



UiO : Universitetsbiblioteket

## Robotløsning (RPA) for effektivisering av arbeidsprosesser på UBO

Elin H. Frøshaug, Digitale tjenester, Universitetsbiblioteket i Oslo



## Hvorfor begynte vi med dette?

- Egenarkivering: En kopi av en vitenskapelig artikkel lastes opp til institusjonens vitenarkiv for tilgjengeliggjøring der
- Alle opplastinger må rettighetsklareres
  - Sjekke Sherpa Romeo-basen for utgivers politikk
  - Sjekke direkte hos utgiver for om den er publisert åpent med en Creative Commons-lisens
  - Registrere mulige lisenser og embargoer i vitenarkivet
  - Kanskje fikse på manglende eller gale bibliografiske data
  - Masse frem og tilbake mellom systemer

## Endring i egenarkiveringspolitikk 2017

- UiO hatt Open Access-politikk i flere år, men relativt begrenset egenarkivering
- Endring i desember 2017, obligatorisk å laste opp kopi av alle fagfelleverderte artikler til DUO via Cristin, dvs f.o.m. artikler med publiseringsår 2017.
- Frist for opplasting for 2017 var 31.januar 2018.
- Enorm økning av opplastinger

## Havnet voldsomt bakpå

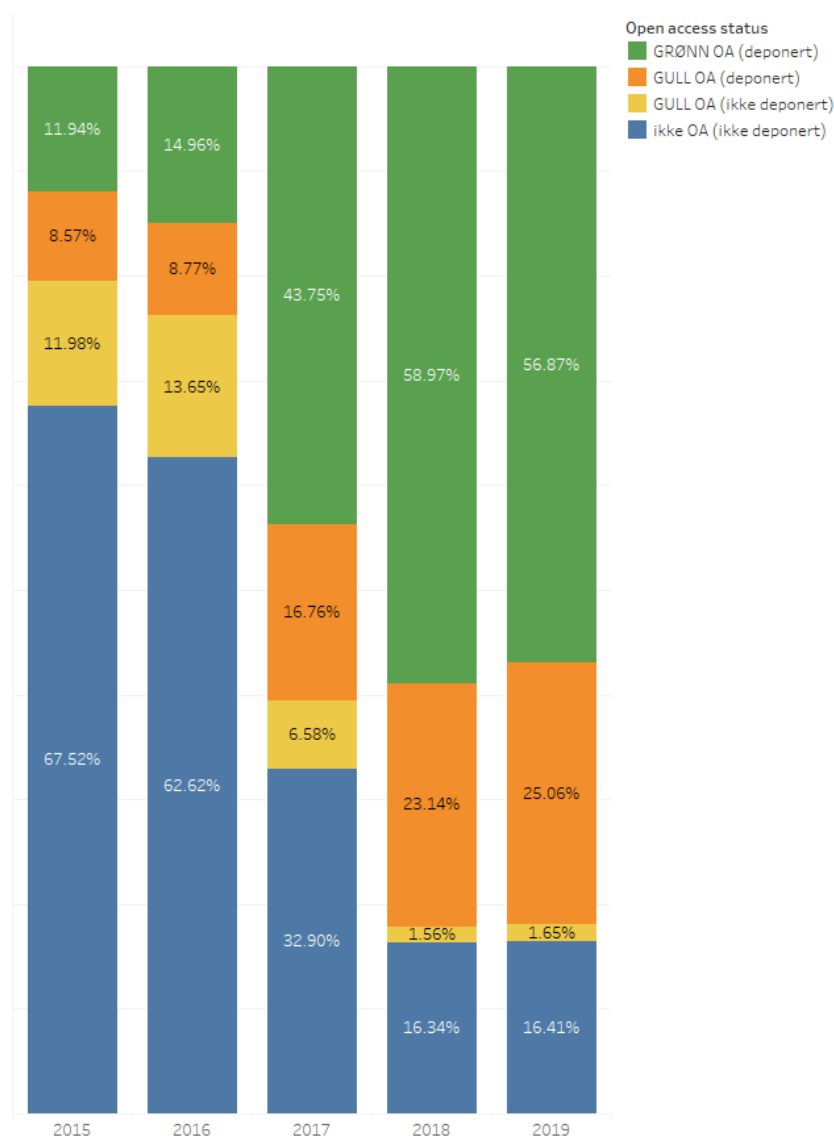
- Per 1. mars 2018: 3000 opplastinger til behandling
- $3000 * 10 \text{ min} = 30.000 \text{ min} / 60 = 500 \text{ timer} / 37.5 = 13.3$   
uker / 4 = 3,3 mnd
- Visste vi ikke fikk alt – totalt antall artikler over 5000, hva da med neste år, og neste år
- I tillegg kapitler og rapporter
- Status:
  - veldig lang behandlingstid
  - vonde skuldre
  - dårlig stemning

### Status Open Access - prosentfordeling

Det vises prosentandel av spesifisert mengde av publikasjoner per år. Rapportstatus "J" gir antall rapporterte NVI-publikasjoner. Andre valg viser NVI-kandidater.

