäter (av papper) har jämförts.

### **Teknik**

Frågeformulärets egenskaper liknar till stor del det som används vid CATI-, CAPI- och DATI-intervjuer. Utrymme finns för att inkludera konsistenskontroller samt komplexa hopp- och förgreningsinstruktioner. Vid CASAQ-användning står dock ingen intervjuare mellan uppgiftslämnarnas svar och registrerade data.

---

<sup>30</sup> Automation Department, CBS, Nederländerna

<sup>31</sup> se även "Datorstöd vid Datainsamlingen - Fallstudier"; CASAQ

## Kommunikation

Kommunikationen sker, liksom för CAPI, genom telekommunikation eller genom att disketter skickas per post.

Telekommunikation har vissa egenskaper som inte erhålls vid diskettinsamling:

- \* Frågeformuläret kan skickas ut till uppgiftslämnarna automatiskt.
- \* Frågeformuläret kan nå ut till geografiskt avlägsna uppgiftslämnare på kort tid.
- \* Sparar tid som annars skulle gå åt till posthantering.
- \* Uppgifterna registreras direkt i statistikproducentens dator utan särskild diskethantering. Den kan vara relativt omfattande vid stora undersökningar.
- \* Undersökningar som är baserade på en viss händelse kan underlättas, tex kan man i en konsumentundersökning automatiskt få fram inköp som gjorts av personer som använder elektroniska datakataloger, Sara Kiesler och Lee Sproull<sup>32</sup> (1986).

Vid diskettinsamling behöver däremot inte uppgiftslämnaren ha möjlighet att koppla in sin dator på telenätet. Diskettinsamling är mest fördelaktig då mycket data ska hämtas från relativt få uppgiftslämnare eftersom styckkostnaden är given.

## Exempel på CASAQ-system

Keller et al (1989) beskriver vissa rutiner på CBS: disketter med frågeformuläret skickas till uppgiftslämnarna, de svarar på frågorna vid sin dator där svaren granskas efter vissa kriterier och skickar tillbaka disketten med svaren till CBS.

EIA undersöker användningen av CASAQ i sitt Petroleum Supply Reporting System. Det visade sig att i ett litet urval av uppgiftslämnare hade de flesta tillgång till en dator. Disketter skickades ut till varje företag. Disketterna innehöll följande:

- \* Ett frågeformulär som skulle besvaras varje månad.
- \* Menydrivna procedurer som skulle hjälpa uppgiftslämnaren att få fram den information som behövs från andra filer eller länkade datorer.
- \* Förprogrammerade rutiner för att ringa upp EIA-kontoret då frågeformuläret är komplett, knappa in passordet och överföra uppgifterna till EIA:s dator.

---

<sup>32</sup> Carneige-Mellon University, USA

På SCB används CASAQ (diskettinsamling) vid fiskeri- och skolstatistiken. Då många av uppgiftslämnarna inte har samma datasystem som SCB uppstår en rad problem med konverteringen av data. Även utbildning av uppgiftslämnarna krävs för att systemet ska fungera tillfredsställande. Detta har lett till att användningen av CASAQ inte har haft så stor genomslagskraft. Vid tex insamlingen av slutbetygen från gymnasieskolorna har endast ett fåtal disketter kommit in och flertalet av de insända disketterna har innehållit många fel<sup>33</sup>.

DAVID (DATainsamling VIa Diskett) är en datainsamlingsteknik som har utformats på SCB i en första version, men ännu ej har provats i produktion. DAVID kan användas vid dataregistrering med samtidig granskning. Det färdiga formulärprogrammet kan kopieras och distribueras utan krav på några licenser då ett licensfritt programspråk används. Endast vid formulärkonstruktionen är tillgång till en C-kompilator nödvändig, Lennart Augustsson, Hans Irebäck och Bengt-Åke Jönsson<sup>34</sup> (1990).

### **Vinstpotential**

Liksom telefon- och besöksintervjuer är CASAQ mycket flexibelt. Frågeformuläret kan ha en formell eller en informell stil. Det kan inkludera förklarande material, "prompts", felkontroller, menyer, förgreningar och hopp, Kiesler et al (1986).

