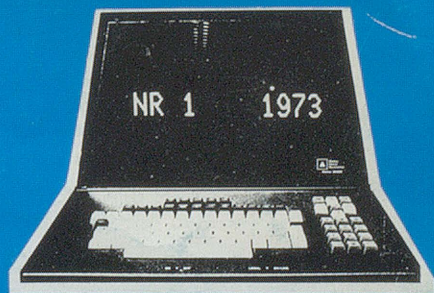


humanistiske data

Meldingsblad for
NAVF's EDB-senter
for humanistisk forskning

Norges Almenvitenskapelige Forskningsrad

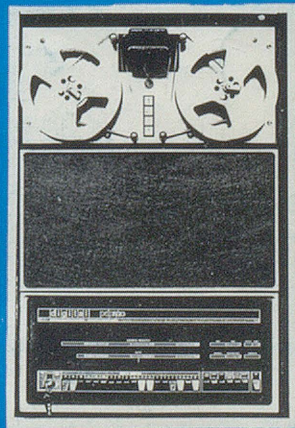
relus



INNHOOLD

1. Bakgrunn for opprettelsen av EDB-senteret
v/professor Eva Sivertsen
2. NAVF's EDB-senter for humanistisk forskning

Opprettelsen av senteret.
Senterets lokalisering.
Håsetting og arbeidsoppgaver.
Senterets styre og personale.
Senterets virksomhet til nå.



HUMANISTISKE DATA blir utgitt av
NAVF's EDB-senter for humanistisk
forskning i Bergen.

Senterets administrative leder,
Jostein H. Hauge, har det
redaksjonelle ansvar for meldings-
bladet.

De som ønsker å få bladet tilsendt,
kan bestille det ved henvendelse
til senterets adresse:
Villavei 10,
Boks 53,
5014 Bergen-Universitetet.

Innlegg kan sendes til samme
adresse.

Lay-out og trykk:

Universitetets trykkeri

Bergen 1973

NAVF's EDB-senter for humanistisk forskning vil med dette nummer starte utgivelsen av et meldingsblad.

Bladet har som sitt primære mål å informere interesserte enkeltforskere og fagmiljøer om den virksomhet som foregår i senteret. Det vil således bli orientert om de tilbud i form av konsulent-tjenester, kurs, databehandling, programmer og data som senteret kan tilby eller bistå med opplysninger om.

Et annet viktig mål er å presentere det arbeid som pågår innenfor humanistisk databehandling ved våre universiteter og virke for kontakt miljøene imellom.

I meldingsbladet tar en også sikte på å gi oversikter over utviklingen av programmer for humanistiske databehandlingsoppgaver i utlandet og informere om det forskningsmaterialet som er tilgjengelig for datamaskinell behandling.

Vi håper at bladet også vil bli benyttet som et organ for spørsmål og debatt på sitt felt og at de som er aktive innenfor EDB i de humanistiske fag vil bidra med rapporter om og redegjørelser for sine respektive forskningsområder.

Det er meningen at bladet skal utkomme minst 2 ganger pr. år, og vårt ønske er at vi med tiden kan tilby et kvartalsskrift.

Meldingsbladets første nummer vil bli viet NAVF's EDB-senter, dets målsetting, arbeidsoppgaver, tekniske utrustning og virksomhet til nå. Formannen i senterets styre orienterer om bakgrunnen for opprettelsen av senteret.

I neste nummer vil det bli gitt et oversyn over og en karakteristikk av igangværende EDB-tiltak innen de humanistiske fag i vårt land. Som resultat av sitt kartleggingsarbeid kan senteret gi en første oversikt over de samlinger av humanistiske data i maskinleselig form som finnes i fagmiljøene og tekniske spesifikasjoner vedr. deres lagring. Det vil bli særskilt opplyst om de restriksjoner som knytter seg til den eksterne bruk av dem og hvor en kan komme i kontakt med dem som forvalter datamaterialet.

