

# Vision og strategi for søgning

16/03/2021



# KOMBIT

Kommunernes it-fællesskab

---

Titel:	Vision og strategi for søgning
Sideantal:	21
Sidste ændringsdato:	16/03/2021
Sidste ændring af:	Mads Bondo Dydensborg
Eksporteringsdato:	16/03/2021
Eksporteret af:	Mads Bondo Dydensborg
Rumnavn:	Mads Bondo Dydensborg
URL:	<a href="https://dbcjira.atlassian.net/wiki/spaces/~78680411/pages/2074247386">https://dbcjira.atlassian.net/wiki/spaces/~78680411/pages/2074247386</a>

## Indholdsoversigt

<b>1</b>	<b>Vision</b> .....	<b>3</b>
1.1	Den almindelige slutbruger .....	3
1.2	Interaktion med biblioteket anno 2025 .....	4
1.2.1	Eksempler på interaktion .....	4
1.3	Elementer i visionen .....	6
1.3.1	Visionens pejlemærker .....	6
1.3.2	Supplerende mål .....	7
<b>2</b>	<b>Strategi</b> .....	<b>7</b>
2.1	Principper for realisering .....	7
2.2	Den brugerrettede model for data, API'er og services – fra kompleksitet til enkelhed .....	8
2.3	Kvalitet og åben evaluering – brugerrettet objektiv kvalitet .....	9
2.4	Søgning via naturligt sprog, dialog og stemmestyring .....	9
2.5	Integration af brugerdata og bibliotekernes kuratering .....	10
2.5.1	Eksempel: Enkelhed og åbenhed fremmer innovation .....	11
2.6	Heterogene data og adgang til resurser uden for sektoren .....	11
<b>3</b>	<b>Projektoverblik</b> .....	<b>12</b>
<b>4</b>	<b>Roadmap</b> .....	<b>19</b>
<b>5</b>	<b>Udvikling af eksisterende søgeinfrastruktur</b> .....	<b>21</b>

# 1 Vision

Bibliotekerne skal, som bekendt, fremme oplysning, uddannelse og kulturel aktivitet med udgangspunkt i rollen som kurator og inspirator. Bibliotekerne lever i vid udstrækning op til rollen som den vigtigste formidler af litteratur til fritids- og lystlæsning for både børn og voksne og til at understøtte den generelle evne og lyst til læring gennem læsning og anden form for aktivitet med bibliotekernes digitale og fysiske materialer og tilbud. Bibliotekerne gør således en stor samfundsmæssig forskel, men at fortsætte med dette hænger i højere og højere grad sammen med evnen til at være relevant i det digitale rum og kunne fastholde og øge antallet af brugere, besøg og udlån der. Et centralt element i det digitale rum er søgning i og præsentation af biblioteksindhold i bred forstand, og søgning er desuden med til at binde det fysiske og digitale bibliotek sammen.

## Visionen

Gennem FBI skal der udvikles og udstilles en søgeinfrastruktur der

- opfylder muligheden for at indfri og overgå den almindelige slutbrugers forventninger til søgning i og præsentation af bibliotekernes samlinger, aktiviteter og tilbud
- understøtter brug af data, indhold og services i og uden for sektoren
- danner basis for leverandørers og interessenters innovation i og uden for sektoren
- medfører en endnu større aktivering af samlingerne, herunder de elektroniske kilder og long tail

Denne vision dækker søgning i den bredeste forstand, inklusiv fx datadrevne services der, baseret på brugerens historik, præferencer og nysgerrighed, kan komme med anbefalinger, forslag, mv. Den dækker såvel fremsøgning som præsentation af indhold.

Visionen lægger et helt afgørende fokus på, at der skal skabes løbende resultater, der skaber værdi for brugerne og kan anvendes hurtigt og enkelt af alle leverandørerne af bibliotekernes brugergrænseflader.

For at fremme muligheden for at mange slags leverandører kan udvikle grænseflader og biblioteksprodukter ovenpå søgeinfrastrukturen, skal FBI investere i en række af forenklinger, der langt bedre gør det muligt at ramme brugernes behov og tilgang. Forenklingerne – og en konsekvent og ambitiøs brugerorientering – skal understøtte udviklingen af nye supplerende interaktionsmodeller for søgning, repræsentationer af data, datamodeller og API'er.