En annan fördel är att CASAQ lätt kan standardiseras liksom postenkäter (av papper).

När uppgiftslämnaren väl har tillgång till en dator kan de relativt låga marginalkostnaderna innebära att CASAQ blir ett ekonomiskt mer lönsamt alternativ till telefonintervjuer, besöksintervjuer eller utskick av postenkäter.

Kiesler et al (1986) menar att CASAQ kan reducera kostnaderna eftersom det automatiserar övergången från rådata till data som är läsbart i datorn. Systemet kan kombinera fördelarna med intervjuer (tex "prompts" och komplexa hopp mellan frågorna) tillsammans med fördelarna med postenkätsundersökningar (tex standardisering och anonymitet).

CASAQ-undersökningen eliminerar behovet av att en person (tex en intervjuare) eller en teknologi (tex OMR) står mellan uppgiftslämnaren och datorn som lagrar och analyserar data. Vissa frågor med öppna svar måste dock fortfarande kodas manuellt. Denna uppgift kan underlättas genom datoriserade textsökningsprogram.

Kiesler et al (1986) skriver att jämfört med besöks- och telefonintervjuer minskar postenkäter prestigebiasen, dvs uppgiftslämnarens önskan att "vara till lags". Uppgiftslämnarens vilja att enbart visa upp sina goda sidor tenderar att reduceras eftersom postenkäter är relativt anonyma. En nackdel med postenkäter är dock att svarsfrekvensen minskar i förhållande till besöks- och telefonintervjuer.

---

<sup>33</sup> Enligt Annica Göransson, I/UK, SCB, Örebro

<sup>34</sup> U/ADB, SCB, Örebro

En CASAQ-undersökning borde därför, menar Kiesler et al, på samma sätt som en undersökning utförd med hjälp av postenkät minska prestigebiasen. Eftersom ingen intervjuare medverkar till att data insamlas kan uppgiftslämnaren inte heller påverkas av intervjuarens person. Det som fortfarande kan påverka uppgiftslämnaren är dock faktorer som frågeformulärets utformning (grafik, frågekonstruktion etc) och ev medföljande brevs utseende (brevhuvud, papperskvalitet etc).

CASAQ-undersökningen är, liksom telefon- och besöksintervjuer, interaktiv. Sådana moment som "prompts" kan öka uppgiftslämnarnas uppmärksamhet med resultatet att deras engagemang blir större än vad det blir vid en postenkät av papper.

### **Målpopulation och Informationsbehov**

CASAQ är mest lämpad till en målpopulation som utgörs av företag, institutioner och organisationer som har tillgång till datorer. För dessa uppgiftslämnare kan denna teknik kanske spara mycket tid samtidigt som tidpunkten för uppgiftslämnandet är självvald. Metoderna är däremot svårare att tillämpa vid individ- och hushållsintervjuer då uppgiftslämnaren måste ha tillgång till en dator.

CASAQ innebär nästan inga begränsningar på informationsbehovet då systemet är mycket flexibelt och även mycket stora datamängder lätt kan samlas in.

Uppgiftslämnarens inställning till datorer är mycket viktig. Då de uppgiftslämnare som kan komma ifråga vid en CASAQ-undersökning måste ha tillgång till en dator och alltså redan borde vara vana datoranvändare är detta kanske ett mindre problem.

## **4. DATORSTÖD I FRAMTIDEN?**

### **4.1. Talteknologi (Röstsimulering och Talidentifiering)**

National Bureau of Standards (NBS) har nyligen publicerat en rapport i USA om datoranvändningen inför 2000-talet för BC. I rapporten skrivs att CATI och CAPI som de nu fungerar bara är övergångsmetoder för datainsamling. Det är istället talet som är det naturligaste sättet att samla in data på. Talteknologi inkluderar både röstsimulering och talidentifiering, Nicholls II (1989).

