


Data science og søgning

# Præsentation

- Árni Loftsson
- Randers Bibliotek
- Digital konsulent (“det centrale datakontor”)
- Bibliotekar/programmør

# To projekter

- "1000 hits hjælp!" – inspiration på bibliotekernes hjemmesider
- "Vi ved hvad I vil ha" – bedre søgning på bibliotekernes hjemmesider.

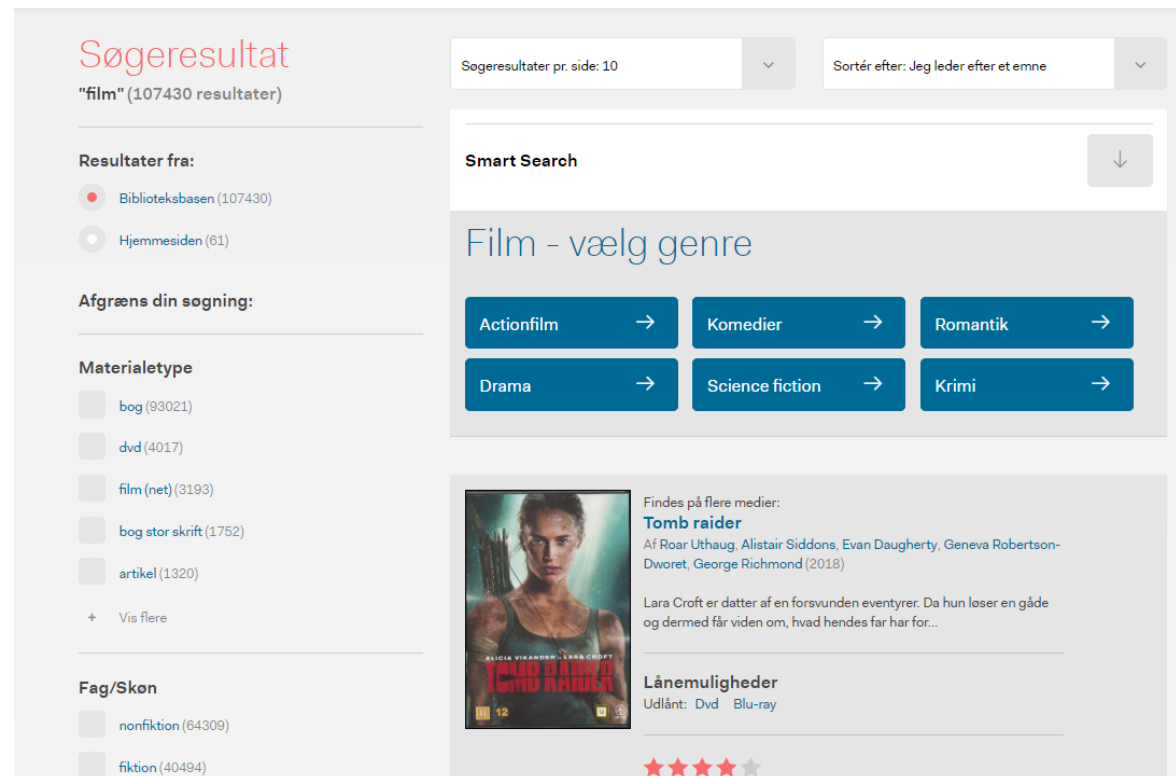


**1000 hits  
- hjælp!**  
Inspiration på DDBCMS hjemmesider

**2017**

**RANDERS  
BIBLIOTEK**

1000 hits - hjælp! | Rapport 2017



**Søgeresultat**  
"film" (107430 resultater)

Søgeresultater pr. side: 10 | Sortér efter: Jeg leder efter et emne

**Resultater fra:**

- Biblioteksbasen (107430)
- Hjemmesiden (61)

**Afgræns din søgning:**

**Materialetype**

- bog (93021)
- dvd (4017)
- film (net) (3193)
- bog stor skrift (1752)
- artikel (1320)
- + Vis flere

**Fag/Skon**

- nonfiktion (64309)
- fiktion (40494)

**Smart Search**

**Film - vælg genre**

- Actionfilm →
- Komedier →
- Romantik →
- Drama →
- Science fiction →
- Krimi →

**Findes på flere medier:**  
**Tomb raider**  
Af Roar Uthaug, Alistair Siddons, Evan Daugherty, Geneva Robertson-Dworet, George Richmond (2018)

Lara Croft er datter af en forsvunden eventyrer. Da hun løser en gåde og dermed får viden om, hvad hendes far har for...