En søgeinfrastruktur der er nemmere at tage i anvendelse for de nuværende leverandører vil også kunne invitere nye spillere ind, hvilket vil muliggøre større konkurrence i markedet og vil komme bibliotekerne til gode i form af muligheden for lavere totalomkostninger til it-løsninger. Enkelhed og en søgeinfrastruktur, der rammer de funktionelle og datamæssige krav på tværs af alle former for biblioteksgrænseflader, vil forenkle arbejdet med den enkelte grænseflade og vil minimere dobbeltudvikling. Komplexiteten skal således indkapsles i søgeinfrastrukturen for at gøre anvendelsen af den mere enkel, nemmere og dermed billigere.

Visionen inkluderer ligeledes mål om at søgeinfrastrukturen skal realisere de behov og ønsker der kan forbedre lokale grænseflader som DDB-CMS-NEXT, BiblioteketAPP, EASY-suiten og Læsekompasset – ligesom børnegrænseflader og nationale løsninger som [bibliotek.dk](http://bibliotek.dk) skal tilgodeses. Visionen rummer således en stærk intention om omfattende inddragelse af de interessenter der allerede laver grænseflader ovenpå FBI, så både kendte og fremtidige krav kan indarbejdes når visionen realiseres.

## 1.1 Den almindelige slutbruger

Visionen – og dermed også strategien – er *primært* målrettet hvad vi har valgt at kalde "den almindelige slutbruger". Denne anvender biblioteket primært til fritids- og lystformål, som læsning af skønlitteratur, lytning til lyd-bøger, visning af film eller tilegnelse af faglitteratur relateret til hobby, interesser eller tilsvarende. Det er brugere der i en

årrække har haft for dårlige betingelser i søgeadgangen til bibliotekernes samlinger. Dette har været dokumenteret i forskellige sammenhænge og fremgår blandt andet af dataopsamlinger i [bibliotek.dk](http://bibliotek.dk) og har været i fokus i forskellige biblioteksprojekter, fx i Randers folkebibliotek. Disse brugere har måtte leve med en uhensigtsmæssig ranking af søgeresultater og for komplekse krav til søgeadfærd.

Det medfører dog ikke, at andre brugerbehov skal negligeres. Alle typer af brugsscenarier kan få glæde af visionen og strategiens implementering, men det er denne type bruger der er den primære driver af vision og strategi, særligt når det handler om at gennemføre de nødvendige forenklinger i selve søgningen.

Visse steder tager vision og strategi højde for at visning kan justeres til andre behov end den almindelige brugers, fx gymnasieeleven der ønsker at få vist forlag direkte med resultater, eller lignende. Sådanne brugerbehov afdækkes løbende og vil blive leveret i udviklingsprojekter og i den almindelige vedligeholdelse.

## 1.2 Interaktion med biblioteket anno 2025

Innovationen indenfor brugergrænseflader peger entydigt på at fremtidens søgning ikke kun kommer til at foregå ud fra et søgefelt. Søgning kommer til at foregå gennem eksempler, noget man allerede kender, bibliotekets kuratering og en række andre brugsscenarier og typer af søgning, vi ikke kender endnu.

