

Språkteknologi för sammanfattningar och ökad läsbarhet

Arne Jönsson, Christian Smith



Arbetsgång

- Målgruppssegmentering
- Persona- och scenarioutveckling
 - PUGH-analys
- Interaktions- och tjänstedesign
- Tjänsteutveckling
 - Sammanfattare
 - Omskrivning till lätt svenska
 - Synonymhanterare
- Utvärderingar



FriendlyReader

Mål:

- Öka den digitala delaktigheten genom att underlätta för personer med lässvårigheter att tillgodogöra sig textuellt baserad information på Internet
- Utveckla en tjänst baserat på konceptet design-för-alla



Målgruppsegmentering

Segment baserade på litteratur och samtal med:

- Syncentral
- Utredare inom funktionsnedsättningar
- Forskare inom syn och dyslexi
- Utbildare inom specialpedagogik med inriktning mot syn
- Koordinatorer för studenter med funktionsnedsättningar
- Specialpedagog med inriktning mot dyslexi
- Personer med dyslexi och synnedsättningar



Resultat från målgruppsegmenteringen

Målgruppsegmentering

- 7 segment
- Valde ut 3:
 - 3) Skärmförstoring,
 - 5) Saknar flyt i läsningen,
 - 6) Undviker läsning



Persona

Tor (Sekundärpersona)



Figur 4.
http://www.freedigitalphotos.net/images/Younger_Men_g118-Smiling_Young_Man_p45811.html

Tor är 34 år gammal och bor centralt i Norrköping. Han jobbar som tapetserare och hans huvuduppgifter är att klä om och stoppa möbler. Det fungerar trots sin synskada ganska bra, så länge han inte ska göra något väldigt smått och pilligt. För några år sedan fick han en sjukdom som satte sig på synnerven som gjorde att hans syn försämrades kraftigt. Förr hade han ett aktivt liv med båt, skidor, fotboll och mycket annat. Nu när han inte kan se så bra har han fått dra ner på mycket men kan fortfarande göra lite av det som till exempel åka skidor eller åka båt om han inte är ensam i båten. Innan Tor blev synsvag läste han inte så mycket, men i och med hans synskada började han lyssna på ljudböcker, vilket han tycker är otroligt givande.

Internet använder Tor inte mycket varken på jobbet eller hemma, men när han gör det tar det mycket energi. Oftast använder han inga hjälpmedel för det. När det blir alldeles för jobbigt använder han dock sitt förstöringsprogram om det är något viktigt som han måste läsa. Han använder aldrig talsyntes eftersom han då känner sig sjuk och avstår därför hellre. Att läsa tar som sagt mycket energi och därför händer det ofta att han slutar läsa innan texten är slut, helt enkelt



Personor och scenarier

- Personor (människor med läs- och skrivsvårigheter, synsvaga, människor med annat modersmål än svenska)
- Scenarier (personor i interaktion med olika koncept av sammanfattningsprogrammet)
- Andra scenarier för att visa på olika konceptidéer
- Baserat på intervjuer
 - Två personor i segment 3), en vardera i 5) och 6)



Scenario

Tor som är synsvag

Tor har fått ett uppdrag av en kund att laga en gammal stol från 1700-talet. Han har aldrig gjort detta förr och behöver ta reda på hur han ska göra för att laga stolen utan att riskera att den går sönder. Han börjar söka på internet och finner en intressant sida. Han för sin förstörade mus till en länk som verkar spännande, där han upptäcker en lång text om hur han ska gå tillväga för att laga 1700-talsstolen. Innan han börjar läsa gör han bakgrunden gul för att inte bli bländad av skärmljuset. Precis när han ska börja läsa kommer han ihåg sin nya funktion, EasyReader och blir lättad över att slippa läsa hela dokumentet. Han förstörar och sammanfattar texten en del och läser den, vilket går mycket bra eftersom att han orkar läsa hela texten nu, till skillnad från förr. Tor behöver ändå inte läsa hela originaltexten eftersom han redan har ett hum om hur han ska reparera stolen.