**Lånemuligheder**  
Udlånt: Dvd Blu-ray

★★★★☆

# Brug af data

- Undersøgelse af de problemer brugerne har
- Lave tekniske løsninger baseret på data
- Test af løsninger efter implementering

# Undersøgelse af søgeadfærd



(italien)

ting/collection/870970-basis:52651360,25

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:53291406,4

ting/collection/870970-basis:53053025,5

ting/collection/870970-basis:53096735,7

ting/collection/870970-basis:52651360,11

ting/collection/870970-basis:52600596,13

ting/collection/870970-basis:53025099,33

ting/collection/870970-basis:52624126,41

ting/collection/870970-basis:53088910,78

ting/collection/870970-basis:52651360,11

ting/collection/870970-basis:52651360,16

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:52600596,24

ting/collection/870970-basis:52600596,24

ting/collection/870970-basis:52651360,28

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:52600596,16

# Undersøgelse af søgeadfærd

- 70% af de poster brugerne klikker på er de 3 første poster i søgeresultatet.
- 50% af de poster brugerne klikker på er den 1. post i søgeresultatet.
- 85% af posterne brugerne klikker på kommer på side 1 (de 10 første søgeresultater).

# Hvor effektiv er vores søgning?

The screenshot shows the Randers Bibliotek website interface. At the top, there is a search bar with the text "italien" and a search button. The website header includes the logo "RANDERS BIBLIOTEK", a clock icon for "Åbningstider", and a "Log ind" button. Below the search bar, there is a navigation menu with links for "Forsiden", "Inspiration", "Det sker", "Låner", "Info", "Digitalt", "Børn", and "Studerende".

The main content area displays the search results for "italien" (2188 resultater). On the left side, there are filters for "Resultater fra:" (Biblioteksbasen (2188), Hjemmesiden (0)), "Afgræns din søgning:", and "Materialetype" (bog (787), cd (musik) (379), ebog (335), tidsskrift (net) (148), lydbog (net) (108), and "Forfatter" (giuseppe verdi (112), giacomo puccini (93), inhann sebastian barb (83)).

The search results are displayed in a list format. The first result is "Passione! : mennesker og mad i Italien" by Morten Beiter (2014). It is described as a culinary journey through Italy from the south to the north. The result shows a book cover with a man in a white shirt and a red apron. Below the title, it says "Lånemuligheder" and "Online: Ebog Bog". There are five stars below the result.

The second result is "Jamie i Italien : italiensk køkken når det er bedst" by Jamie Oliver (2018). It is described as over 130 Italian recipes for the year. The result shows a book cover with Jamie Oliver. Below the title, it says "Lånemuligheder" and "Udlånt: Bog". There are five stars below the result.

At the bottom of the page, there is a red button that says "Spørg Biblioteksvagten".



#### Resultater fra:

- Bibliotekbasen (2189)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (789)
- cd (musik) (379)
- ebog (334)
- tidskrift (net) (148)
- lydbog (net) (108)
- + Vis flere

#### Fag/Skøn

- fiktion (822)
- nonfiktion (802)

#### Målgruppe

- voksenmaterialer (2002)
- bornematerialer (286)

#### Forfatter

- giuseppe verdi (112)
- giacommo puccini (93)
- johann sebastian bach (83)
- wolfgang amadeus mozart (79)
- antonio vivaldi (74)
- + Vis flere

#### Årstal

- 2018 (192)
- 2017 (209)
- 2016 (227)
- 2015 (146)
- 2014 (166)



Bog:  
**Jamie i Italien : italiensk køkken når det er bedst**  
Af Jamie Oliver (2018)

Over 130 italienske opskrifter til hele året

#### Lånemuligheder

Udlånt: Bog



Findes på flere medier:  
**Den skønne Helene**  
Af Birgitte Jørgov (2016)

Krimi. Helene - løst ansat ved Nationalmuseet - finder på speget vis ud af, at et opmagasineret elfenbenshovede er et...