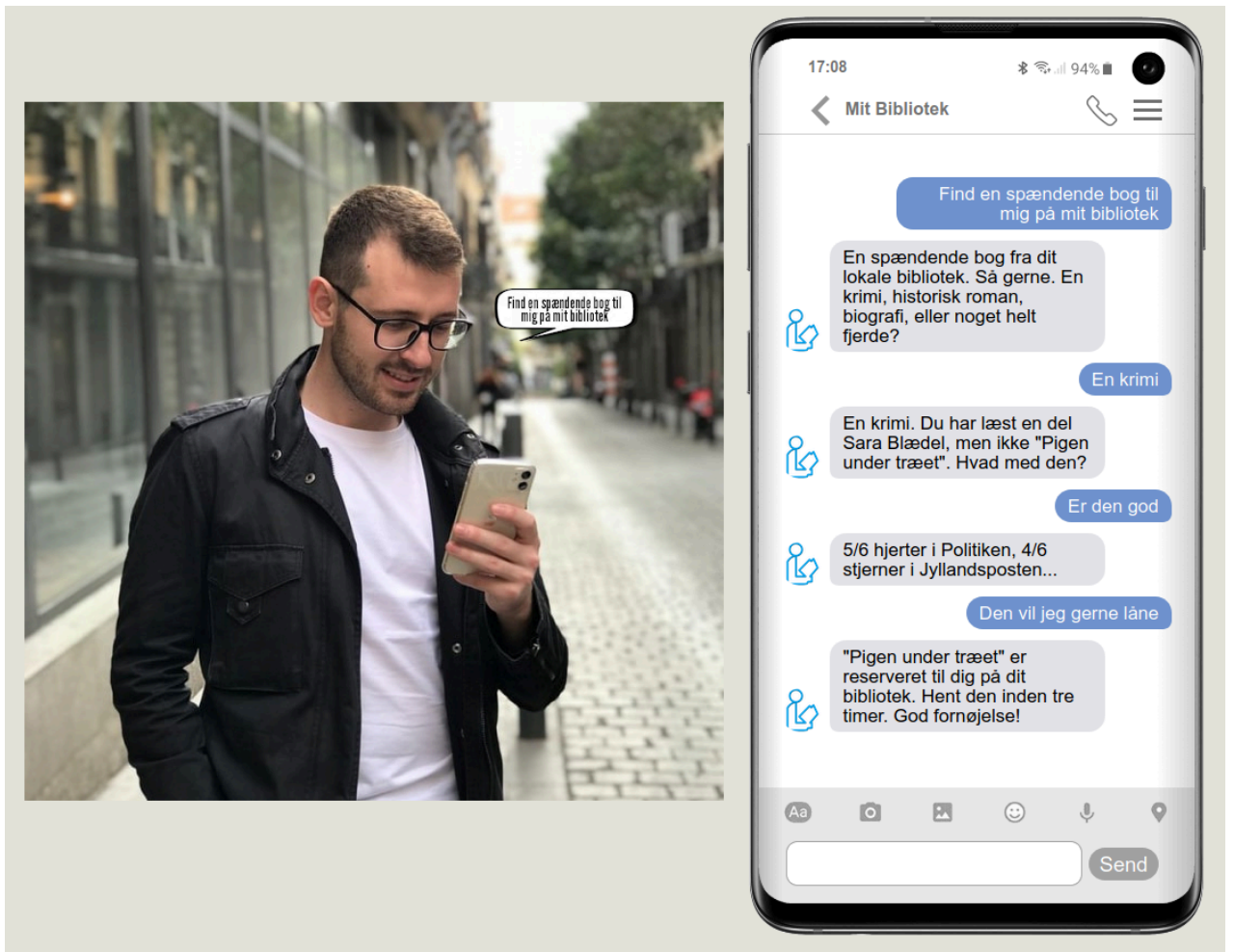
Samlingerne skal gøres lettere tilgængelige for de +60% af borgerne der ikke benytter bibliotekerne og endnu mere relevant for dem der gør. Fokus skal ikke ligge på at diskutere tekniske detaljer. Der skal i stedet opstilles relevante metrikker til at evaluere hvor gode systemerne er, og hvilke tiltag og ændringer der forbedrer systemerne i forhold til vores fælles mål.

