

# Slutrapport för projektet Vidareutveckling av Folkets lexikon med fokus på tillgänglighet

Viggo Kann  
viggo@algoritmica.se

1 juli 2014

## 1 Inledning

Internet har gjort det möjligt att skapa stora verk (som lexikon) genom att massor av människor drar var sitt strå till stacken. Ett verk som kräver en enorm arbetsinsats kan därmed skapas utan att varje enskilt bidrag behöver vara så stort. Ju fler personer som lämnar bidrag desto mindre behöver varje bidrag bli. På engelska kallas detta för crowdsourcing. Det mest kända verket som byggts upp på detta sätt är encyklopedin Wikipedia.

År 2008 startade projektet Folkets engelsk-svenska lexikon och i mars 2009 gjordes detta Folkets lexikon tillgängligt för allmänheten på <http://folkets-lexikon.csc.kth.se>

Detta stora svensk-engelska och engelsk-svenska lexikon kan utvidgas av folket genom att användarna bedömer översättningsförslag som presenteras automatiskt när man gör uppslagningar i lexikonet.

I detta fortsättningsprojekt, som startade i juli 2013, har Folkets lexikon utvecklats på många sätt. Projektet blev klart juni 2014.

## 2 Mål och syfte

Projektets mål och syften kan sammanfattas i tre punkter.

1. Tillgänglighetsanpassa Folkets lexikon för att nå ut till nya grupper av användare

såsom synskadade och äldre.

2. Förbättra översättningsförslagen som användarna tar ställning till, generera nya frågor för disambiguering av flertydiga ord och tillhandahåll statistik som visar hur lexikonet växer för att säkra fortsatt intresse hos användarna, båda att slå upp i och att själva bidra till Folkets lexikon.
3. Byt ut de gamla serverna som tjänsten Folkets lexikon går på för att den stora mängden uppslagningar ska kunna hanteras med kort svarstid även under den tid på dagen då belastningen är som störst.

## 3 Projektbeskrivning

Projektet har drivits av Viggo Kann på Algoritmica Hollman & Kann HB och finansierats av .SE-stiftelsens Internetfond. Det bygger på två tidigare projekt som också finansierats av Internetfonden.

Över hundratusen Internetanvändare har bidragit genom att rösta på översättningsförslag eller själva skicka in förslag på översättningar och andra tillägg eller ändringar av Folkets lexikon.

Projektet bestod av fem delar som genomförts efter varandra, se beskrivningen av leverablerna nedan. I projektansökan gav vi två alternativ till sista leverabeln. Antingen skulle en

PDF-version av tas fram eller också skulle homonymer och homografer hanteras med hjälp av användarfrågor. I december 2013 valde vi det andra alternativet.

### 3.1 Ekonomi

Av projektmedlen från Internetfonden på 150 000 kronor har 40 000 kronor använts för inköp av två servrar och 110 000 kronor för projektarbete, som utförts av Viggo Kann. Dessutom har Algoritmica Hollman & Kann HB medfinansierat projektarbetet med 110 000 kronor. Detta betyder att projektbudgeten har följts.

## 4 Leverabler

### 1. Presentation av statistik för användarbidrag

Statistik för användarbidrag och användarbedömningar samt antalet nya och ändrade ord samlas in dagligen och presenteras för användaren första gången hen går till Folkets lexikon och med jämna mellanrum därefter om användaren gör flera uppslagningar eller bidrag. Detta gör att användarna uppmärksammas på att Folkets lexikon utvecklas och att användarbidrag påverkar lexikonet. Första leverabeln var klar i augusti 2013.

### 2. Införskaffande och installation av två nya servrar

Två nya servrar har köpts in och satts upp. Tjänsten har flyttats över till den ena av dessa medan den andra servern används som testmiljö och reservserver om den första skulle drabbas av något fel. Servrarna har tre års garanti. Webbadresserna utåt är oförändrade, så användarna märker inget annat än att svarstiderna minskar. Andra leverabeln var klar i januari 2014.

### 3. Anpassning för synskadade

Alla bilder och knappar på hela uppslagningssidan är nu fokuserbara (dvs kan nås via tangentbordet utan musklick) och försedda med alternativtext. Uppslagsorden är märkta med språk så att skärmläsare som kan läsa både svenska och engelska förstår. Färgval och kontrast har setts över. Tredje leverabeln var klar i februari 2014.

### 4. Förbättrad filtrering av förslag för bedömning

Över 25000 dåliga bidrag (översättningar och synonymer) har tagits bort ur bidragsdatabasen. När nya bidrag görs granskas nu dessa med avseende på språk innan dom visas för andra användare. Fjärde leverabeln var klar i april 2014.