NBS föreslog två mål för användningen av talteknologi vid datainsamlandet:

- \* Helt automatiserade telefonintervjuer.
- \* Utveckla talteknologi som en hjälpfunktion vid besöksintervjuer och observationsundersökningar. Beroendet av intervjuarens maskinskrivningsvana och tangentbordets utformning reduceras.

Talteknologi prövas främst i institutions- och företagsundersökningar. Användning av talteknologi vid individ- och hushållsundersökningar är svårare.

Nicholls II (1989) nämner några exempel som härstammar från den privata sektorn:

- \* Feinberg (1988)<sup>35</sup> har beskrivit tester av att använda talidentifiering för att registrera öppna svar i marknadsundersökningar utförda med hjälp av CATI-teknik. Uppgiftslämnarens tillåtelse behövs naturligtvis. De digitaliserade svaren kan sedan korsklassificeras med uppgiftslämnarens karakteristik.
- \* Stonecash (1988)<sup>36</sup> beskriver en av flera politiska opinionsundersökningar i USA som använder fullt automatiserade telefonintervjuer med automatisk uppringning och automatisk registrering av svaren. Den låga kostnaden vid dessa fullt automatiserade system uppmuntar antagligen till vidareutveckling, men svarsfrekvensen har inte nått upp till den vid de flesta statliga undersökningar. Det avgörande vid användning av talteknologi vid individ- och hushållsundersökningar verkar vara uppgiftslämnarens vilja att medverka.

Enligt Mitchell Waldrop (1988) så finns numera en prototyp kallad SPHINX framtagen av Kai-Fu Lee<sup>37</sup> som klarar av att känna igen mänskligt tal med upp till 96% säkerhet. När artikeln skrevs kände datorn igen 997 ord. I framtiden kommer antagligen forskarna att använda sig av sk "artificiell intelligens" för att ytterligare förbättra möjligheterna att få en dator att "förstå" mänskligt tal. Lee's rådgivare Raj Reddy anser det möjligt att talteknologi kommer att ingå som standard i de flesta persondatorer någon gång i framtiden, Waldrop (1988).

---

<sup>35</sup> se "Datorstöd vid Datainsamlingen - Artikelförteckning"

<sup>36</sup> ibid

<sup>37</sup> Carnegie-Mellon University, USA

## **SAMMANFATTNING**

---

### **Teknik**

---

|      |                                                                                                                                                                                                                                                                                                                                                                                                                                                                                                 |
|------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| OCR  | <ul style="list-style-type: none"><li>* Avläser och registrerar blanketter.</li><li>* Nyare läsare kan både läsa handskrivna siffror och bokstäver.</li></ul>                                                                                                                                                                                                                                                                                                                                   |
| OMR  | <ul style="list-style-type: none"><li>* Avläser och registrerar streck- och kryssmarkeringar på blanketter.</li><li>* Kan samtidigt avläsa både fram- och baksida samt blankettens id-nummer.</li></ul>                                                                                                                                                                                                                                                                                         |
| CATI | <ul style="list-style-type: none"><li>* Denna teknik används vid telefonintervjuer där intervjuaren skriver in svaren direkt i datorn.</li><li>* Frågeformuläret innehåller konsistenskontroller samt automatiska hopp- och förgreningsinstruktioner.</li><li>* Vid en decentraliserad organisation skickas de data-registrerade uppgifterna till centralkontoret med hjälp av disketter eller telekommunikation.</li></ul>                                                                     |
| CAPI | <ul style="list-style-type: none"><li>* Denna teknik används vid observationsundersökningar och besöksintervjuer där intervjuaren skriver in svaren direkt i datorn.</li><li>* Frågeformuläret innehåller konsistenskontroller samt automatiska hopp- och förgreningsinstruktioner.</li><li>* Tekniken kräver små kraftfulla bärbara eller handhållna datorer.</li><li>* De dataregistrerade uppgifterna skickas till centralkontoret med hjälp av disketter eller telekommunikation.</li></ul> |
| DATI | <ul style="list-style-type: none"><li>* Denna teknik används vid observationsundersökningar samt besöks- och telefonintervjuer där intervjuaren skriver in svaren direkt i datorn.</li><li>* Frågeformuläret innehåller konsistenskontroller samt automatiska hopp- och förgreningsinstruktioner.</li><li>* Tekniken kräver små bärbara och/eller handhållna datorer.</li><li>* De dataregistrerade uppgifterna skickas till centralkontoret med hjälp av telekommunikation.</li></ul>          |