Open Access status måles for vitenskapelige artikler (ISSN). "Ikke relevante poster" i denne sammenhengen er monografier og antologikapitler.

Med "deponert" menes poster som er lastet opp til arkiv via Cristin av minst en av de deltakende institusjoner på posten.



## Sommer 2018

- Fikk høre at Universitetets senter for informasjonsteknologi (USIT) var i gang med RPA
- Hadde møte med USIT for å høre om robotisering
  - For hva var det egentlig
- USIT helt i starten av dette arbeidet
  - Vi skulle holdes orientert og være i loopen

## RPA - Robotic process automation

- Robotisk prosessautomatisering
- Observere hvordan brukere utfører standardiserte oppgaver i ulike brukergrensesnitt og deretter automatisere disse ved å gjenta oppgavene i automatiseringsprogramvaren
- Roboten er programmert til å utføre oppgaver i en bestemt rekkefølge
- Stokk dum! Ikke noe kunstig intelligens her
- Absolutt alt må kodes

## Hva RPA-roboter gjør best

- Hente informasjon fra forskjellige kilder
- Overføre eller registrere informasjon til forskjellige systemer
- Samhandle med mennesker via epost eller arbeidsflyt
- Generere og distribuere dokumenter og rapporter
- Regelbaserte arbeidsprosesser
- Standardiserte og klart definerte arbeidsprosesser
- Digital og strukturert input



## UiOs RPA-prosjekt

- Forankret i og er en del av 'Program for administrativ forbedring og digitalisering' som har som mål å forenkle, forbedre og fornye administrative arbeidsprosesser for å frigjøre ressurser til kjerneoppgavene
- Programmet er viktig for forankring og tilfangst av prosesser
- USIT eier RPA-løsningen, sikrer drift/sikkerhet/arkitektur
- Nøkkel at det er aktiv deltakelse fra flere miljøer
- UiO har bygget kompetanse selv
- God definering av roller og eierskap til administrative systemer og klare prosesseiere

## Høst 2018

- Programvare valgt: Blue Prism
- UBs bibliotekfaglige arbeidsprosesser ikke en del av forbedringsprosjektet
- Fikk likevel være med på kurs i prosesskartlegging og prosessutvikling
- To fra UB med kurs
- Veldig godt samarbeid
- Kompetanseutvikling på UB

Pointer

Expression: Get\_Sherma\_Romeo\_metadata\_Get\_Sherma\_Romeo\_publisher\_exceptions\_Get\_Crossref\_metadata\_Is\_published\_version\_allowed\_Is\_accented\_or\_submitted\_version\_allowed\_G