#### Lånemuligheder

Online: Ebog Lydbog (net)



Findes på flere medier:  
**Vederfarelse**  
Af Bodo Kirchoff (2018)

Den tidligere forlægger Reither finder en mystisk bog på biblioteket uden titel men med et forfatternavn på. Han tage...

#### Lånemuligheder

Online: Ebog Lydbog (net) Bog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**Venezia**  
Af Hans Gregersen (f. 1946) (2018)

Serie  
Strandbergs italienske kulturguides ;

#### Lånemuligheder

Online: Ebog

basis:53025099 or 870970-  
basis:52600596 or 870970-  
basis:53025099 or 870970-  
basis:52624126 or 870970-  
basis:53088910" (8 resultater)

#### Resultater fra:

- Bibliotekbasen (8)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (4)
- ebog (4)

#### Fag/Skøn

- nonfiktion (6)
- fiktion (2)

#### Målgruppe

- voksenmaterialer (8)

#### Forfatter

- dunoan garwood (2)
- jorgen a. leth (2)
- anders grondahl (1)
- frederik dessau (1)
- karen kristensen (1)
- + Vis flere

#### Årstal

- 2018 (1)
- 2017 (4)
- 2016 (3)

#### Emne

- italien (8)
- rejseforere (3)



Ebog:  
**Oplev Italien**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Online: Ebog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Bog:  
**Italien : seværdigheder, autentiske oplevelser**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**De elskende i Toscana : spændingsroman**  
Af Ole Clifford (2017)

På sit dodsleje genoplever Ellen Winther hele sit liv, med kærlighed, mirakler og spændende skæbner. Barnebarnet Rita...

#### Lånemuligheder

Online: Ebog



Ebog:  
**Livet på rejse : eftertanker fra Italien**  
Af Frederik Dessau (2017)

Med udgangspunkt i den sydtoseanske by Montalino rejser forfatteren rundt til forskellige italienske lokaliteter, bl...