Nu vet Tor hur han ska göra för att laga stolen från 1700-talen och gör det. Kunden blir väldigt nöjd med resultatet och Tor känner sig glad och tycker inte att det är något större problem att vara synsvag.



Pugh-analys

- Hjälperverktyg för att rangordna olika koncept
- Skissade koncept
- Krav identifieras
- Hur bra uppfyller varje koncept de olika kraven (-1, 0, 1)
- Personor och andra fiktiva användare viktas
- Varje krav viktas för varje persona
 - -1 = kommer aldrig att behövas
 - 1 = behövs inte
 - 2 = kan vara bra
 - 3 = måste ha
- "Bästa" alternativet



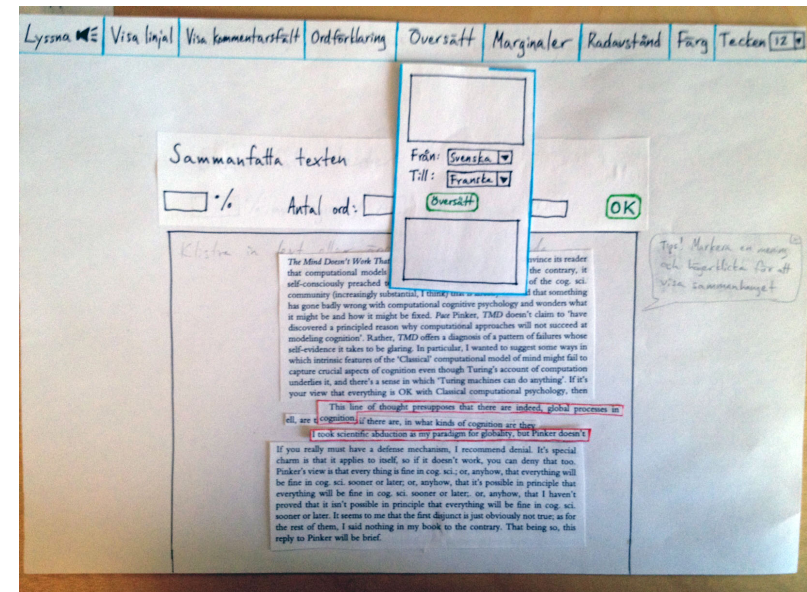
Resultat

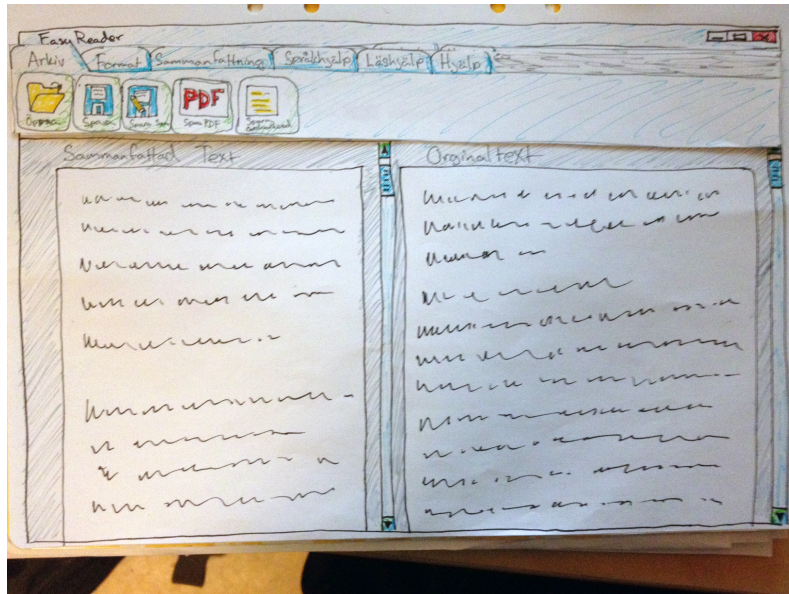
Koncept	Poäng alla	Poäng personor
Läsplatta	2.34	2.37
Manuell bok	2.24	2.14
Plug in desktop	2.05	2.006
Stand alone	2	2
Program på hemsida	1.93	1.93
Smartphone	1.89	1.94
Plug in web	1.77	1.71