Det vil også bli gitt rapporter om kurs, konferanser og studiereiser av interesse for et bredere publikum.

Neste nummer, som planlegges utgitt ved årsskiftet, vil også inneholde en beskrivelse av virksomheten i EDB-senteret høsten 1973.



*Villavei 10, Universitetet i Bergen,
hvor senteret holder til
i 1. etasje.*

BAKGRUNN FOR OPPRETTELSEN AV EDB-SENTERET

Ved professor Eva Sivertsen - formann i NAVF's EDB-komite for humanistisk forskning

NAVF har lenge følt behovet for spesielle tiltak for å bygge ut EDB-feltet innen humaniora. Av forskjellige grunner synes humanistene å ha hatt vanskelig for å ta i bruk dette nye hjelpemiddel, elektronisk databehandling. I 1968 ble derfor NAVF's EDB-komite for de humanistiske fag opprettet. Den var et rådgivende organ for Gruppe A, siden Fagråd A (Rådet for humanistisk forskning).

I sine første år så komiteen det som sin primære oppgave 'å spre kunnskap om EDB til humanistiske forskere og legge forholdene til rette slik at interesserte studenter og forskere kunne få mulighet til å spesialisere seg i EDB' (fra komiteens årsmelding for 1969). Det ble holdt kurs, opprettet spesielle stipendiatstillinger og tilsatt en konsulent.

Etter hvert ble det klart at dette ikke var nok. Det var latente, store behov for rådgivning, opplæring, hjelp til programutvikling, utstyr og koordinering av arbeidet. Men foreløpig var miljøet ved de enkelte læresteder ikke sterkt nok til at disse kunne påta seg ansvaret for utvikling av feltet. Det syntes naturlig for NAVF å påta seg oppgaven i en startperiode.

NAVF fant det mest hensiktsmessig til en viss grad å konsentrere innsatsen til ett sted. Dermed kunne en sikre en mest mulig rasjonell utnyttelse av ressursene, personell så vel som utstyr. NAVF ønsket derfor å opprette et nasjonalt senter for EDB i humaniora, for et begrenset antall år. Forholdene lå best til rette ved Universitetet i Bergen.

Der var det allerede utviklet en viss ekspertise i anvendelsen av EDB i humaniora, særlig innen språkvitenskap. Og der stod en foran anskaffelsen av et nytt stort regneanlegg som sentret ville kunne benytte seg av, i tillegg til sitt eget utstyr.

På denne bakgrunn besluttet NAVF å opprette sitt EDB-senter for humanistisk forskning, med sete i Bergen, men med ansvar for hele landet. Sentret kom i drift i løpet av 1972, og er etter hvert blitt godt utbygd. Det disponerer for tiden utstyr til en samlet verdi av kr. 900.000.-, i tillegg til at det har adgang til Universitetets regneanlegg, har en stab på 7 medarbeidere, hvorav 1 har sin arbeidsplass i Oslo. En nærmere redegjørelse for

organisasjon, målsetting og arbeidsoppgaver for sentret finnes på de følgende sider.

NAVF ønsker ikke å engasjere seg permanent i et slikt prosjekt. Rådet for humanistisk forskning har antydnet en foreløpig tidsramme på 5 år for virksomheten, idet en venter at de enkelte læresteder etter hvert selv vil kunne overta det fulle ansvar for feltet. Sentret bør derfor kunne arbeide for å overflødiggjøre seg selv. Dette betyr først og fremst styrking av de lokale miljøer.

For å oppnå dette, og for å kunne yte best mulig service generelt, trenger vi god kontakt med nåværende og potensielle brukere av sentret i alle deler av landet. Det er bl.a. derfor vi har følt behov for å utgi et meldingsblad.

NAVF'S EDB-SENTER FOR HUMANISTISK FORSKNING

1. Opprettelsen av senteret

Rådet for humanistisk forskning, Norges Almenvitenskapelige Forskningsråd vedtok i 1971 å opprette et nasjonalt EDB-senter for humanistisk forskning. Det ble besluttet at senteret skulle få til disposisjon maskinell utrustning i form av et mindre datamaskinanlegg til sitt arbeid.