---

## Teknik

---

- |                     |                                                                                                                                                                                                                                                                                                                                                                                             |
|---------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| TDE                 | <ul style="list-style-type: none"><li>* Uppgiftslämnarna ringer själva in sina uppgifter genom att trycka in koder på en tonvalstelefon. Därmed data-registreras svaren automatiskt.</li><li>* Frågorna och bekräftelser på svaren läses upp med en datasimulerad röst.</li><li>* Frågeformulärets utseende kan varieras beroende på uppgiftslämnare.</li></ul>                             |
| <hr/>               |                                                                                                                                                                                                                                                                                                                                                                                             |
| CASAQ               | <ul style="list-style-type: none"><li>* Uppgiftslämnaren svarar på ett frågeformulär vid sin egen dator.</li><li>* Frågeformuläret innehåller konsistenskontroller samt automatiska hopp- och förgreningsinstruktioner.</li><li>* De dataregistrerade uppgifterna skickas till centralkontoret med hjälp av disketter eller telekommunikation.</li></ul>                                    |
| <hr/>               |                                                                                                                                                                                                                                                                                                                                                                                             |
| Röst &<br>Talteknik | <ul style="list-style-type: none"><li>* Denna teknik är ännu endast på utvecklingsstadiet.</li><li>* Uppgiftslämnaren rings upp eller ringer själv upp och talar in sina svar. Svaren registreras av en dator som känner igen och förstår mänskligt tal.</li><li>* Denna teknik kan också vara ett hjälpmedel till CAPI-tekniken där intervjuaren kan tala in sina observationer.</li></ul> |
-



---

### Organisationsstruktur

---

OCR      \* Passar i en organisation där all dataregistrering av postenkäter sker på samma ställe.

---

OMR      \* Passar i en organisation där all dataregistrering av postenkäter sker på samma ställe.

---

CATI      \* Centraliserad organisation:  
+ Intervjuarna delar samma hårdvara, programvara, urval och tekniska personal.  
+ Rekryteringen, utbildningen och handledningen av intervjuarna kan lätt samordnas.  
+ Inget speciellt system för kommunikation mellan fältdatorer och centraldatorn behövs.  
\* Decentraliserad organisation:  
+ Intervjuarna kan få variation i sitt arbete då telefon- och besöksintervjuer kan kombineras.  
+ Intervjuaren delar oftast kultur och dialekt med sin omgivning.

---

CAPI      \* Behovet av tekniskt kunnig personal i den decentraliserade organisationen ökar.

---

DATI      \* Behovet av tekniskt kunnig personal i den decentraliserade organisationen ökar.

---

TDE      \* Passar i en organisation där dataregistrering sker på samma ställe.

---

CASAQ    \* Dataregistreringen sker decentraliserat, medan analys och bearbetning sker centralt.