Interaktionsdesign

- Valde ut 3: Plug in desktop, Stand alone, Program på hemsida
 - Implementerade Stand alone och program på hemsida
- Designkoncept
 - Skissa designkoncept baserat på personor och scenarier
 - Utvärdera med personer ur målgruppen





Tjänsten FriendlyReader består av

- Verktyg för dynamiska textsammanfattningar
- Stöd för omskrivning till lätt svenska
- Synonymhanterare
- Tilläggstjänst för ljudpresentation
- Hjälpmedel för att läsa texter
- <http://www.ida.liu.se/projects/friendlyreader/>

Textsammanfattningar

- Ett verktyg för att skapa sammanfattningar
 - Problem att läsa
 - Olika syften
 - Ersätta läsning
 - Ge indikation på om det är intressant att läsa
 - Studera sammanfattningar
 - Kan variera ett antal parametrar för att förbättra sammanfattningarna
- Bygger på vektorrumsmodellen Random indexing och Page Rank

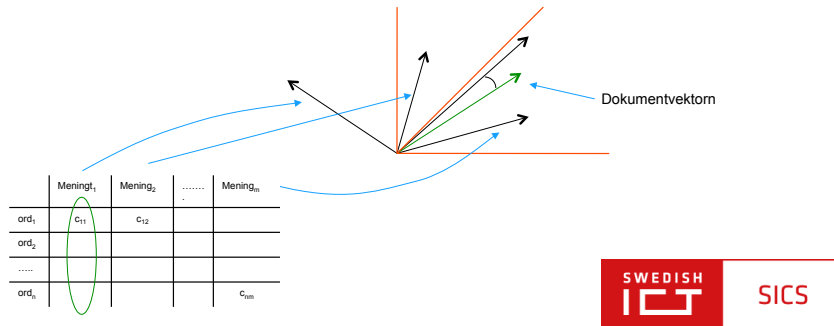
Vektorrumsmodeller

- Representera ords och meningars betydelser som koordinater och riktningar i en stor rymd
- Tekniker som reducerar den lingvistiska variationen och fångar semantiskt relaterade begrepp
- Skapa en matris med textenheter/dokument som kolumner och alla ord i textenheterna som rader
- Cellerna i matrisen är skilda från noll om ordet på en rad finns i textenheten. Gles matris, dvs många nollor.

	Mening ₁	Mening ₂		Mening _{t_m}
ord ₁	c ₁₁	c ₁₂		
ord ₂				
....				
ord _n				c _{nm}

Vektorrumsmodell

- Varje kolumn (eller rad) kan ses som en vektor i en mångdimensionell rymd
- Likhet mellan dokument mäts som närhet i vektorrummet (cos)
 - Sammanfattningar skapas genom att ta de meningar vars vektorer är mest representativa för dokumentet



Random indexing

- Tilldela varje ord en slumpmässig vektor
- Gå igenom hela texten ord för ord
- Uppdatera varje ordvektor med ordvektorena för grannarna
 - $w_{n-2} w_{n-1} w_{n+1} w_{n+2}$
 - $w = w_{n-1} + w_{n+1} + 0,5(w_{n-2} + w_{n+2})$
- Skapa meningsvektorer genom att addera meningens ordvektorer
- Skapa dokumentvektor genom att addera alla meningsvektorena

