

# Crowdsourcing – hvem, hvad og hvorfor?

Der ligger en stor udfordring i at digitalisere (den globale) kulturarv. En række store projekter benytter crowdsourcing, som er at outsource opgaven til befolkningen. Det skaber fremdrift, men hvad er erfaringerne? Og hvad får bibliotekerne ud af det?

Af Birte Christensen-Dalsgaard [bcd@kb.dk](mailto:bcd@kb.dk), Det Kongelige Bibliotek

Godt 330 mio. personers fødselsattester er blevet transkriberet af en frivillig arbejdsstyrke på 160.000 individer i FamilySearch-initiativet<sup>1</sup>, godt 6.000 frivillige har korrekturlæst 7 millioner linjer avisartikel i det australske avis digitaliseringsprojekt<sup>2</sup>, og 67.000 tags eller brugerkommentarer blev tilføjet Library of Congress' billeder på Flickr af godt 5.000 profiler i løbet af de første 10 måneder efter projektet blev iværksat<sup>3</sup>.

Dette er tre af mange eksempler på, hvad man kan opnå, hvis man inviterer befolkningen til at hjælpe med at gennemføre dele af de enorme udfordringer, som ligger i en mere eller mindre struktureret overføring af den analoge kulturarv til den digitale verden.

At folk generelt også er interesseret i at bidrage med materiale og historier, er der også flere eksempler på. Gennem Picture Australia<sup>4</sup> har Nationalbiblioteket fået godt 55.000 billeder, og BBC har gennem deres WW2 Peoples War<sup>5</sup> modtaget 47.000 historier og 15.000 billeder i løbet af de 2 ½ år, det var aktivt.

I Danmark har vi også en del initiativer, som inddrager folk bredere i forstand. Typisk er her tale om projekter, som bygger på deltagelse af foreninger eller skoler, men ikke kun det. Dansk Demografisk database<sup>6</sup> blev påbegyndt i 1992, og siden er der indtastet ca. 10 mio. navne bl.a. baseret på de komplette folketællinger fra 1787, 1801, 1834 og 1838.

Men hvad får folk til at investere timer, dage, uger, år på frivilligt arbejde – som at læse korrektur, som at tage billeder eller som at transskribere vanskeligt læselige dokumenter?

## Crowdsourcing

Brugen af crowdsourcing, som ovenstående er eksempler på, har de sidste par år været genstand for en del eksperimenter. Begrebet er sammensat af "crowd" og "outsourcing" og blev introduceret af Jeff Howe i 2006 i Wired Magazine i artiklen "*The Rise of Crowdsourcing*"<sup>7</sup>.

Den definition, jeg oftest har set, ligger på wikipedia:

*Crowdsourcing is the act of outsourcing tasks, traditionally performed by an employee or contractor, to a large group of people or community (a crowd), through an open call.*

Internettet har muliggjort en meget bred deltagelse, og analyser af deltagerprofiler viser, at der ofte er deltagere fra mange lande.

Ovenfor er allerede nævnt en del eksempler på crowdsourcing. Der er mange yderligere eksempler, og målet i denne artikel er ikke at give en komplet liste over projekter, men snarere at se på nogle af de mekanismer, som får folk til at deltage. Hvem tiltrækkes af denne type aktivitet? Og hvad er det for opgaver, som tiltrækker folk?

En artikel om et crowdsourcing-eksperiment havde titlen: *Passion as a sustainable source of hard work*. Og selv om passion ikke er hele svaret på hvorfor, tyder meget på, at det er en vigtig komponent.

## 3 ANBEFALEDE ARTIKLER

Crowdsourcing. How and Why Should Libraries Do. Af Holley Rose. 2010<sup>8</sup>

For the Common Good: The Library of Congress Flickr Pilot Project. Af David Woodward. 2008<sup>9</sup>

Capturing the power of the crowd and the challenge of community collections. 2010<sup>10</sup>

## Hvad ved vi om deltagerne?

Rose Holley<sup>8</sup> har analyseret 8 crowdsourcing-projekter, herunder de to australske, korrekturlæsning af digitale aviser og Picture Australia, og har givet et bud på profilen af deltagerne, motiverende faktorer samt belønningsmetoder. Der findes flere andre undersøgelser, herunder af Library of Congress' (LC i det følgende) Flickr-eksperimenter<sup>9</sup>, som alle understøtter punkterne nedenfor.

Om deltagerne siges bl.a.:

- 10 % laver 80 % af arbejdet.
- Halvdelen af de aktive har en personlig interesse i emnet, den anden halvdel gør det for at støtte en god sag.
- Syge, handicappede og rekonvalescenter deltog, fordi de kunne gennemføre arbejdet hjemme, og det gav mening og struktur på dagen.
- Man bliver afhængig.

Og om de motiverende faktorer nævnes bl.a.:

- Interessant, sjovt og man lærer.
- Værdigt formål.
- Målet er virkelig udfordrende.