---

---

### Vinstpotential

---

|      |                                                                                                                                                                                                                                                                                                                                             |
|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| OCR  | <ul style="list-style-type: none"><li>* Detta är en snabb och effektiv teknik för att registrera stora datavolymer.</li><li>* Tekniken finns att köpa färdigutvecklad.</li></ul>                                                                                                                                                            |
| OMR  | <ul style="list-style-type: none"><li>* Detta är en snabb och effektiv teknik för att registrera stora datavolymer.</li><li>* Tekniken finns att köpa färdigutvecklad.</li></ul>                                                                                                                                                            |
| CATI | <ul style="list-style-type: none"><li>* Kontroller och hopp- och förgreningsinstruktioner borgar för att datakvaliteten förbättras.</li><li>* Arbetsmoment vid dataregistrering och granskning rationaliseras bort.</li><li>* Ett införande av denna teknik kräver stora investeringar i hårdvara, programvara och utbildning.</li></ul>    |
| CAPI | <ul style="list-style-type: none"><li>* Kontroller och hopp- och förgreningsinstruktioner borgar för att datakvaliteten förbättras.</li><li>* Arbetsmoment vid dataregistrering och granskning rationaliseras bort.</li><li>* Ett införande av denna teknik kräver stora investeringar i hårdvara, programvara och utbildning.</li></ul>    |
| DATI | <ul style="list-style-type: none"><li>* Kontroller och hopp- och förgreningsinstruktioner borgar för att datakvaliteten förbättras.</li><li>* Arbetsmoment vid dataregistrering och granskning rationaliseras bort.</li><li>* Ett införande av denna teknik kräver stora investeringar i hårdvara, programvara och utbildning.</li></ul>    |
| TDE  | <ul style="list-style-type: none"><li>* Möjligheter till att använda automatiska kontroller och hopp- och förgreningsinstruktioner borgar för att datakvaliteten kan förbättras.</li><li>* Dataregistreringsprocessen effektiviseras.</li><li>* Uppgiftslämnaren kan själv välja en passande tidpunkt för att ringa in sina data.</li></ul> |

---

---

### **Vinstpotential**

---

- |       |                                                                                                                                                                                                                                                                                                     |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CASAQ | <ul style="list-style-type: none"><li>* Kontroller och hopp- och förgreningsinstruktioner borgar för att datakvaliteten förbättras.</li><li>* Dataregistreringsprocessen effektiviseras.</li><li>* Uppgiftslämnaren kan själv välja en passande tidpunkt för att besvara frågeformuläret.</li></ul> |
|-------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
- 

- |                     |                                                                                                                                                                                               |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Röst &<br>Talteknik | <ul style="list-style-type: none"><li>* Minskar betydelsen av maskinskrivningsvana.</li><li>* Förenklar dataregistreringen vid intervjuer och observationer med handhållna datorer.</li></ul> |
|---------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
-

---

**Målpopulation  
och Informa-  
tionsbehov**

---

|                  |                                                                                                                                                                                                                                                         |
|------------------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| OCR              | <ul style="list-style-type: none"><li>* Det är viktigt att uppgiftslämnaren får ordentlig information om hur blanketterna ska fyllas i.</li><li>* Det är bäst om endast få tecken krävs för att besvara frågorna.</li></ul>                             |
| OMR              | <ul style="list-style-type: none"><li>* Det är viktigt att uppgiftslämnaren får ordentlig information om hur blanketterna ska fyllas i.</li><li>* Tekniken används till att registrera svaren vid flervälsfrågor.</li></ul>                             |
| CATI             | <ul style="list-style-type: none"><li>* Mer komplexa frågeformulär kan tillåtas.</li><li>* Intervjuarens inställning till datorer är mycket viktig.</li></ul>                                                                                           |
| CAPI             | <ul style="list-style-type: none"><li>* Mer komplexa frågeformulär kan tillåtas.</li><li>* Uppgiftslämnarens inställning till datorer är mycket viktig.</li></ul>                                                                                       |
| DATI             | <ul style="list-style-type: none"><li>* Mer komplexa frågeformulär kan tillåtas.</li><li>* Uppgiftslämnarens och intervjuarens inställning till datorer är mycket viktig.</li></ul>                                                                     |
| TDE              | <ul style="list-style-type: none"><li>* Uppgiftslämnaren måste ha tillgång till en tonvalstelefon.</li><li>* Denna teknik kräver korta och enkla svar.</li><li>* Tekniken är bra om samma uppgiftslämnare används med återkommande intervall.</li></ul> |
| CASAQ            | <ul style="list-style-type: none"><li>* Uppgiftslämnaren måste ha tillgång till en dator.</li><li>* Uppgiftslämnarens inställning till datorer är mycket viktig.</li></ul>                                                                              |
| Röst & Talteknik | <ul style="list-style-type: none"><li>* Uppgiftslämnarens inställning till datorer är avgörande.</li></ul>                                                                                                                                              |

---

## **Några Synpunkter**

I denna rapport har vi behandlat datorstöd vid datainsamlingen, med dess för- och nackdelar, ur en datoranvändares och statistikers synvinkel. Datorstöden har ur många synpunkter visat sig vara kostnads- och arbetsbesparande.