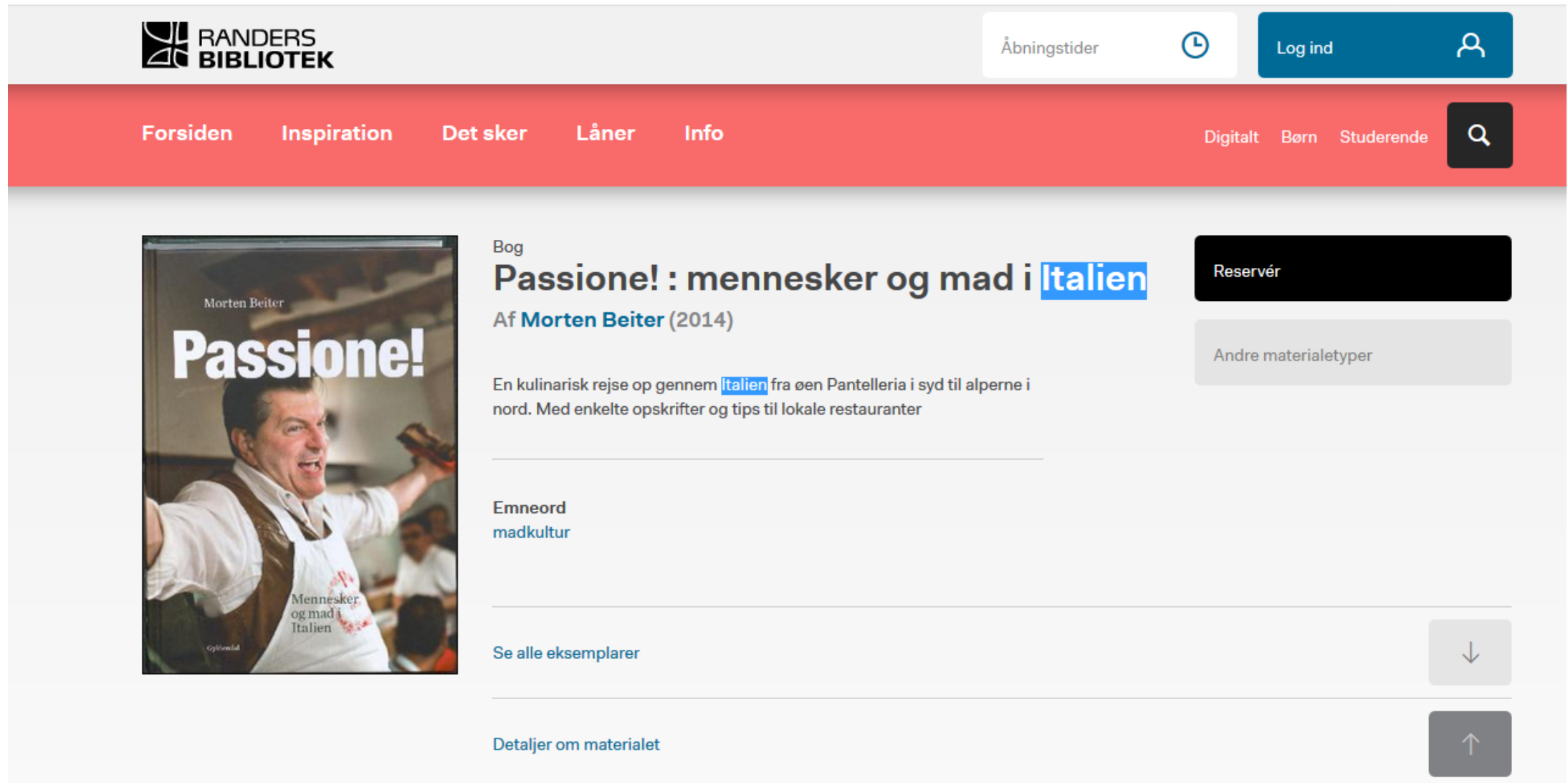
#### Lånemuligheder

Online: Ebog



# Hvor effektiv er vores søgning?


- Det vi ved er at for hver 100 sider med søgeresultater vi viser til brugerne så klikker de på 30 poster.
- Dvs. brugerne klikker på 3 ud hver 100 poster vi viser dem.
- Brugere laver de samme søgninger igen og igen. De 1000 mest hyppige søgninger udgør 30-40% af alle søgninger

# Hvad er problemet med den nuværende søgning



**RANDERS BIBLIOTEK**

Åbningstider  [Log ind](#) 

[Forsiden](#) [Inspiration](#) [Det sker](#) [Låner](#) [Info](#) [Digitalt](#) [Børn](#) [Studerende](#) 

Bog

**Passione! : mennesker og mad i Italien**


Af [Morten Beiter](#) (2014)


En kulinarisk rejse op gennem [Italien](#) fra øen Pantelleria i syd til alperne i nord. Med enkelte opskrifter og tips til lokale restauranter

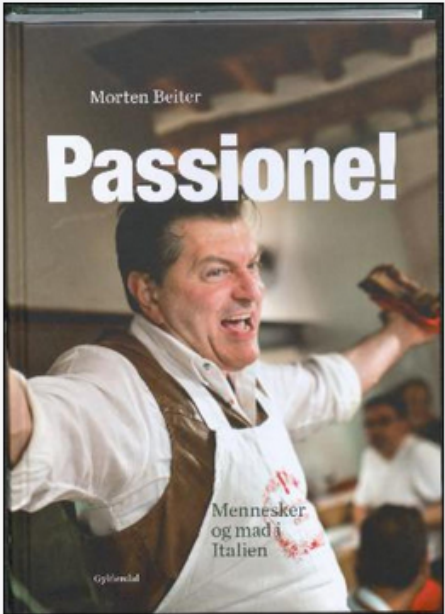
[Reservér](#)

[Andre materialetyper](#)

Emneord  
[madkultur](#)

[Se alle eksemplarer](#) 

[Detaljer om materialet](#) 



# Hvad er problemet med den nuværende søgning

- Vores nuværende søgning er baseret på et match i mellem søgeord og et eller flere ord i metadataerne.