### 5. Bättre hantering av homonymer och homografer

När Folkets lexikon skapades las synonymer från Folkets synonymlexikon och länkar till associationslexikonet Saldo in på alla betydelser av varje ord. Det betyder att en post i Folkets lexikon kan länka till upp till åtta poster i Saldo. Alla ord (omkring 20 stycken olika ordtyper) med sex eller fler Saldolänkar har gått igenom manuellt och länkar till felaktiga länkar och synonymer har tagits bort. För övriga flertydiga ord har disambiguerande användarfrågor konstruerats för både Saldolänkar och synonymer. Totalt har 4000 frågor om Saldolänkar och 12000 frågor om synonymer skapats. Femte leverabeln var klar i juni 2014.

## 5 Resultat

De fem leverablerna ovan har levererats enligt ursprungliga specifikationer och enligt plan, förutom att inköpet av servrarna senarelades

två månader eftersom vi behövde undersöka marknaden före köpet. Resultaten beskrivs ovan för varje leverabel.

Folkets lexikon finns tillgängligt för alla på <http://folkets-lexikon.csc.kth.se>

## 6 Utvärdering och analys

### 6.1 Utvärdering av resultat

Folkets lexikon är mycket välanvänt och tycks uppskattat av användarna. Användningsstatistik finns på webbsidan (under Om Folkets lexikon). Sedan 2012 har en miljon användarbedömningar gjorts och 170 000 användarförslag har kommit in. Över 150 000 användare har svarat på bedömningsfrågor och 1 000 är så hängivna att dom har gjort mer än hundra bedömningar. Efter användargranskningen har 20 000 nya ord lagts till och 42 000 existerande ord har ändrats.

Tidigare har användare hört av sig och klagat på att många översättningsförslag som användarna fått ta ställning till varit dåliga. Dessa klagomål har nu helt försunnit.

Samtliga uppställda mål för projektet har uppnåtts. Vi kan se att 170 000 nya användare (IP-adresser) har gjort uppslagningar i lexikonet under de fyra månader som gått sedan det tillgänglighetsanpassades, vilket tyder på att Folkets lexikon nått nya grupper.

### 6.2 Förslag på förbättringar

Flera av förbättringsförslagen som nämns i Patrik Liljas exjobbssrapport, kapitel 7, har genomförts i detta projekt. Några förslag i rapporten återstår dock men är svårare att åtgärda. Det gäller bland annat möjligheten för användare att ta bort hela ord och att redigera grammatikkommentarer.

Den alternativa sista leverabeln i detta projekt är också en tänkbar framtida förbättring,

nämligen att generera ett läsbart Folkets lexikon för nedladdning i PDF-format med traditionell lexikonlayout.

## 7 Framtida arbeten

Folkets lexikon är gjort så att det ska kunna fortleva med minimalt underhåll. Folkets egna förslag och bedömningar gör att innehållet utvecklas och förhoppningsvis blir både större och bättre. Med de nya serverna som köptes i detta projekt är tjänsten säkrad till och med utgången av år 2016.

Det finns flera webbsidor och appar som använder Folkets lexikon, antingen webbservicen eller som har laddat ner lexikonet som XML. Det visar att det finns behov av ett fritt tillgängligt svensk-engelskt lexikon och att driftiga användare arbetar vidare med materialet på olika sätt, precis som vi önskat.

## 8 Referenser

Lars Borin, Markus Forsberg & Lennart Lönngrén: SALDO 1.0, Språkbanken, Göteborgs universitet, 2008,

[http://spraakbanken.gu.se/personal/markus/publications/saldo\\_1.0.pdf](http://spraakbanken.gu.se/personal/markus/publications/saldo_1.0.pdf)

Viggo Kann & Magnus Rosell: Free construction of a Swedish dictionary of synonyms, NoDaLiDa 2005, Joensuu, 2005,

<http://www.nada.kth.se/theory/projects/infomat/rapporter/kannrosell105.pdf>

Patrik Lilja: Folket har makten över ordet – ett ordboksredigeringsystem för allmänheten, Examensarbete, KTH CSC, 2011, TRITA-CSC-E 2011:071,

[http://www.nada.kth.se/utbildning/grukth/exjobb/rapportlistor/2011/rapporter11/lilja\\_patrik\\_11071.pdf](http://www.nada.kth.se/utbildning/grukth/exjobb/rapportlistor/2011/rapporter11/lilja_patrik_11071.pdf)