Random indexing, exempel

Exempel *Sven gick i skogen. Han bet hunden. Den skrek*

1. Tilldela en gles slumpvektor (ofta 300 dimensioner) till varje ord

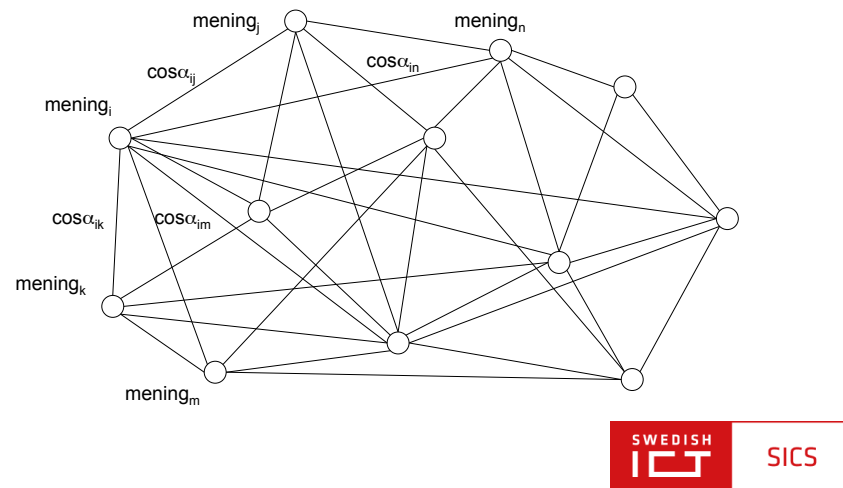
...
Han (0 -1 0 1 0 -1 0 0 0 1)
bet (0 0 0 0 0 1 1 -1 -1 0)
hunden (-1 0 1 0 0 1 0 -1 0 0)

2. Kontextvektor för *bet* (tar bara hänsyn till närmsta grannar)
 (-1 -1 1 1 0 0 0 -1 0 1)
3. En meningsvektor skapas genom att slå ihop alla kontexter i en mening.

Ytterligare förbättring

- Använd måttet på hur nära meningsvektorer är varandra för att ytterligare stärka viktiga meningar
- Skapa en graf över alla meningar med avståndet mellan meningar som vikter mellan meningarna
- Uppdatera vikterna mellan meningar utifrån grafen med hjälp av Page Rank
- Välj de meningar med högst rank (uppdaterad vikt)

Page Rank, illustration



Träning av sammanfattaren

- Utöka med stor förtränad rymd
- Mindre slump, bättre kvalitet (jmf. guldstandard)
- Mer robust



Resultat

- Lika bra som mänskliga sammanfattare
 - Skapat guldstandarder
 - Normalt mellan 60-70% överlapp mellan människor



Automatisk omskrivning till lätt svenska

- Syntaktiska regler
 - Mönster som är vanliga i lätta texter men inte i svåra
 - Transformation svår → lätt
 - REPL//NP-DT NN → NP-NN §P(#)
 - Byt ut nominalfras innehållande artikel och substantiv till endast substantiv, dvs ta bort artikeln. "Jag tycker om den tröjan" → "Jag tycker om tröjan."
- Text → ordklasstagning → parsning → omskrivning

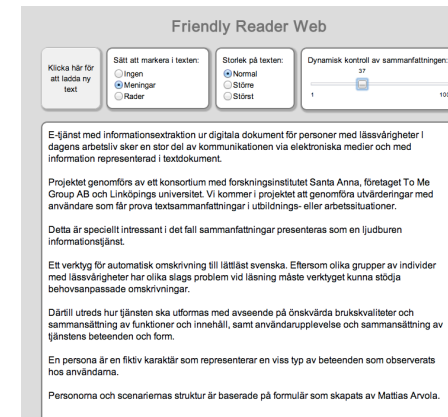


Synonymer

- Studier av synonymer
 - Längd
 - Frekvens
 - Grad av synonymitet
- Resultat
 - Ökad läsbarhet utifrån läsbarhetsmått
 - Många fel
- Ersätter inte ord med synonymer utan presenterar en lista med tänkbara synonymer



Webbtjänst



PDFstöd

3. Eye Tracking

Eye tracking is a method with many possible applications. The main concept associated with the method is that the eyes provide a kind of direct link to the cognitive processes and by analyzing the movement of the eye it is possible to gain insight into the cognitive state of a person executing a certain task. The eye's movement is a result of both goal driven and stimulus driven processes (Dachowski, 2007), and depends strongly on the type of cognitive task that is being performed. In our studies we will measure:

- **Fixations**, the period of time where the eye is relatively still (about 200-300 ms).
- **Fixation duration**. Just and Carpenter (1980) formed a hypothesis that an object or a text is processed exactly as long as a fixation lasts, and therefore implies a relatively easy access to cognitive processing. However, this is not uncontroversial, and the hypothesis has been questioned (Holmqvist, 2011; Rayner, 1998; Reichle et al., 1998).