# Løsning:

- Fremadrettet så skal vi opsamle store mængder data vedrørende brugernes søgeadfærd på vores hjemmesider og bruge dem på intelligent vis til at lave bedre søgeresultater.
- = Data science

# ”Vi ved hvad I vil ha” projektet

The screenshot displays a library search interface with the following sections:

- Resultater fra:** Bibliotekbasen (2188) and Hjemmesiden (0).
- Afgræns din søgning:** A section for refining the search.
- Materialetype:** Filters for bog (767), od (musik) (379), ebog (336), tidskrift (net) (148), and lydbog (net) (108).
- Forfatter:** Filters for giuseppe verdi (112), giaoomo puocini (93), johann sebastian baoh (83), wolfgang amadeus mozart (79), and antonio vivaldi (74).
- Årstal:** Filters for years 2018 (189), 2017 (209), 2016 (227), 2015 (146), and 2014 (188).
- Smart Search:** Search bar with 'italien', filters for Søgninger - år (12), Søgninger - 30 dage (6), Sogesoore (0.25), and Status (Inaktiv).
- Alternativ søgestreng og reference:** A section for alternative search terms.
- Fremhæv og boost:** A section for highlighting and boosting results.
- Sortering:** A section for sorting results.
- Buttons:** 'Go to overview' and 'Gem' buttons.

The search results section shows two items:

- Passione! : mennesker og mad i Italien** by Morten Beiter (2014). Description: En kulinarisk rejse op gennem Italien fra oen Pantelleria i syd til alperne i nord. Med enkelte opskrifter og tips ti...  
Lånemuligheder: Bog: Online, Ebog, Bog.
- Turen går til Venedig & Norditalien** by Preben Hansen (f. 1956) (2015). Description: Turen går til : Politikens rejsebøger :  
Lånemuligheder: Tilgængelig: Bog.

# ”Vi ved hvad I vil ha” projektet

- Et forsøg på fikse søgning manuelt for de mest populære søgninger
- De 1000 mest populære søgninger udgør 30-40% af alle søgninger
- Modul i DDBCMS
- Gør det muligt at angive alternativ søgestreng, booste bestemte materialer og udvælge bestemte materialer på de første pladser.
- Nødløsning her og nu.

#### Resultater fra:

- Bibliotekbasen (2189)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (789)
- cd (musik) (379)
- ebog (334)
- tidskrift (net) (148)
- lydbog (net) (108)
- + Vis flere

#### Fag/Skøn

- fiktion (822)
- nonfiktion (802)

#### Målgruppe

- voksenmaterialer (2002)
- bornematerialer (286)

#### Forfatter

- giuseppe verdi (112)
- giacommo puccini (93)
- johann sebastian bach (83)
- wolfgang amadeus mozart (79)
- antonio vivaldi (74)
- + Vis flere

#### Årstal

- 2018 (192)
- 2017 (209)
- 2016 (227)
- 2015 (146)
- 2014 (166)



Bog:  
**Jamie i Italien : italiensk køkken når det er bedst**  
Af Jamie Oliver (2018)

Over 130 italienske opskrifter til hele året

#### Lånemuligheder

Udlånt: Bog



Findes på flere medier:  
**Den skønne Helene**  
Af Birgitte Jørgov (2016)

Krimi. Helene - løst ansat ved Nationalmuseet - finder på speget vis ud af, at et opmagasineret elfenbenshovede er et...

#### Lånemuligheder

Online: Ebog Lydbog (net)



Findes på flere medier:  
**Vederfarelse**  
Af Bodo Kirchoff (2018)

Den tidligere forlægger Reither finder en mystisk bog på biblioteket uden titel men med et forfatternavn på. Han tage...