Arbetsmiljön har dessutom förändrats, enligt många till det bättre, då många tråkiga och monotona arbetsmoment har reducerats. För de som förut hela dagarna registrerade och kodade in data framför en dator har kanske arbetsskador och allergier minskat. Genom CATI, CAPI och DATI har dock datoranvändningen spridit sig till fler användare som tex intervjuarkåren. Om tex TDE och tal-teknologi i framtiden blir mer utbrett kanske det på sikt minskar behovet av människor framför datorer. Gemensamt för alla typer av datorstöd är att behovet av mer tekniskt kunnig personal ökar samtidigt som behovet av mindre kvalificerad personal minskar.

Undersökningarna och testerna av datorstöd har främst varit teknikorienterade. Utvecklingen har inte nått lika långt med avseende på kvalitetsfrågor. Risken finns att dessa frågor försummas om datorstödet medför stora kostnadsbesparingar eller att tekniken utvecklas alltmer för teknikens egen skull.

Problemen med datorstöd vid datainsamlingen ligger kanske i högre grad på en annan nivå som rör hela frågekomplexet kring datorns roll i samhället. Dessa frågor rör tex:

- \* Sårbarheten genom beroendet av datoranvändning.
- \* Datorns inverkan på den personliga integriteten.
- \* Energi- och råvaruförbrukning.
- \* Datorers betydelse för individens fysiska och psykiska hälsotillstånd.

Användningen av datorstöd vid datainsamlingen har kanske, i praktiken, endast en marginell inverkan på just dessa punkter eller kan kanske tom innebära vissa förbättringar. Dessa problemställningar ligger utanför vår rapport men de är dock viktiga att ha i åtanke.

I framtiden tror vi att datorstödet vid datainsamlingen kommer att öka. Det är dock viktigt att hela tiden framhålla kvalitén så att inte tekniken tar över.



## KÄLLFÖRTECKNING

- \* Augustsson, L., Irebäck, H., Jönsson, B.-Å.; (1990); DAVID: Data-insamling Via Diskett; Microbladet; Nr 1; sid 10-11
- \* Bergsten, O.; (1990); Manuell Dataregistrering - en Arbetsmiljömässig Omöjlighet eller blev Datoriseringens Förlorare - Vinnare?; Rapport; D/DB; Statistiska Centralbyrån; Örebro
- \* Biggar, R.S.; (1988); The Design and Development of a Computer Administered Questionnaire Utilising the dBASE III Language; American Statistical Association 1988 Proceedings of the Statistical Computing Section; sid 259-262
- \* Couper, M., Groves, R.M., Jacobs, C.A.; (1990); Building Predictive Models of CAPI Acceptance in a Field Interviewing Staff; Paper Presented at the Bureau of the Census' Fifth Annual Research Conference; mars 1990
- \* Keller, W.J., Metz, K.-J., Bethlehem, J.G.; (1989); The Impact of Microcomputers in Survey Processing at the Netherlands Central Bureau of Statistics; CBS report Netherlands Central Bureau of Statistics, Automation Department, Voorburg
- \* Kiesler, S., Sproull, L.S.; (1986); Response Effects in the Electronic Survey; The Public Opinion Quarterly/Organ of the American Association for Public Opinion Research; 50(3); sid 402-413; hösten 1986
- \* Nicholls II, W.L.; (1989); The Impact of High Technology on Data Collection; CATI Research Report No. GEN-1; Computer-Assisted Interviewing Central Planning Committee; Bureau of the Census; USA
- \* O'Muircheartaigh, C., Aye Maung, N., Bushnell, D., Martin, J., Murphy, M.; (1990); A Comparative Evaluation of Computer Assisted Interviewing Software; Project Team; Joint Centre for Survey Methods; London; England; 4 juli 1990
- \* Otterborg, K.; (1980); Maskinsystem för Optisk Läsning; ADB information; D/DC; Statistiska Centralbyrån
- \* Ponikowski C.H., Copeland, K.R., Meily, S.A.; (1989); Applications for Touchtone Recognition Technology in Establishment Surveys; US Bureau of Labor Statistics; Document 6:1; januari 1989
- \* SCB; (1987); Teknisk Specifikation för Serieversionen av Fältmaskinen i DATI-projektet; D/DC/PM; Statistiska Centralbyrån
- \* SCB; (1990); Rapport från DATI-projektets produktionsprov i Arbetskraftsundersökningen augusti 1989 - januari 1990; DATI-Projektet, Statistiska Centralbyrån; Örebro