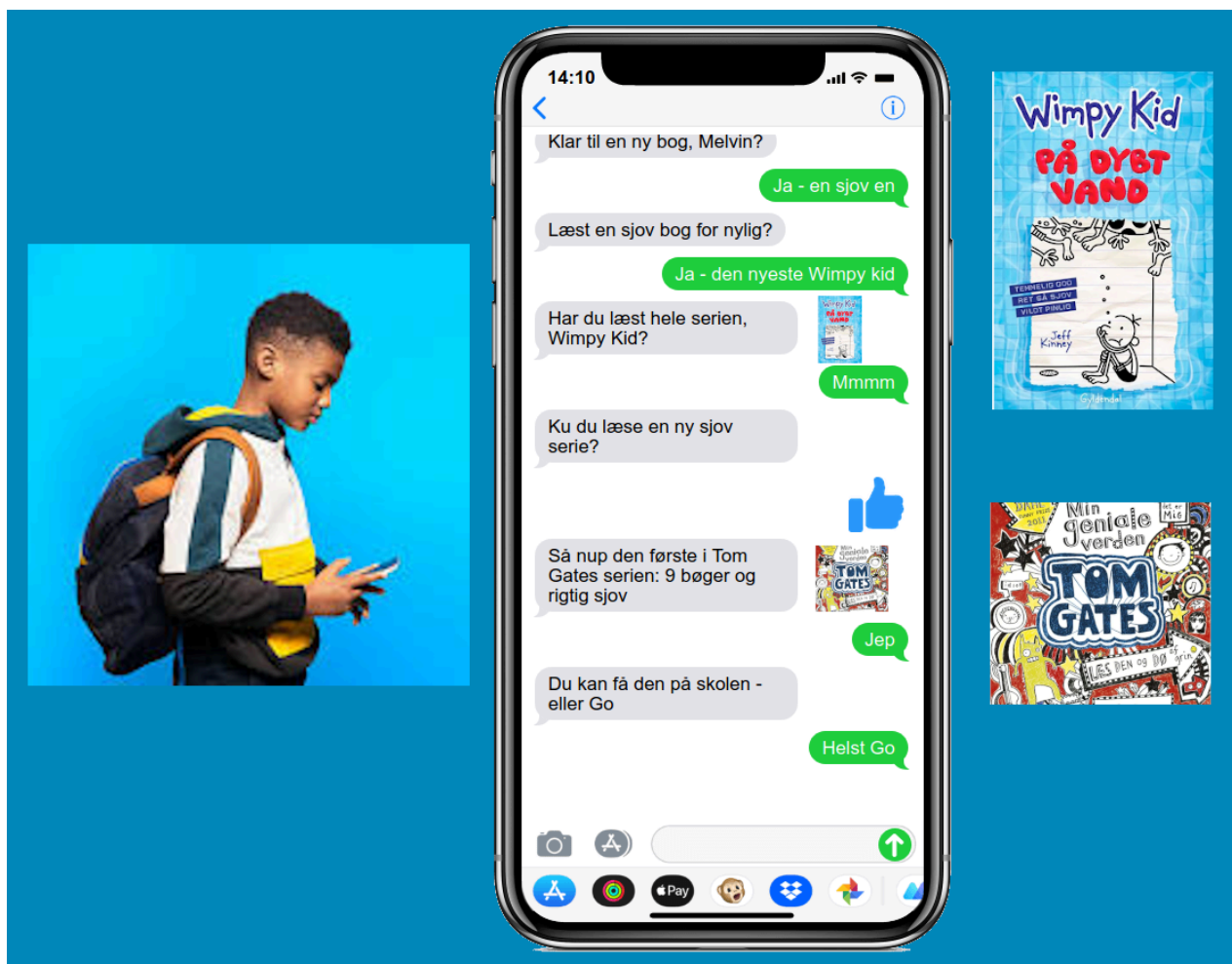
Visionen understøtter naturligvis "søgefelt" og hvad deraf følger, men infrastrukturen/platformen skal også understøtte tilgange der er markant anderledes. De kommercielle indholdsudbydere har ikke tænkt sig at stoppe med at innovere og den retning de sætter vil have betydning for brugernes forventninger og opfattelser af hvordan søgesystemer skal fungere – infrastrukturen/platformen skal kunne understøtte også sådanne tilgange, og må altså ikke tænkes for snævert.

Bemærk at den eksisterende søgeinfrastruktur bevares i denne vision, så de professionelle interessenter, dem med traditionelle søgebehov og kompetence i at søge specifikt i de meget omfattende datamængder, skal fortsat have dækket deres behov.

### 1.2.1 Eksempler på interaktion

Som illustration af de interaktionsmuligheder der skal kunne understøttes, har vi valgt at medtage nogle dialogbaserede interaktioner igennem dokumentet. De samme cases kan også understøttes med andre interaktionsmetoder, fx websites.





### 1.3 Elementer i visionen

Visionen skal understøtte bibliotekernes ønske om at bibliotekernes samlinger åbnes op for en langt større gruppe end tilfældet er i dag.

For at opnå dette, skal gode ideer kunne realiseres langt enklere og af flere aktører end i dag. Det skal være muligt for langt flere at lave innovation inden for området og ramme en større del af den potentielle brugergruppe, samt give den eksisterende brugergruppe flere og bedre muligheder for at opdage og forbruge bibliotekernes materialer og formidling af viden og oplevelser.

At åbne op for mere innovation på området, gerne fra flere aktører, er således en kerneforudsætning for at visionen kan realiseres. Dette kræver at barrieren til entry på området sænkes betydeligt. En kendt barriere er kompleksiteten i de underliggende data og i de eksisterende søgesystemer, API'er, mv - i flere sammenhænge endda også i grænseflader.

I arbejdet med visionen er følgende pejlemærker, supplerende mål og principper blevet identificeret. Disse danner baggrunden for strategien.

#### 1.3.1 Visionens pejlemærker

- En brugerrettet, stærkt forenklet model for søgning.