NAVF har forpliktet seg til å opprettholde sitt engasjement i EDB-senteret i minst 5 år fra 1.1.1973. De første medarbeidere tiltrådte i løpet av mai/juni 1972.

2. Senterets lokalisering

EDB-senteret ble lagt til Bergen, og det var et ønske fra NAVF at det ble etablert et nært samarbeid med Universitetet i Bergen. For å regulere dette samarbeidet ble det inngått en avtale mellom de to institusjoner.

Ifølge denne avtale får EDB-senteret og dets samarbeidende

brukere rett til ytelser fra universitetets EDB-avdeling på like fot med universitetets egne institutter, herunder adgang til den programutrustning og maskinkraft som universitetet disponerer. Det ble også gitt tillatelse til å knytte senterets datamaskin til universitetets nye anlegg, UNIVAC 1110.

Universitetet i Bergen har stilt lokaler til disposisjon for senterets virksomhet. Senteret holder til i Villavei 10, i nær tilknytning til de sentrale universitetsbygninger.

3. Målsetting og arbeidsoppgaver

Rådet for humanistisk forskning vedtok 13.11.1972 endelige retningslinjer for virksomheten i EDB-senteret. Da den målsetting og de arbeidsoppgaver som her er satt frem, kan ha almen interesse, refereres de i sin helhet:

"Retningslinjer for NAVF's EDB-senter for humanistisk forskning:

Målsetting

NAVF's EDB-senter for humanistisk

forskning skal ha til hovedoppgave å stimulere og hjelpe forskere innen humaniora til å ta EDB i bruk i forskningsarbeidet. Senteret skal regionalt og sentralt søke å legge forholdene til rette for en fornuftig utnyttelse av EDB som hjelpemiddel i de humanistiske fag. Det langsiktige mål er å utvikle lokale miljøer som kan arbeide selvstendig ved de lokale institusjoner, slik at disse kan overta NAVF's engasjement på området.

Arbeidsoppgaver

Senteret driver informasjons- og kursvirksomhet innen de enkelte fagområder og på forskjellige nivåer om bruk av datamaskiner i humanistisk forskning.

Senteret gir konsulent-service i tilknytning til humanistiske forskningsprosjekter som gjør bruk av datamaskiner.

Senteret gir programmerings-, maskin- og puncheservice innenfor sin kapasitet.

Senteret utvikler programutrustning for humanistisk forskning på nasjonal basis.

I samarbeid med fagmiljøene skal sentret bidra til en systematisk oppbygging av dataarkiv ved å utarbeide retningslinjer og metoder for oppbygging og utnyttelse av dem. Sentret skal registrere og gi beskrivelse av dataarkiv og generelt medvirke til at flest mulig får adgang til slike arkiv.'

4. Senterets styre

Før senteret ble opprettet, hadde Rådet for humanistisk forskning i flere år hatt i arbeid en EDB-komite for de humanistiske fag. Komiteens oppgave er å yte bistand til NAVF ved behandling av EDB-saker og ved selvstendige initiativ å treffe tiltak som kan føre til utbredelse av EDB som et hjelpemiddel i de humanistiske fag.

Fra den tid senteret kom i arbeid, fikk komiteen i oppgave å være styre for det.

For tiden har styret denne sammensetning:

Formann: Professor Eva Sivertsen, (Engelsk institutt, Universitetet i Trondheim).

Medlemmer: Professor Ottar Dahl, (Historisk institutt, Universitetet i Oslo). Dosent Dag

Gundersen, (Norsk Leksikografisk institutt, Universitetet i Oslo). Førsteamanuensis Kolbjørn Heggstad, (Nordisk institutt, Universitetet i Bergen). Prosjektleder Arne B. Johansen, (Hardangerviddaprojektet, Historisk museum, Universitetet i Bergen). Professor Geir Kjetsaa (Slavisk-baltisk institutt, Universitetet i Oslo). Professor Svein Nordbotten, (Institutt for informasjonsvitenskap, Universitetet i Bergen). Vit.ass. Tore Simonsen, (Musikkvitenskapelig institutt, Universitetet i Trondheim).