The fixation duration indicates the effort needed for the cognitive processing, but the average fixation duration varies depending on the task and stimuli. **The more complicated a text is, the longer the average fixation durations, and factors like stress might result in shorter fixations** (Holmqvist, 2011).

According to Rayner (1998), the average fixation duration is not an adequate measure since it underestimates the duration that the fixations last. **The first fixation is often longer than the following fixations on the same word, and the mean duration to reread in many cases slightly too low.** Rayner (1998) claims that the first fixation duration is a better measure of cognitive processing.

In usability research, many short fixations imply that information that was expected to be found is missing (Shinke and Wilson, 2007).

All words of a text are not fixated during reading. **Long words are more likely to be fixated than short ones** (Just and Carpenter, 1980), but other aspects such as frequency and predictability from context are also given to be a reason for shorter fixations or word skipping (Reichle et al., 1998).- **Pupil size, which increases during problem solving and correlates to the difficulty of the task which implies that this could be used as a measure of cognitive activity** (Heus and Poth, 1964).

The diameter of the pupil can indeed be used to measure cognitive workload, though one has to be aware of the problems this method involves. **The pupil size is sensitive to various states of the participant and the environment, factors that should be accounted for in the experimental design.**

Except for cognitive workload, pupil size increases as an effect of emotion, anticipation, pain or drug influence, and it might decrease due to factors like fatigue, diabetes or high age. **The environment factors can**

be controlled for by ensuring that the presented stimuli are of the same brightness and content and that the lighting of the room is kept constant (Holmqvist, 2011).

4. Procedure

The study was conducted on 20 students, 13 men and 10 women. They were all native Swedish speakers without any writing or reading disability and with normal or corrected-to-normal vision. The average age was 23.2 (SD = 2.76). The experiment consisted of four parts: answering a questionnaire, text reading, error marking and text rating.

4.1. Equipment

The eye tracking equipment used for this study was SMI View RED II 50 Hz Pupil/Corneal reflex camera mounted underneath a 19" computer monitor. The software used for recording and analyzing the eye tracking data were View X, Experiment Center 3.0 and BeGaze 2.

4.2. Texts

The texts used in the tests were four texts from the Swedish popular science magazine *Forskning och Framtid*. The summaries are in Swedish and produced by the automatic text summarizer CoCoSu (Smith and Jönsson, 2011). **CoCoSu is based on Random Indexing and a modified version of the Weighted Page Rank algorithm, which is used for selecting the sentences that are most relevant in the text** (Smith and Jönsson, 2011). The algorithm calculates a rank based on the Random Indexing vectors, which makes sentences that are similar in content support each other, and eventually result in a ranking of the sentences by their importance. The output of the summarizer was not in any way formatted, other than being divided into paragraphs in order to enhance readability. The texts were previously tagged for errors by Karperson et al. (2012). **The texts were summarized to a summary level of 33% meaning that 33% of the original text remained chosen in order to get as many errors as possible in a text, while keeping it at a reasonable length that is still readable** (Karperson et al., 2012).

The four texts varied in length from 11 to 14 sentences and the number of tagged errors varied from 6 to 12 per text. In total there were 34 errors. The error types and number of errors for each type that were present in the texts were:

- 1(c) Erroneous anaphoric reference - Pronouns, a total of 4 errors
- 2. Absent cohesion or context, a total of 16 errors
- 3(a) Broken anaphoric reference - Noun-phrases, a total of 4 errors
- 3(c) Broken anaphoric reference - Pronouns, a total of 10 errors

The remaining error types were not present in the texts. Table 1 shows the amount of tagged cohesion errors for each text and the number of sentences for each text. The row labeled *Percentage* represents the ratio of the number

Användarstudier

- Sammanfattningar
- Omskrivningar till lätt svenska
- Kombinationer
 - Sammanfattning → Lätt svenska
 - Lätt svenska → Sammanfattning
- Bättre att först skriva om texten innan den sammanfattas