Som belønning fremhæves:

- Navne på deltagerne nævnes på websiden.
- Man rankeres efter hvor meget man deltager.
- Certifikat.
- Gaver som t-shirts, bøger eller gavekort.
- Rejse for at møde personalet bag projektet.

### Hvad får biblioteket ud af det?

Lykkes det at få frivillige engageret i løsningen af en opgave, er der mange oplagte fordele. I dette vil jeg nøjes med at nævne tre:

- *Øget synlighed af egne samlinger.*

I LC-projektet beskrives, hvordan at synligheden på Flickr gav en væsentlig forøgelse af besøg til LCs egen website. Folk nøjes altså ikke med bare at se på billederne på Flickr, men følger linket. LC rapporterer en stigning på 20 % i de 5 måneder, hvor projektet blev målt i forhold til det foregående år. De rapporterer samtidig, at en specialsamling, hvor billederne på Flickr løbende blev suppleret, havde et øget besøgstal på 60 %.

En meget interessant observation fra LC var, at billeder på Flickr kom væsentlig højere op på Google's hitliste. De nævner et eksempel, hvor søgningen hos Google gav en femte plads via Flickr, mens det direkte LC-link var på side 5.

- *Bliver del af nye virtuelle communities.*

De frivillige i disse projekter bliver en del af "fanskaren", de hjælper og mærker, at de kan bidrage på en værdifuld måde.

- *Værdiforøgelse af data, som både gør dem mere interessante og som muliggør nye måder at bruge og finde dem.*

Dette punkt er bedst illustreret af et eksempel. Blandt Library of Congress' billeder på Flickr var et af et krigsskib med titlen 'Rochester'. Navnet var ikke verificeret, det eneste man vidste med sikkerhed var, at det kom fra Bain News Service. I løbet af 24 timer havde brugerne ikke alene bidraget at verificere skibets navn, men også med oplysninger om at skibet oprindeligt hed USS New York, var bygget i 1893, hed USS Saratago i perioden 1911-1917, hvorefter det blev til USS Rochester. Yderligere blev rapporteret at det i december 1941 blev sunket af den amerikanske flåde, så det ikke faldt i japanske hænder og også hvor det var sunket og at det var et populært dykkermål.




Library of Congress' billede af 'USS Rochester' på Flickr.

### Overvejelser i forbindelse med crowdsourcing

Hvorfor fejler nogle projekter? Flere biblioteker har arbejdet med, at få brugerne til at tilføje tags til deres katalog, men uden det store held samtidig med, at fx LibraryThing er succesfuld og nu har aftaler med biblioteker om at de kan importere tags.

En forklaring kan være, at den mulige skare af frivillige er begrænset til det enkelte bibliotek, og dermed er der ikke nok frivillige og interesse til at skabe det nødvendige moment. En anden, at folk ikke ser nytten af deres arbejde. Alle artikler jeg har læst peger på en meget vigtig faktor: Problemet skal være stort og udfordrende, ja nærmest uoverkommeligt.

Det er vigtigt, at der er en form for belønning, ikke nødvendigvis i kroner og øre, men i anerkendelse, netværk eller viden. Endelig skal værktøjerne til løsning være nemme, sjove og intuitive. 

### Noter

[1] <http://www.familysearch.org>

[2] Distributed Proofreaders: <http://www.pgdp.net/c>; se også:

Rose Holley (2010): *Tagging full text searchable articles: An overview of social tagging activity in historic Australian Newspapers*, August 2008-August 2009. D-Lib Magazine, vol 16, no 1/2, 2010.

[doi:10.1045/january2010-holley]

[3] [http://www.flickr.com/photos/library\\_of\\_congress](http://www.flickr.com/photos/library_of_congress)

[4] Picture Australia: <http://www.pictureaustralia.org>

[5] <http://www.bbc.co.uk/ww2peopleswar> og WW2 People's War: Project History: How the site worked: User Journey. [http://www.bbc.co.uk/ww2peopleswar/about/project\\_07.shtml](http://www.bbc.co.uk/ww2peopleswar/about/project_07.shtml)

[6] Dansk Demografisk database, <http://www.ddd.dda.dk>

[7] <http://www.wired.com/wired/archive/14.06/crowds.html>

[8] Rose Holley (2010) Crowdsourcing: How and Why Should Libraries Do It? D-Lib Magazine, vol 16, no 3/4, 2010. [doi: 10.1045/march2010-holley]

[9] Michelle Springer, Beth Dulabahn, Phil Michel, Barbara Natanson, David Reser, David Woodward, and Helena Zinkham (2008): *For the Common Good: The Library of Congress Flickr Pilot Project, October 30, 2008*. [www.loc.gov/rr/print/flickr\\_report\\_final.pdf](http://www.loc.gov/rr/print/flickr_report_final.pdf)

[10] *Capturing the power of the crowd and the challenge of community collections*

<http://www.jisc.ac.uk/media/documents/publications/programme/2010/communitycollectionscrowd.pdf>