- Suggestere og recommendere i stærkt fokus for at sikre at samlingerne åbnes op og for at long tail bliver aktiveret, samt for at give mange veje ind i samlingerne, herunder understøtte bibliotekernes kuraterings- og formidlingsarbejde.
- Søgning via dialogbaseret naturligt sprog, udtrykt som fx ønsker tilpasset den enkelte bruger, med viden om denne.
- En søgeplatform der i endnu højere grad kan integrere eller udnytte data fra andre kilder, såsom fuldtekster, databaser, kvalitetskilder på nettet, mv.
- En søgeplatform der endnu nemmere kan integreres i nuværende systemer, som fx DDB-CMS og en række andre systemer, hjemmesider, apps, mv. af flere anvendere.

### 1.3.2 Supplerende mål

- En platform for opsamling og udlevering af brugerspor og BI data på tværs af alle FBI systemer, hvortil eksterne data nemt kan integreres. (Naturligvis med respekt for brugernes ønsker i relation til fx privacy).
- Infrastruktur og evalueringssystemer der giver interessenter mulighed for at få indsigt i og kontinuert påvirke søgeplatformens kvalitet på en objektiv og systematisk måde.
- Optimering af søgeresultater igennem automatiserede processer (såsom Machine Learning/AI (Kunstig intelligens)) baseret på testsæt udarbejdet af fagfolk, samt brugeradfærd.
- Prioritering af kendte problematikker, hvor bibliotekerne ønsker særlige filtreringer og håndtering, fx i forbindelse med manglende forsider, tilgængelighed ift e-resurser og fysiske bøger, mv.
- Søgeinfrastrukturen skal fortsat og udbygget kunne understøtte bibliotekernes ønsker for brugerafgrænsninger og filtreringer, så forskellige brugscenarier understøttes. Et biblioteksprodukt skal kunne veksle mellem nationale og lokale afgrænsninger. Lokale afgrænsninger skal afspejle bibliotekernes heterogene virkelighed, fx at kunne afgrænse til den del af samlingen, der er tilgængelig - fysisk og/eller digitalt .
- Understøttelse af særlige visningsbehov for fx studerende, elever på ungdomsuddannelser, børn, mv.
- Øget fokus på digitale materialers integration i samlingen da fysiske materialer kun er en del af de samlinger som borgerne har adgang til.
- Øget fokus på udstilling/integration af information via Google, Bing, og andre markante søgeplatforme, hvor mange biblioteksbrugere søger.
- Reduktion af udviklings- og vedligeholdelsesomkostninger på services, infrastruktur, mv. De samlede omkostninger til it-plattformen i FBI skal ikke stige, selvom der etableres en ny brugerrettet søgeinfrastruktur.

## 2 Strategi

Visionen skal realiseres igennem følgende hovedområder:

- En enkel, bruger-/anvendelsesrettet, data- og søgemodel i interaktionslaget.
- En fælles/delt evalueringsmodel for kvalitet.
- Søgning via dialog og naturligt sprog.
- Integration af brugerdata og lokal kuratering.
- Integration af heterogene data og resurser uden for sektoren.

Det skal fremhæves, at vision og strategi tager afsæt i tilstedeværelsen af komplekse datamodeller i databrønden, som ikke forlades, da de tjener mange varierede formål, fx præcise og avancerede søgebehov, men at det ses som afgørende, at brugerne og leverandører af slutbrugergrænseflader ikke skal kende disse.

### 2.1 Principper for realisering

Følgende principper skal ligge til grund for alle aktiviteter der skal udmønte strategien:

- En tilgang der er drevet af brugerinddragelse og -undersøgelser, samt anvendelse af adfærdsdata.

- Løbende og tæt dialog med biblioteker, interessenter og leverandører i implementeringen af vision og strategi.
- Komplexiteten skal holdes i FBI's indre, hvor kvalitetsmetadata og avancerede matchsystemer understøtter den enkelhed, og brugervenlighed som brugere og leverandører skal opleve.
- DBC skal i videst mulige omfang ligestille eksterne anvendere med DBC selv – interne services og funktioner åbnes op for andre og inkluderes i Den Åbne Platform.
- En holistisk tilgang til søgning og visning, hvor datagrundlag ses bredere, herunder brug af fuldtekster og fx ustrukturerede data skabt i lokale biblioteker.

Principperne udledes naturligvis i projekter, mv. Princippet om løbende og tæt dialog med interessenter tænkes blandt andet udmøntet i

- større og tidligere åbenhed om tekniske løsninger,
- etablering af alpha-sites/services i større omfang end idag,
- tidlig involvering af leverandører i udarbejdelse af API'er, fastlæggelse af SLA aftaler, mv

og andre tiltag der peger i samme retning.