Link

Block

Process

Page

Action

Decision

Choice

Calculation

Multi Calc

Data Item

Collection

Loop

Note

Anchor

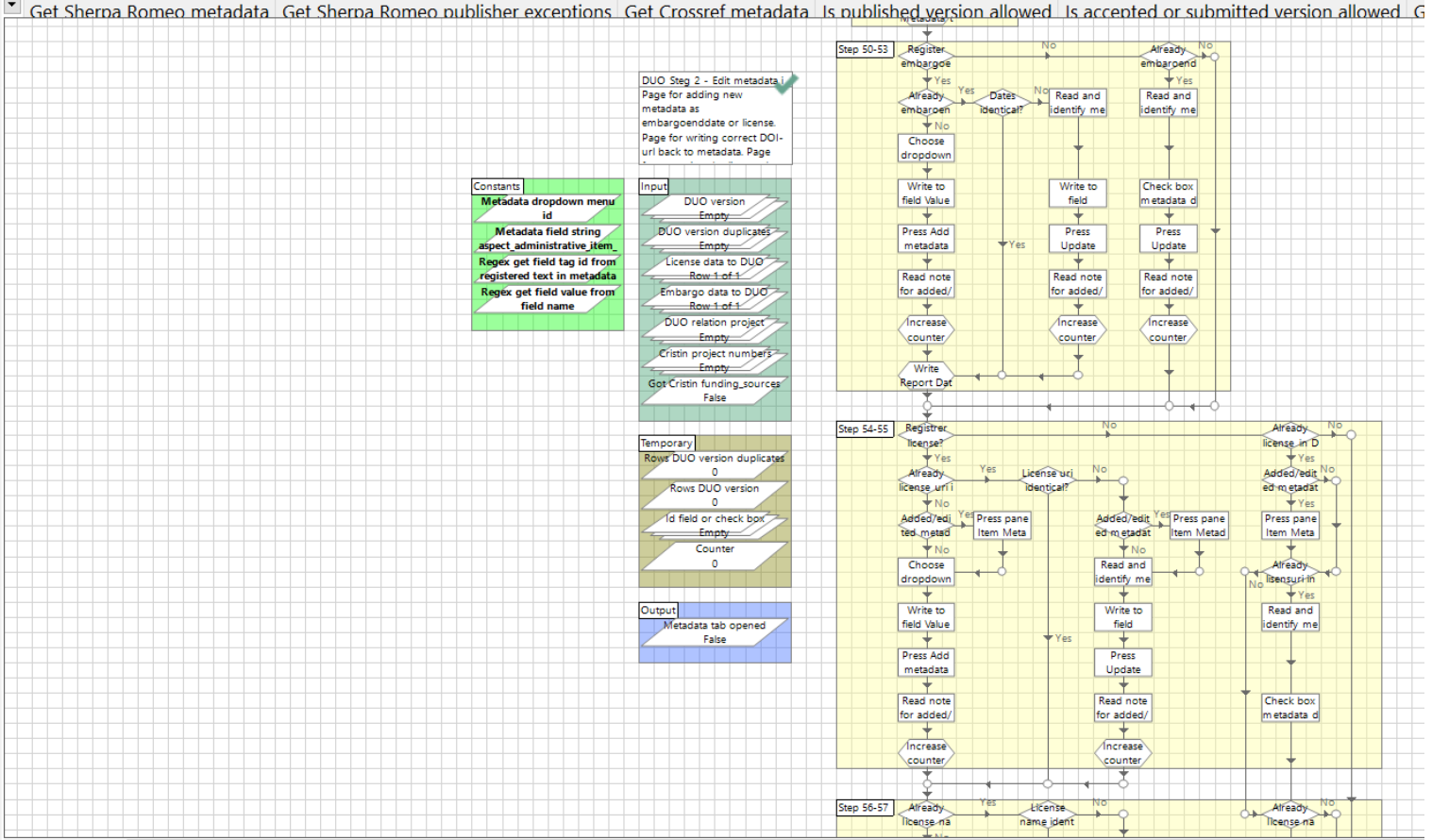
End

Alert

Exception

Recover

Resume



| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver                             |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk                                 |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Tittel i metadatafelt vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)                           |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek                                |

| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver                             |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk                                 |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Tittel i metadatafelt vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)                           |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek                                |

8419  
140 timer

## Dublettsjekk

- Mange dubletter i arbeidsflyten
- Tre kilder til dubletter
  - En post i arbeidsflyt for hver gang det lastes opp fil
  - Dubletter i Cristin
  - Sletting, retting, opplasting på ny post
- Tok ut lister, jobbet i excel, conditional formatting
- Sjekket for hver vi behandlet

## Hva gjør Robin

- Henter alle titler i tabellen for Ledige oppgaver i arbeidsflyten
- Finner like titler
- Sjekker om Cristin-id er lik
  - Hvis lik sletter alle som har færre filer enn den som har flest
- Leverer i rapport potensielle dubletter
  - Lik tittel, men ulik Cristin-id
  - Lik tittel, men likt antall filer
  - Lik tittel, men en av postene på steg 'Velg samling'

| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver                             |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk                                 |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Tittel i metadatafelt vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)                           |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek                                |

2010

8419  
140 timer

134 timer



## Steg 3

- Dette er titler vi har godkjent for DUO
- Henter titler på steg 'Velg samling' i arbeidsflyten
- Henter tilhørighetsinformasjon ved Cristin API
- Sjekker avdelings-, underavdelings, gruppe-numre opp mot en mappingtabell over DUO samlinger
- Krysser av for samlinger artikkel skal speiles i
- Publiserer
- Rapport over alle publiserte og over de som den ikke finner samling for

| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver                             |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk                                 |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Tittel i metadatafelt vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)                           |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek                                |

2010  
134 timer

8419  
140 timer

2240

168 timer

| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver   |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk   |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Studentenes egenregistrerte tittel ved levering vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)   |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek  |