#### Lånemuligheder

Online: Ebog Lydbog (net) Bog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**Venezia**  
Af Hans Gregersen (f. 1946) (2018)

Serie  
Strandbergs italienske kulturguides ;

#### Lånemuligheder

Online: Ebog

basis:53025099 or 870970-basis:52600596 or 870970-basis:53025099 or 870970-basis:52624126 or 870970-basis:53088910" (8 resultater)

#### Resultater fra:

- Bibliotekbasen (8)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (4)
- ebog (4)

#### Fag/Skøn

- nonfiktion (6)
- fiktion (2)

#### Målgruppe

- voksenmaterialer (8)

#### Forfatter

- dunoan garwood (2)
- jorgen a. leth (2)
- anders grondahl (1)
- frederik dessau (1)
- karen kristensen (1)
- + Vis flere

#### Årstal

- 2018 (1)
- 2017 (4)
- 2016 (3)

#### Emne

- italien (8)
- rejseforere (3)



Ebog:  
**Oplev Italien**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Online: Ebog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Bog:  
**Italien : seværdigheder, autentiske oplevelser**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**De elskende i Toscana : spændingsroman**  
Af Ole Clifford (2017)

På sit dødsleje genoplever Ellen Winther hele sit liv, med kærlighed, mirakler og spændende skæbner. Barnebarnet Rita...

#### Lånemuligheder

Online: Ebog



Ebog:  
**Livet på rejse : eftertanker fra Italien**  
Af Frederik Dessau (2017)

Med udgangspunkt i den sydtoseanske by Montalino rejser forfatteren rundt til forskellige italienske lokaliteter, bl...

#### Lånemuligheder

Online: Ebog



# Mere generel løsning

- Brug data omkring hvad brugerne rent faktisk klikker på til at forbedre ranking:

nicole kidman (søgning)

ting/object/870970-basis:52290171,6 (på plads nr. 6 i søgeresultatet)

ting/object/870970-basis:52290171,6 (på plads nr. 6 i søgeresultatet)

Udtræk af vores logfil.

# Flere data fra logfilen

(italien)

ting/collection/870970-basis:52651360,11

ting/collection/870970-basis:52651360,16

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:52600596,24

ting/collection/870970-basis:52600596,24

ting/collection/870970-basis:52651360,28

ting/collection/870970-basis:54133235,2

ting/collection/870970-basis:52600596,16

# Algoritmen:

- Når et materiale længere ned i søgeresultatet får mange klik tæller det positivt og den rykker op.
- Når et materiale højt i søgeresultatet ikke får nogen klik så tæller det negativt og den rykker ned.
- Algoritmen kan indstilles til forskellig følsomhed dvs. at et materiale rykker hurtigt eller langsomt op og ned.

#### Resultater fra:

- Bibliotekbasen (2189)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (789)
- cd (musik) (379)
- ebog (334)
- tidskrift (net) (148)
- lydbog (net) (108)
- + Vis flere

#### Fag/Skøn

- fiktion (822)
- nonfiktion (802)

#### Målgruppe

- voksenmaterialer (2002)
- bornematerialer (286)

#### Forfatter

- giuseppe verdi (112)
- giacommo puccini (93)
- johann sebastian bach (83)
- wolfgang amadeus mozart (79)
- antonio vivaldi (74)
- + Vis flere

#### Årstal

- 2018 (192)
- 2017 (209)
- 2016 (227)
- 2015 (146)
- 2014 (166)



Bog:  
**Jamie i Italien : italiensk køkken når det er bedst**  
Af Jamie Oliver (2018)

Over 130 italienske opskrifter til hele året

#### Lånemuligheder

Udlånt: Bog



Findes på flere medier:  
**Den skønne Helene**  
Af Birgitte Jørgov (2016)

Krimi. Helene - løst ansat ved Nationalmuseet - finder på speget vis ud af, at et opmagasineret elfenbenshovede er et...

#### Lånemuligheder

Online: Ebog Lydbog (net)



Findes på flere medier:  
**Vederfarelse**  
Af Bodo Kirchoff (2018)

Den tidligere forlægger Reither finder en mystisk bog på biblioteket uden titel men med et forfatternavn på. Han tage...