- \* Shanks, M.J.; (1989); Information Technology and Survey Research: Where Do We Go From Here?; Journal of Official Statistics; 5(1); sid 3-21
- \* van Bastelaer, A., Kerssemakers, F., Sikkel, D.; (1988); Data Collection With Hand Held Computers: Contribution to Questionnaire Design; Journal of Official Statistics; 4(2); sid 141-154
- \* Waldrop, M.M.; (1988); A landmark in Speech Recognition; Science; Research News; 240 (4859); sid 1615
- \* Werking, G.S., Tupek, A., Clayton, R.; (1988); CATI and Touchtone Self-response Applications for Establishment Surveys; Journal of Official Statistics; 4(4); sid 349-362

Samtliga artiklar i denna källförteckning finns arkiverade i Biblioteket; U/STM; SCB; Örebro.



**R & D Reports** är en för U/ADB och U/STM gemensam publikationsserie som fr o m 1988-01-01 ersätter de tidigare "gula" och "gröna" serierna. I serien ingår även **Abstracts** (sammanfattning av metodrapporter från SCB).

**R & D Reports Statistics Sweden**, are published by the Department of Research & Development within Statistics Sweden. Reports dealing with statistical methods have green (grön) covers. Reports dealing with EDP methods have yellow (gul) covers. In addition, abstracts are published three times a year (light brown /beige/ covers).

Report published earlier during 1990:

- |                   |                                                                                                                                |
|-------------------|--------------------------------------------------------------------------------------------------------------------------------|
| 1990:1<br>(grön)  | Calibration Estimators and Generalized Raking Techniques in Survey Sampling (Jean-Claude Deville, Carl-Erik Särndal)           |
| 1990:2<br>(grön)  | Sampling, Nonresponse and Measurement Issues in the 1984/85 Swedish Time Budget Survey (Ingrid Lyberg)                         |
| 1990:3<br>(grön)  | Om justering för undertäckning vid undersökningar med urval i "rum och tid" (Bengt Rosén)                                      |
| 1990:4<br>(gul)   | Data Processing at the Central Statistical Office - Lessons from recent history (M. Jambwa)                                    |
| 1990:5<br>(beige) | Abstracts I - sammanfattning av metodrapporter från SCB                                                                        |
| 1990:6<br>(grön)  | Sequential Poisson Sampling from a Business Register and its Application to the Swedish Consumer Price Index (Esbjörn Ohlsson) |
| 1990:7<br>(grön)  | Kvalitetsrapporten 1990 - huvudrapport (NN)                                                                                    |
| 1990:8<br>(grön)  | Kvalitetsrapporten 1990 - bilagor (NN)                                                                                         |

Kvarvarande beige och gröna exemplar av ovanstående promemorior kan rekvireras från Inga-Lill Pettersson U/LEDN, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 49 56.

Kvarvarande gula exemplar kan rekvireras från Ingvar Andersson, U/ADB, SCB, 115 81 STOCKHOLM, eller per telefon 08-783 41 47.