2010  
134 timer

8419  
140 timer

2240  
168 timer

474  
12 timer

## Steg 2 Robin henter saker i DUO

- Henter saker på steg 'Godkjenn/Avvis/Endre'
- Sorterer, finner de eldste
- Henter ut metadata fra DUO, altså det som kom med fra Cristin
  - Tittel, DOI, ISSN, dokumenttype, artikkelversjon, volum, hefte, sidetall, artikkelnummer
  - Interne DUO-id'er
  - Lisens og embargo hvis allerede registrert
  - Prosjektnumre

## Steg 2 Robin sjekker med Cristin

- API
- Sjekker at posten fortsatt finnes
- Henter DOI
- Henter prosjektkoder
- Henter tilhørighet

## Steg 2 Robin sjekker Sherpa Romeo og Crossref

- API
- Sherpa Romeo
  - Henter ut publisher policy
  - Hva er lov hvor for hvilken artikkelversjon
- Crossref
  - Henter ut lisenser, publiseringsdatoer (online og trykt)
  - Bibliografisk informasjon, sammendrag

## Steg 2 Robin kalkulerer

- Opplastet versjon lov i arkiv i følge Sherpa Romeo?
- Hvis embargo i Sherpa Romeo
  - Regner ut sluttdato for denne på bakgrunn av publiseringsdato i Crossref
- Hvis lisenser i Crossref eller Sherpa Romeo
  - Ut fra forkortelser lager fullstendig lisenslenke og navn

## Steg 2 Robin ordner i DUO

- Registrerer embargo og lisenser
- Sletter duplikate metadata-felt for artikkelversjon
- Fikser bibliografiske data
- Fikser prosjektnumre
- Sammendrag



## Steg 2 Robin sender oss rapport

- Egen fane: Ikke tidsskrift-artikler, ikke UiO-tilhørighet
- Den beste opplastede versjonen som er lov
- Om embargo og/eller lisens lagt til
- Om vi har den i DUO fra før
- Postprint lastet opp, men det er CC-lisens i Crossref
- Er lisens eller embargo, men finner ikke ut hva eller hvilken
- Finner ikke info i Sherpa Romeo eller Crossref

| 1  | DUO W | Link      | Item in | Cristin | Duplica   | Conclu | Source   | Check      | Check 2   | Sherpa     | Crossref        | Embargo       | License    | SR Tern   | Title       | Full titl   |
|----|-------|-----------|---------|---------|-----------|--------|----------|------------|-----------|------------|-----------------|---------------|------------|-----------|-------------|-------------|
| 2  | 37037 | https://w | 133493  | Yes     | No        | Post   | SR       | Check lise |           |            | Not CC          | Embargo, publ |            |           | Long- and   | Long- and   |
| 3  | 37017 | https://w | 133468  | Yes     | No        | Post   | SR       |            |           |            | PV CC           | Embargo adde  |            | Publisher | Towards a   | Towards a   |
| 4  | 37018 | https://w | 133469  | Yes     | No        | Post   | SR       |            | articleid | CC from SI | PV CC           |               | License ac | Published | Feasibility | Feasibility |
| 5  | 37019 | https://w | 133470  | Yes     | No        | Post   | SR       | Check lise |           |            | License info mi |               |            | When acc  | Probing su  | Probing su  |
| 7  | 37039 | https://w | 133495  | Yes     | No        | Post   | SR       |            |           |            | Not CC          | Embargo adde  |            | Published | Imagining   | Imagining   |
| 8  | 37043 | https://w | 133499  | Yes     | No        | Post   | SR       |            |           | CC from SI | PV CC           | Embargo adde  | License ac | Must link | Colourful   | Colourful   |
| 9  | 37008 | https://w | 133459  | Yes     | No        | Pre    | SR       |            |           |            | License info mi |               |            | Published | Critics, Mc | Critics, Mc |
| 10 | 37021 | https://w | 133472  | Yes     | No        | PV     | Crossref |            |           |            | CC              |               | License ac |           | The role o  | The role c  |
| 11 | 37009 | https://w | 133460  | Yes     | No        | PV     | Crossref |            | articleid |            | CC              |               | License ac |           | Validator   | Validator   |
| 13 | 37023 | https://w | 133474  | Yes     | No        | PV     | Crossref |            | articleid |            | CC              |               | License ac |           | Intracellu  | Intracellu  |
| 14 | 37042 | https://w | 133498  | Yes     | DOI missi | PV     | SR       | Check lise |           |            | No result beca  | Embargo, publ |            | Published | The Pigm    | The Pigm    |
| 16 | 37047 | https://w | 133503  | Yes     | No        | PV     | SR       |            |           |            | Not CC          |               |            | Publisher | Height De   | Height De   |
| 17 | 37016 | https://w | 133467  | Yes     | Yes       | PV     | Crossref |            |           |            | CC              |               | License ac |           | MRI of tru  | MRI of tru  |
| 18 | 37007 | https://w | 133458  | Yes     | DOI missi |        |          |            |           | No result  | No result beca  |               |            |           | Adapting    | Adapting    |
| 19 | 37000 | https://w | 133451  | Yes     | DOI missi |        |          |            |           | No result  | No result beca  |               |            |           | A Hybrid C  | A Hybrid C  |
| 20 | 37004 | https://w | 133455  | Yes     | No        |        |          |            |           | No result  | No result       |               |            |           | A Self-Org  | A Self-Org  |
| 22 | 36998 | https://w | 133449  | Yes     | No        |        |          |            |           | No result  | No result       |               |            |           | Physical A  | Physical A  |
| 23 | 37006 | https://w | 133457  | Yes     | No        |        |          |            |           | No result  | No result       |               |            |           | A Voice In  | A Voice In  |
| 26 | 36996 | https://w | 133447  | Yes     | No        |        |          |            |           | Check exc  | License info mi |               |            |           | Vocal con   | Vocal con   |