Styret har nedsatt et arbeidsutvalg som består av professor Sivertsen, førsteamanuensis Heggstad og professor Nordbotten.

5. Senterets personale

Senteret har i 1973 følgende administrative og tekniske personale:

Administrativ leder for senteret (Jostein H. Hauge).

Sjefskonsulent som leder for det tekniske personale (George M. Gillow).

Konsulent med hovedoppgave innenfor informasjons- og kursvirksomhet, samt programutvikling (Geir Berge).

Lokal konsulent for brukere i Oslo med særlig oppgave å være kontaktledd mellom senteret og de EDB-orienterte fagmiljøer i Oslo (Ivar Fonnes, arb.sted: Hist. inst., Univ. i Oslo).

Programmerer for programarbeid på senterets og universitetets dataanlegg (Sigbjørn Hellene).

Operatør i 1/2-stilling til maskinelle driftsoppgaver (Per Vestbøstad).

Kontorfullmektig som også tar del i dataregistrering for senterets brukere (Else Janson).

Ved senteret arbeider dessuten i 1973 en programmerer, Aaron Gurski, med særlig oppgave å utvikle programmer til bruk ved de prosjekter innen klassisk filologi og arkeologi som senteret samarbeider med.

6. Den maskinelle utrustning

Som en del av senterets etablering, bevilget NAVF midler til innkjøp av et mindre datamaskinanlegg. Datamaskinen er av typen PDP 11/20 levert av Digital Equipment Corporation. Maskinen er installert i senterets lokaler. PDP 11/20 er utstyrt med et kjerne-minne på 24K 16-bits ord. Denne

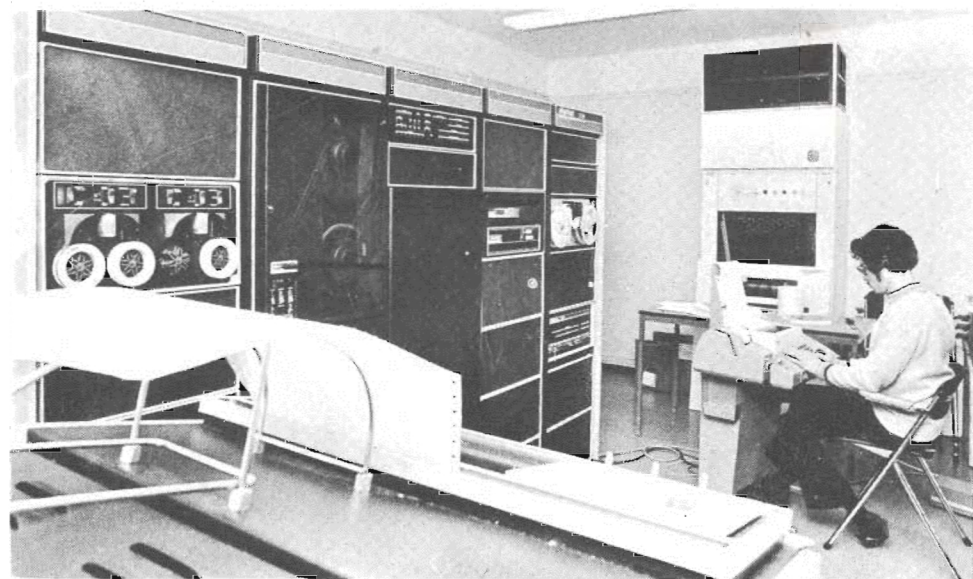
størrelsen på kjerneminnet gjør det f.eks. mulig å kjøre et time-sharing-system med programmerings-språket BASIC-PLUS.