## 2.2 Den brugerrettede model for data, API'er og services – fra kompleksitet til enkelhed

En brugerrettet datamodel betyder at den skal ramme brugernes opfattelse af, hvordan et givent medie generelt opleves og erkendes, når man søger efter dem, lader sig inspirere, relaterer og på andre måder ønsker at tilegne sig et værk og dets kontekst, et emne eller ophav til et værk. Sådanne væsentlige søge- og visningsentiteter skal understøtte de forestillinger og modeller som helt almindelige brugere har for det pågældende medie og de værker, der er publiceret i dette medie.

Der vil, for alle materialer, være et egentlig værk, som har de basale karakteristika som en bruger forbinder med et værk (titel, ophav, emne, mv). Tilsvarende for en serie. Værk og fx serie vil være egentlige, persistente entiteter der kan adresseres entydigt fremadrettet. Dermed kan referencer til værker fx gemmes i apps, andre søgesystemer, grænseflader, mv. og altid bruges til at referere det pågældende værk.

Det er som fremhævet et mål at datamodellen nemt kan overføres til de datamodeller der findes i de dominerende søgeplatforme på internettet, således at disse data nemmere kan udstilles og bringes i spil fra disse platforme.

I det hele taget skal datamodellen være udformet således at det kræver minimal indsigt i de enkelte datakilders ofte meget heterogene bibliografiske praksis, metadata modeller og formatmæssige lagring for at kunne fremsøge og forbruge data. Hensigten er altså at holde kompleksiteten internt i FBI platformen, og stille en langt mere anvendelig model for data til rådighed for omverdenen men også internt i platformen, hvor det er muligt og ikke gør det svært at opfylde andre og mere avancerede brugerbehov.

Den brugerrettede model skal direkte indeholde de data der efterspørges i fx grænseflader målrettet slutbrugere og det skal være meget nemmere at specificere hvilke specifikke data man har behov for, når man anvender API'er og services, og hurtigt og enkelt tilgå data stillet til rådighed. Dette skal fremme generelt udbredte søgemodeller fra nettet, fx GraphQL eller tilsvarende, således at fx visning af søgeresultater kan tilpasses konkrete brugsscenarier.

Forenklingen i API'er drives iøvrigt i vid udstrækning igennem som konsekvens af en forenklet datamodel. Men, der skal også udstilles API'er der lever op til best practice for udstilling af services i dag. I praksis vil det fx betyde en fuld overgang til REST/JSON, samt relevante autentificeringsmekanismer, der understøtter valgfri komposition af API'er i innovative løsninger designet "i kanten" af branchen: Det skal være muligt for brugere at lade apps og andre grænseflader agere på vegne af dem, uden at skulle afkræves autentificering igen og igen.

Det strategiske mål er her naturligvis at gøre infrastrukturen langt mere tilgængelig for personer, virksomheder, eller andre, der ikke besidder dyb bibliografisk indsigt eller indsigt i de i bibliotekssektoren verden over repræsenterede komplekse søgemodeller som fx CQL. Men, det er også et strategisk mål at forenkle infrastrukturen internt i den slutbrugerrettede del af FBI, både arkitektur mæssigt, og i forhold til større enkelhed i de services der udstilles i den brugerrettede del, og med fokus på at reducere omkostninger ved udvikling, vedligehold og drift.



## 2.3 Kvalitet og åben evaluering – brugerrettet objektiv kvalitet

Det er et strategisk sigte, at kvaliteten defineres igennem systemer der, så objektivt som muligt, gør at brugernes oplevelser, såvel som fagpersoners viden, mv. kan overføres direkte til de mål der ønskes evalueret på. Valideringen af kvaliteten skal åbnes op for langt flere interessenter, så disse, via etablerede systemer, kan inddrages direkte i evalueringen af kvaliteten af søgningerne og forbedringerne af den, fx løbende i faggruppefora, indeholdende f.eks. repræsentanter for bibliotekerne, leverandører, mv.

Det nye i denne tilgang er, at en ønsket kvalitetsforbedring på søgesiden kan måles og gennemføres ud fra den effekt som en ændring har, i stedet for at blive vurderet efter om søgeindekserne er optimale. Ofte kan søgeindekserne ikke evalueres uafhængigt af den søgemaskine der anvender dem og den måde brugerne anvender søgemaskinen på. Så der er behov for at udvide genstandsfeltet for søgeevaluering. Der skal således tages udgangspunkt i mere objektive og brugerrettede modeller, hvor der i forskellige brugerperspektiver måles på, hvad der fremkommer ved en søgning.