## Hva gjør vi etter Robins rapport?

- Sjekker at filen er riktig
- Fikser det som Robin ikke har klart å fikse
- Kanskje bytter til publisert versjon hvis det er en åpen artikkel og postprint er lastet opp
- Sletter hvis ikke lov
- Klikker Take task og Accept slik at posten legger seg på 'Velg samling' og Steg 3-prosessen gjør den ferdig

| Dublettsjekk  | Steg 1          | Steg 2  | Steg 3  | Studentoppgaver   |
|---|-----------------|---|---|---|
| Ledige oppgaver i arbeidsflyt                                   | Organiser filer | Selve rettighetsklareringen   | Koble til riktig samling og publisere                               | Tittelsjekk   |
| Sletter dubletter og gir liste over andre potensielle dubletter | Tre tasteklikk  | Sjekke SR<br>Sjekke Crossref<br>Finner lisens<br>Finner embargo<br>Retter/legger til data i DUO | Sjekke hvilket institutt forfatter hører til i medfølgende metadata | Studentenes egenregistrerte tittel ved levering vs forside på oppgave |
| Robot (22.3.19)   | Robot (27.2.19) | Robot (10.11.2019)  | Robot (31.5.2019)   | Robot (19.2.2020)   |
| 4 min   | 1 min           | 8 min 30 sek  | 4 min 30 sek  | 1 min 30 sek  |

2010  
134 timer

8419  
140 timer

3550

2240  
168 timer

474  
12 timer

503 timer

## Hva nå? Alma og innlån

- Kommet kortere enn vi trodde før jul
- 'Locate failed'
- Godkjent for kartlegging

## Hvorfor Locate failed

- Søke etter bøker utenfor Alma
- Mye copy/paste
- Mye repetitivt arbeid
- Mange enkle
- La mennesker gjøre de vanskelige søkene
- En god del sjekker før bestilling vurderes
  - Siste tre år, skal til innkjøp
  - Gjentatte bestillinger på samme tittel vurderes for innkjøp

## Hva starter vi med

- Robot sortere excel-filene
  - Nye locate failed, finne alle utgivelser innen siste tre år
  - Finne gjentatte titler
- Søke i Oria, Libris og netpunkt
  - Verifisere
  - Leverer lenke til treff
- Bestille i Sverige og Danmark også? Oppdatere bestillingen i Alma?
  - Vi kartlegger om det er mulig

## Hva har vi lært?

- La roboten gjøre rutineoppgavene
- La mennesker gjøre verdiskapende oppgaver
- La roboten være en del av teamet
  - Roboten gjør sin del, vi gjør vår del
- Roboten gjør det den får beskjed om, ikke noe mer
- Roboten er kjapp
- Kan jobbe til alle døgnets tider
- Forberedt på å gjøre endringer, tilpasninger
- Utvide prosessen underveis



## Oppsummert

- RPA kan brukes på langt flere prosesser enn tradisjonelle administrative prosesser
- Vi tror vi har flere kandidater
- Og når det gjelder de vi har i produksjon...

**We can not go back!**