Ved installering hadde maskinen en platestasjon med faste hoder som har en lagerkapasitet på 262K ord. Denne platestasjonen brukes til lagring av operasjonssystem og hyppig brukte data-filer. Stasjonen har en gjennomsnittlig aksess-tid på 20 ms. I løpet av høsten 1973 er maskinen blitt supplert med en platestasjon med utskiftbare plater, som hver tar 2,4 millioner karakterer. Dette utstyret har en midlere aksess-tid på 70 ms.

Karakteristisk for denne maskintypen er bruken av en egen type magnetbånd, såkalt 'DEC-tape'. PDP 11/20 har utstyr for bruk av to slike magnetbånd med samme kapasitet som platestasjonen med faste hoder.

Ettersom båndene er fil-strukturert, kan denne enheten betraktes som en langsom platestasjon. Magnetbåndene som brukes, er såvidt rimelige at de egner seg godt for lagring av brukerprogram, subrutine-sett osv. For lagring av større mengder data og for distribusjon av data har senteret en 9-kanals magnetbåndstasjon. Den leser 36.000 tegn/

*Senterets
datamaskin,
PDP 11/20.*



sekund. De største magnetbåndene kan lagre 12 millioner karakterer.

Anlegget omfatter også en papirbåndleser og en linjeskriver. Papirbåndleseren kan behandle 5, 7 og 8 kanals papirbånd og de 6 kanals papirbånd som mange trykkerier bruker (TTS-bånd). Leserhastigheten er ca. 300 tegn pr. sekund. Linjeskriveren har en linjebredde på 80 kolonner og kan skrive 250-800 linjer/minutt. Utstyret har fullt 96-karakterer ASCII-tegnsett. Det kan legges til at NAVF dessuten har delfinansiert en IBM 1403 linjeskriver på Universitetets EDB-avdeling. Denne linjeskriver, som har utskiftbar tegnkjede, kan i dag også levere

utskrift i gresk tegnsett. Senere utvidelser av tegnsett er mulig.

Senteret har dessuten utstyr for tilkobling av tre enheter for datakommunikasjon - en for synkron overføring og to for asynkron. Utstyret kan enten brukes lokalt eller over telefonnettet. Den synkron inngang brukes for tiden til kommunikasjon med UNIVAC 1110 via en 4-tråds fast oppkoblet telefonlinje.

Sommeren 1973 ble det innkjøpt en billedskjerm-terminal av typen Delta-Data. Terminalen vil i første rekke bli brukt i forbindelse med programkonstruksjon og redigering av datasett.

7. Senterets virksomhet til nå

a) PROGRAMUTVIKLING PÅ PDP 11/20

Senterets maskinutstyr ble formelt overtatt etter avsluttede godkjenningsprøver i september 1972.

I det året som er gått, har det vært drevet et kontinuerlig arbeid med programutvikling med grunnlag i maskinens operasjonssystem DOS. Programmering har hovedsaklig vært utført i ASSEMBLY og FORTRAN IV, men i den siste tiden er også BASIC-PLUS blitt implementert på anlegget (se nedenfor).

Arbeidet har i første omgang hatt som mål å utvikle et fleksibelt sett med standardprogram som er basale i humanistisk databehandlingsarbeid.

Av slike kan nevnes innlesingsprogram for papirbånd, som kan ta hånd om bånd med ulike koderepresentasjoner, korrektur- og redigeringsprogram, samt indeks- og sorteringsprogram. Det er f.eks. skrevet program som sorterer data i core og program som fletter sammen sorterte filer. Senteret vil om en tid dessuten få et eget programsystem for sortering fra leverandøren. Et viktig redskap

for humanistene er såkalte maskeprogram for gjenkjenning av ulike typer formelle strukturer i et tekstmateriale. De første programmer av denne type er utviklet. Innenfor rammen av et prøveprosjekt er det utviklet program for sortering av arkeologiske funn-data fra Bryggen i Bergen.