Läsbarhet hos sammanfattningar

- LIX, ordvariationsindex
- OVIX, andel unika ord
- NR, nominalkvot
- PN, andel bestämda substantiv
- Medelmeningslängd
- Medelordlängd
- Extra långa ord

$$LIX = \frac{n(ord)}{n(meningar)} + \left(\frac{n(ord > 6 \text{ bokstäver})}{n(ord)} \times 100 \right)$$
$$OVIX = \frac{\log(n(ord))}{\log(2) - \frac{\log(n(unika \text{ ord}))}{\log(n(ord))}}$$
$$NR = \frac{n(nomen) + n(prepositioner) + n(participer)}{n(pronomen) + n(adverb) + n(verb)}$$



Textkvalitetsstudier, Läsbarhet

- Sammanfattningar vs original
- Tre olika genrer
 - DN
 - Forskning och framsteg
 - Försäkringskassan
- Bättre läsbarhet enligt LIX hos sammanfattningar

$$LIX = \frac{n(ord)}{n(meningar)} + \left(\frac{n(ord > 6 \text{ bokstäver})}{n(ord)} \times 100 \right)$$



Textkvalitetsstudier, kohesion

Feltyper:

- Felaktig anaforsisk referens
- Avsaknad av kontext eller koherens
- Trasig/missledande anaforsisk referens
- Felaktiga citat
- Ofullständiga listor och upprädningsor
- Klippta meningar



Resultat

- Avsaknad av kontext och koherens samt felaktiga och trasiga/missledande referenser är de vanligaste feltyperna i automatiskt sammanfattade texter
- Feltyperna förhåller sig olika i relation till sammanfattningsgrad
- Vidare användarstudier behövs för att se hur feltyperna påverkar läsbarhet och förståelse för slutanvändaren
- Ögonrörelsemätningar



Studier av textkvalitet

- Högskoleprovet
- 60 personer, studenter
- Sammanfattningar och omskrivningsregler
- 30% sammanfattning, full text, gissning
- Antal rätta svar, tid att läsa samt subjektiv skattning utifrån frågeformulär
- Flest rätta svar med originaltexterna. Sammanfattningarna bättre än att gissa
- Sammanfattningarna ansågs betydligt bättre än de omskrivna texterna.
- Originaltexterna ansågs bättre än de bearbetade texterna
- Gick betydligt fortare att läsa de bearbetade texterna



Studier av användbarhet

- Personer med lässvårigheter.
- Tidig version av stand-alonetjänsten och fokuserat på användningen av tjänsten.
- I en studie använde 11 dyslektiker tjänsten.
 - Förfrågningar till 60 personer med lässvårigheter men endast 11 svarade.
- I stort sett nöjda med tjänsten, många uppskattade t.ex. möjligheten att få texten uppläst.
- Förslag till förbättringar som implementerats i den slutgiltiga versionen.
- Intervjuer med dyslektiker och observationer, 6 personer
- Även denna studie visade att tjänsten är uppskattad och ytterligare förslag till förbättringar presenterades, vilka realiserats i den slutliga tjänsten



Webblättnläst

- Syfte: Göra textuell information på internet lättare tillgänglig för personer som har problem att läsa och förstå svenska
- Rangordna texter efter hur lättlästa de är → prioritera att läsa de texter som är lättast att läsa utifrån olika kriterier
- Utveckla teknik för att automatiskt rangordna svenska webbsidor efter läsbarhet
- Enkelt kunna anpassas till olika läsbarhetsmått, användargrupper och texttyper



Arbetsätt

- Samma som EasyReader
 - Persona, scenarier, PUGH-analys
 - Designkoncept
 - Teknikutveckling
 - Användarstudier
- Olika viktning
 - Relevans, google
 - Läsbarhet, LIX, OVIX



Webblättläst

[Information](#)

lättläst

Sök

Sortera efter det här måttet: LIX Visa i siffror hur lättläst en sida är: ☐ Lix ☐ Ovix

☐ Spara inställningar

Genom att välja "spara inställningar" följer du av att du klickar på "Sök" ovan godkänner du att sidan sparar cookies på din dator



Mer information

- <http://www.ida.liu.se/projects/friendlyreader/>
- <http://www.ida.liu.se/projects/webblattlast/>

