Kapittel 2 Målgrupper og produktplan

2.1 Målgrupper

Ved utarbeidinga av denne rapporten har dei fylgjande brukargruppene vore i fokus:

- offentleg administrasjon (Sametinget, dei samiske kommunane, m.fl.)
- samisk næringsliv
- samisk presse og samiske forlag
- privatpersonar
- skular og utdanning

I tillegg har det vore viktig å undersøkja alternativ for dei som ikkje har fått skriveopplæring i samisk, og undersøkja om det ved hjelp av språkteknologi kunne vera mogleg å hjelpa dei i gang med å skriva.

2.2 Pr foformar og program

Målet med rapporten har vore å gjera greie for kostnadene ved å utvikrj E0 0 -fått

2.3 Produktplan

I følgje spesifikasjonen for forprosjektet skulle utgreiinga dekkja retteprogram

Kapittel 3 Oversikt over eksisterande ressursar

3.1 Kva slags ressursar trengst?

For å kunna laga dei verktya som er skissert (stavekontroll, orddeling, synonymordbok og grammatikkontroll), skulle ein ideelt sett hatt det følgjande:

- allment:
  - eit stort og omfattande namnemateriale som inneheld personnamn og geografiske namn som er relevante: samiske person- og stadnamn i Finland, Sverige og Noreg, vanlege finske, svenske og norske stadnamn og personnamn, og eit godt utval utanlandske namn, firmanamn, merkenamn og namn på programvare.
- stavekontroll:
  - grammatikk med presis skildring av alle bøyingsklasser, komplette

- Nielsen: skildrar alle sider ved grammatikken, men mykje er arbeidsamt Œ.0 1 Tf ( ) ˉ

norske omsetjingar og eldre (no frårådde) ord og termar. Språkavdelinga har sagt ja til å la materialet brukast i prosjektet. Materialet er lagra i ein Lotus Notes-database, teiknsettet er 8-bits (Levi2).

- Sameradioen har ein database med ulike ord og frasar, først og fremst ord og uttrykk som ikkje finst i eksisterande, trykte ordbøker (eller med ny informasjon

filformat som ikkje lett kan brukast som det er (QuarkXPress). Det er mogleg at endringar i produksjonen kan gjera dette lettare når prosjektet startar.

  - Davvi Girji har òg ein del materiale som er ferdig delt, men det same problemet med filformat gjeld her òg. I tillegg har Davvi Girji i utgangspunktet takka nei til å gje frå seg slike tekstar.

- feilkorpus (stavekontroll og grammatikkontroll):
  - Både språkavdelinga og opplæringsavdelinga har eit ganske stort materiale med originaltekstar før språkvask, og i mange tilfelle same tekst etter språkvask. Dette materialet vil bli gjort tilgjengeleg for prosjektet.
  - Nils Øivind Helander har føreslått eit samarbeid mellom dette prosjektet og fleire skular for å få inn tekstar både før og etter retting. Det vil i så fall gje eit stort materiale som ein kan byggja på i arbeidet med retteprogrammet. Eit eventuelt samarbeid må utformast etter at arbeidet med retteprogramma har starta.

I språkteknologiprosjektet har ein nordsamisk språkteknologisk komponent vorte utvikla. Han er fullstendig dekkjande når det gjeld bøyingsmorfologien, men er uferdig for både

Kapittel 4 Leverandøroversikt med tilbod

Med leverandør er meint firma som kan levera dei ferdige produkta slik dei er

4.1.3 Polderland

Kontaktinformasjon: Polderland Language and Speech Technology, Nederland
www.polderland.nl

Sterke sider: eit relativt godt etablert firma i språkteknologibransjen, og ein

5.2 Lulesamisk

		Pakke 1	Pakke 2	Pakke 3

Kapittel 6 Organisering

6.1 Hovuddeltakarar i prosjektet

Dei viktigaste partane i prosjektet vil bli:

- Språkkonsulentane og dei samiske språkmiljøa
- Universitetet i Tromsø
- Eitt eller fleire firma
- Prosjektkoordinator

Dei ulike deltakarane vil arbeida i hop på hovudkomponentane, og heile

andre viktige programma å integrera med er InDesign for forlag og aviser, og OmniPage for OCR-behandling. Alle firmaa har uttrykt interesse for å integrera mot dei andre programma i prosjektet, men det vil bli dyrare enn å basera seg på løysingar med open kjeldekode. Og sidan ikkje eit einaste av firmaa som har levert eit forslag til budsjett har noka erfaring med å integrera med program med open kjeldekode, er det ein større risk enn å basera seg på eksisterproe og fungerproe løysingar med open

skal jamleg ha møter med ei styringsgruppe som sametinget har peikt ut. Prosjektleiaragktør absolutt ha erfaring av språkteknologiprosjekt, og helst retteprogramprosjekt frå før.

6.2.5 Språkkonsulent

Språkkonsulentagkskal først og fremst vera eit språkleg orakel for prosjektgruppa, særleg når det gjeld den normative sida av skriftspråket. Språkkonsulentkskal her

Kapittel 7 Prosjektplan

7.1 Tekniske føresetnader:

8.1 Merknader til budsjettet

Den desidert største utgiftsposten er det eigne prosjektarbeidet, og innanfor den posten ligg canhdtyngda av arbeidet på grunnleggjande arbeid for å få det eksisterande materialet eigna til retteprogram.

Budsjettet ligg sannsynlegvis noko over det som ville ha vore tilfelle for eit firma med eksisterande infrastruktur og med etablerte rutinar for arbeid med retteprogram. Utgangspunktet for planlegginga har vore å byggja opp kompetanse, organisering og databasar som ligg nære eller i det samiske miljøet. Dette vil betala seg grunnllang sikt, men krev litt ekstra satsing i starten.