Det er også ønskelig for senteret å kunne behandle de mest vanlige formater av magnetbånd på maskinen. I første omgang er det utviklet program for lesing av IBM-bånd. Arbeidet med konvertering av de vanligste båndformat for UNIVAC 1110 pågår.

Et programbibliotek for matematiske operasjoner er tilrettelagt.

Oppgavene forbundet med å skape et tilfredsstillende telekommunikasjonsprogram for PDP har hatt høy prioritet. Dette arbeidet har to sider: 1. Program for kommunikasjon med UNIVAC 1110. 2. Program for kommunikasjon med skrive-maskinterminaler og billedskjermterminaler av ulike typer.

For kommunikasjon med UNIVAC 1110 er det for tiden i drift et program som simulerer terminaltypen UNIVAC 1004. Systemet har vært i drift fra mars. Men da dette systemet har flere ulemper (umuliggjør interaktiv bruk, parallell

programutvikling og har restriksjoner i karaktersett etc.) vil en gå over til et integrert kommunikasjonssystem basert på operasjonssystemet DOS-COMTEX. Dette systemet gjør det mulig å bruke UNIVAC 1110 og PDP 11/20 interaktivt fra flere terminaler og dessuten å starte dataoverføring mellom UNIVAC og plate, magnetbånd, linjeskriver på PDP.

Systemet er ferdig planlagt, og deler av det er programmert og testet.

I august i år ble det holdt en Nordisk sommerskole i datamaskinell språkbehandling i Bergen. Arrangør var Nordisk institutt, Universitetet i Bergen ved førsteamanuensis Kolbjørn Heggstad.

Under sommerskolen stilte senteret utstyr og medarbeidere til disposisjon for arrangementet, og var ansvarlig for det maskintekniske opplegg. Det ble i denne tiden utarbeidet en serie applikasjonsprogram i BASIC-PLUS, bl.a. program for morfologisk analyse, program for testing av generative grammatiske regler og program innenfor studiet av semantiske relasjoner. Som ledd i den praktiske delen av denne sommerskolen brukte deltakerne fire terminaler til PDP for øvelser i programmering og to terminaler til UNIVAC 1110.

b) PROGRAMUTVIKLING PÅ UNIVAC 1110

Til behandling av større datamengder vil senteret benytte Universitetets anlegg UNIVAC 1110. Det samme gjelder arbeid hvor samarbeidet mellom senteret og brukere ved Universitetet i Bergen etter en tilretteleggingsfase skal føres videre av instituttene selv. For en del oppgavers vedkommende vil en søke en arbeidsdeling mellom de to maskinanlegg alt etter data-behandlingens art og volum.

For tiden har senteret under utførelse konkordansprogram skrevet i COBOL - som i første omgang skal brukes for behandling av gammelengelske tekster. Videre er det såkalte Brown-Corpus, som inneholder en tekstsamling på 1 million ord bestående av amerikansk prosa av ulike genrer, tilrettelagt for bruk på universitetets anlegg.

Til analyse av dette materialet er det utviklet programmer for gjenfinning av språklige strukturer som fyller brukerspesifiserte krav.

Som ledd i sitt samarbeid med et forskningsprosjekt innen klassisk filologi er det utarbeidet program for innlesing og utlistering av greskspråklig materiale samt for senere korrigerende av data.

Foto: Morgenavisen



Programmerer Hellene og sjefskonsulent Gillow (til høyre, stående) instruerer kursdeltakere under øving i språklig databehandling.

Etter en fase med et prøveopplegg for databehandling av arkeologisk materiale fra Bryggen i Bergen på senterets maskin, har man nå tatt fatt på arbeidet med et permanent opplegg for behandling av materialet på Universitetets anlegg.

Mange av de brukere som senteret samarbeider med, vil ha behov for et databasesystem med mulighet for direkte adgang til databasen. Aktuelle systemer evalueres for tiden i samarbeid med andre brukergrupper.