Søgekvalitet skal vurderes med udgangspunkt i testsæt (testcollections) der udgøres af et antal materialer der repræsenterer udvalgte søgninger, med tilhørende vurdering af gode og dårlige hits på søgningen. Testsættet kan bruges direkte til at vurdere kvaliteten af en søgemaskine der kun søger i testsættet, ved hjælp af etablerede metrikker inden for området, herunder "recall" og "precision". Blandt andet med udgangspunkt i disse metrikker, kan testsæt således anvendes til at optimere en søgemaskine maskinelt ved brug af machine learning.

Eksempler på søgninger fra et testsæt		
Narrativ	Søgestreng	Ønsket resultat
Du læser serien De syv søstre af Lucinda Riley og leder efter bogen "Månesøsteren". Den danske oversættelse er første valg, men den engelske udgave er også relevant.	månesøsteren	Bogen Månesøsteren af Lucinda Riley
Du leder efter materialer til en skoleopgave og kigger efter bøger og tekster skrevet af Voltaire, men også sekundær litteratur om ham og hans værker. Dokumenter om oplysningstiden er også relevante. Du foretrækker materialer på dansk eller engelsk, men materialer på fransk er også interessante.	voltaire	Materialer der handler om Voltaire og hans værker

For hver søgning angives relevansen for et antal potentielle resultater, der således kan bruges til at evaluere hvor gode resultaterne fra en konkret søgemaskine er.

I en tilgang til søgning, der tager afsæt i testsæt, vil fokus blive flyttet til eksempler på søgninger og en faglig vurdering af hvad et godt søgeresultat er. Det betyder også at hvor brugere eller fagfolk finder problematiske søgninger, vil dette direkte kunne gøres til input til at forbedre søgningen, ved at udvide testsættet med disse søgninger og forventede resultater. Denne tilgang vil også understøtte at forskellige behov (eksempelvis skønlitteratur, faglitteratur, børnegrænseflade, særlig mediesøgning, mv.) kan repræsenteres af forskellige testsæt og vil gøre det muligt at implementere et antal søgemaskiner rettet mod forskellige behov, med minimale udviklingsomkostninger.

Tilgangen med testsæt har endvidere den fordel, at heterogene datasæt formentlig bedre kan udnyttes, idet en maskinel optimering ikke vil være begrænset af regelstyrede indekser på metadata.

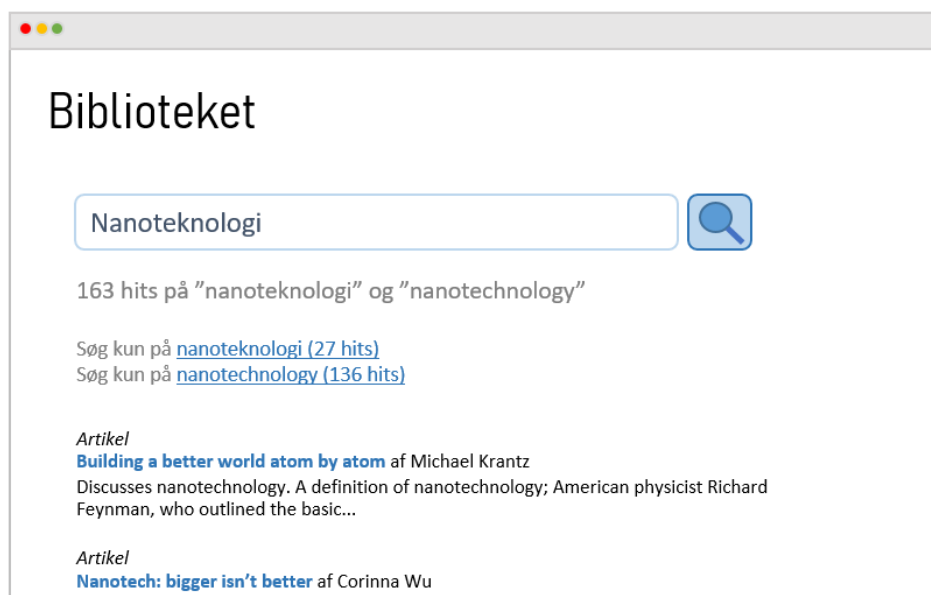
## 2.4 Søgning via naturligt sprog, dialog og stemmestyling

Digitale assistenter der "næsten" kan forstå naturligt sprog eller systemer der er stemmestyret, er allerede nu et hverdagsfænomen på fx smartphones. Chatbots driver interaktioner overalt omkring os, lige fra reklamationer i

webshops, til interne systemer der styrer forretningsprocesser. Slutbrugeren vil allerede om få år være vant til at kunne styre sin mobiltelefon, sit hus, sit fjernsyn og sin bil med stemmen.