#### Lånemuligheder

Online: Ebog Lydbog (net) Bog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**Venezia**  
Af Hans Gregersen (f. 1946) (2018)

Serie  
Strandbergs italienske kulturguides ;

#### Lånemuligheder

Online: Ebog

basis:53025099 or 870970-  
basis:52600596 or 870970-  
basis:53025099 or 870970-  
basis:52624126 or 870970-  
basis:53088910" (8 resultater)

#### Resultater fra:

- Bibliotekbasen (8)
- Hjemmesiden (0)

#### Afgræns din søgning:

#### Materialetype

- bog (4)
- ebog (4)

#### Fag/Skøn

- nonfiktion (6)
- fiktion (2)

#### Målgruppe

- voksenmaterialer (8)

#### Forfatter

- dunoan garwood (2)
- jorgen a. leth (2)
- anders grondahl (1)
- frederik dessau (1)
- karen kristensen (1)
- + Vis flere

#### Årstal

- 2018 (1)
- 2017 (4)
- 2016 (3)

#### Emne

- italien (8)
- rejseforere (3)



Ebog:  
**Oplev Italien**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Online: Ebog



Bog:  
**Turen går til Venedig & Norditalien**  
Af Preben Hansen (f. 1956) (2018)

Serie  
Turen går til : Politikens rejsebøger ;

#### Lånemuligheder

Tilgængelig: Bog



Bog:  
**Italien : seværdigheder, autentiske oplevelser**  
Af Dunoan Garwood (2018)

Serie  
Lonely Planet ;

#### Lånemuligheder

Tilgængelig: Bog



Ebog:  
**De elskende i Toscana : spændingsroman**  
Af Ole Clifford (2017)

På sit dødsleje genoplever Ellen Winther hele sit liv, med kærlighed, mirakler og spændende skæbner. Barnebarnet Rita...

#### Lånemuligheder

Online: Ebog



Ebog:  
**Livet på rejse : eftertanker fra Italien**  
Af Frederik Dessau (2017)

Med udgangspunkt i den sydtoseanske by Montalino rejser forfatteren rundt til forskellige italienske lokaliteter, bl...

#### Lånemuligheder

Online: Ebog

# Udfordringer

- Skal håndtere nye materialer som ikke har så meget data endnu
- Skal kunne forudsige at et materiale vil blive populært for en given søgning. F.eks. ny bog af en populær forfatter.
- Skal kunne glemme gamle tendenser som ikke er relevante mere.
- Skal bibeholde f.eks. aspekter af et søgeresultat. F.eks. søgning på "rom" kan være byen eller drikken. Hvis det ene aspekt dominerer så skal der stadigvæk være mulighed for det andet aspekt at fremtræde i søgeresultaterne.

# Muligheder fremadrettet

- Opdagelse af materialer som er langt nede i søgeresultatet men som er meget relevante for en given søgning.
- Opdagelse af søgninger hvor brugeren har behov for at specificere yderligere for det giver et relevant resultat. F.eks. søgningen "film"
- Opdagelse af søgninger som rammer ved siden af.
- Personalisering

# Generelle principper

- Søgning er ofte den eneste kontakt en given bruger har med biblioteket. Mister vi dem der så har vi ofte mistet den eneste mulighed for at nå dem.
- Vi skal betragte søgning og interaktion via hjemmesiden som ligeså vigtig som hver anden form for brugerkontakt.
- Den bedste løsning er sandsynligvis 95% maskinel – 5 % manuel. Dvs. der vil være behov for at kombinere den viden bibliotekarer har i deres omgang med brugerne med den viden vi kan lære og automatisere udefra data.

# Hvordan kommer vi videre

- Randers Bibliotek har logfiler fra starten af 2018. Det kan bruges som første spadestik til test og udvikling.
- Vi skal hurtigst muligt implementere logning af vores søgninger og søgeadfærd i alle DDBCMS biblioteker.
- Den indbyggede boost funktion i brønden giver muligheder for test af konceptet uden større mængder af udvikling.
- Lav A/B test for at afgøre om en given metode er bedre end en anden.