En viktig delaktivitet har det også vært å tilrettelegge sorteringsprogram for de datatyper som er vanlig i humanistisk forskningsarbeid.

c) PROGRAMUTVIKLING I OSLO

Fra 1. mars 1973 har senteret hatt i arbeid en heltidsansatt konsulent i Oslo. Konsulenten har som en av sine hovedoppgaver å arbeide for utvikling av programutrustning både ved selvstendig programmering og ved å tilrettelegge standardprogram for humanister på det maskinelle utstyr Universitetet i Oslo disponerer. Den konsulent som nå innehar stillingen, cand. philol. Ivar Fonnes, har i flere

år arbeidet med databehandling innen sitt fagfelt, historie, og har utviklet en samling programmer kalt TEXT for behandling av tekstmateriale for språklige og innholdsmessige analyser

I løpet av den tid konsulentordningen i Oslo har vært i funksjon, har konsulenten gått videre i programutviklingen innenfor rammen av TEXT-systemet. Det pågår likeledes arbeid med å knytte en forbindelse fra TEXT-programmene til den statistiske programpakken DDPP med tanke på en fleksibel statistisk behandling av språkdata.

Konsulenten har foruten samarbeid med flere historiske og språklige forskningsprosjekter særlig tatt opp arbeidet med overføring av de mest brukte programsystemer til det nye datamaskinanlegget CYBER på Kjeller, som humanistene ved Universitetet i Oslo vil ta i bruk i økende omfang.

d) KONTAKT MED FAGMILJØENE

Senteret har i løpet av det første året, som på mange vis er en etableringsperiode, søkt å skaffe seg et relativt grundig oversyn over det EDB-arbeid som i dag utføres i vårt land innenfor den humanistiske fagsektor. Det har vært en oppgave for senteret å få klarhet i det behov for bistand som humanistene har i sitt arbeid og hvil-

ke felt senteret ut fra sine forutsetninger skal yte en innsats på. I løpet av vinteren 1973 ble det holdt kontaktmøter med humanistene i Oslo og Trondheim hvor senteret ble presentert og deltok i samtaler om EDB-virksomheten i de ulike fag. De syn på og forventninger til vårt EDB-senter som en del fagmiljøer fremsatte, gav senteret en nyttig bakgrunn for det videre planleggingsarbeid.

Det samme kan sies å være resultatet av de drøftinger som kom i stand i forbindelse med presentasjonen av senteret ved NAVF's nasjonale seminar for nordisk språkvitenskap høsten 1972.

Den største delen av senterets kontaktvirksomhet har ellers vært drevet uformelt ved kontakt med enkeltpersoner og prosjekter, ikke minst i Bergen.

Den kontakten som er opprettet her, har også resultert i et nærmere samarbeid med flere forskere og prosjekter hvor senteret har sagt seg villig til å utføre et prøveprosjekt i samarbeid med brukeren. Med et par prosjekter har senteret også innledet et samarbeid hvor prosjektenes programmerer arbeider i senteret. Ordningen gjelder foreløpig for 1973.

I den tid senteret har vært i arbeid, har det vært holdt en rekke orienteringsmøter og omvisninger i senteret, likeledes seminar i spesialemner innenfor språklig databehandling og i bruk av EDB innen musikkvitenskap.

Som tidligere nevnt ytet senteret i august i år teknisk bistand i form av utstyr og databehandling for Nordisk sommerskole i datamaskinell språkbehandling. Det er meningen å redegjøre nærmere for en del av disse aktivitetene i et senere nummer.

Senteret har for sitt videre arbeid behov for å følge med i den tekniske utvikling og vurdere det program- og maskinutstyr som utvikles ut fra den nytte produktene kan ha i humanistisk sammenheng. Det kan nevnes at senteret bl.a. har interessert seg for nye teknikker for dataregistrering og for bruk av datamaskiner i trykkeriarbeidet. Disse emner tar en sikte på å belyse nærmere i et senere nummer.