Det er et strategisk pejlemærke at kunne understøtte søgning med interaktioner der centrerer omkring naturligt sprog, lagt ind i dialoglignende rammer. Igennem FBI skal udstilles services der kan anvendes til at styre en dialog inden for de hovedforretningsprocesser (anbefaling, søgning, reservation, udlån) der giver mening at understøtte på denne måde.

Det er også et strategisk sigte at søgningen inden for synonymer og oversættelser imellem sprog, bliver forbedret betydelig.



Amanda har brug for litteratur om nanoteknologi til en gymnasieopgave. Hun går til sit biblioteks hjemmeside og søger på 'nanoteknologi' og bliver gjort opmærksom på at hun ser hits på både "nanoteknologi" og "nanotechnology". Hun kan se hvor mange hits der er på hvert søgeord, og hun har mulighed for at vælge, at der kun skal søges på nanoteknologi, som hun oprindeligt skrev. Men hun kan se, at der er mange flere hits på det engelske ord så hun er glad for resultatet.

## 2.5 Integration af brugerdata og bibliotekernes kuratering

Personalisering af services som suggestere og recommendere opleves som uomgængeligt af forbrugere af kommercielle services i dag. Ambitionen for FBI skal sættes højt: En bruger skal ikke opleve på nationalt niveau at få anbefalet et materiale som brugeren har forbrugt hos sit lokale bibliotek, eller i en kommerciel service allerede. For at skabe sammenhæng imellem det nationale og lokale niveau skal der sikres integration af brugerspor (under hensyntagen til brugerens præferencer for privacy, mv) på tværs af sektoren. Ultimativt skal brugeren have mulighed for at brugerens forbrug af indhold i hele FBI sektoren, og udenfor, kan påvirke suggestere og recommendere direkte.

Samtidig skal lokale biblioteker have mulighed for at påvirke suggestere og recommendere og søgeresultater med fx lokalt registreret og kurateret materiale, ligesom de skal have mulighed for at oprette relationer til disse, der efterfølgende kommer til udtryk i suggestere og recommendere. Ultimativt skal et lokalt bibliotek kunne sikre at et lokalt materiale bliver eksponeret overfor "deres" brugere når brugeren eksponeres for relateret materiale. Det skal naturligvis være enkelt for de lokale aktører at sikre at materiale de vægter højt, kan eksponeres i suggestere og recommendere, og dermed eksponere disse materialer, så long tail bliver eksponeret og aktiveret.

Både for at sikre imod "ekko-kammer" effekt og for at kunne eksponere lokalt materiale skal suggestere og recommendere kunne styres fra grænseflader til at dække forskellige kontekster for den enkelte bruger, fx hvis en bruger anvender et bibliotek med en særlig faglitteratur eller fokus, skal vedkommende ikke have de samme anbefalinger, mv., som på det lokale folkebibliotek. Der skal desuden være solid understøttelse af afgrænsninger, således at en recommender fx kan anvendes til kun at anbefale materialer der er umiddelbart tilgængelige i brugerens konkrete kontekst (hvis fx brugeren fysisk befinder sig på en lokal filial), har en bestemt type, osv. i det omfang det er muligt.

### 2.5.1 Eksempel: Enkelhed og åbenhed fremmer innovation

#### Eksempel: Udviklingen af BoFiTrack

Flemming er glad for at bruge biblioteket, men forbruger også mange medier gennem fx Mofibo, Kobo og Netflix. Han er glad for FBI's recommendere, men oplever at han ofte anbefales materiale han har læst eller set allerede.

FBI platformen har et API og en brugergrænseflade der tillader brugerne at indlægge informationer om materialer de har læst eller set og bruge disse data i de personlige anbefalinger til den enkelte bruger. Flemming har været en ivrig bruger af Goodreads igennem mange år og har registreret de bøger han har læst der. Flemming får en ide og spørger en ven, Thomas, der arbejder med software om man ikke automatisk kunne eksportere databasen over Flemmings læste bøger fra Goodreads ind i FBI platformen?

Thomas tager et kig på API'et. Data og API er nemme at forstå og i løbet af et par timer har han lavet et lille script, som Flemming kan køre på sin PC, der overfører informationerne automatisk. Flemming bliver meget begejstret over den forbedrede oplevelse i recommenderne, og fortæller om det i et Goodreads forum for danskere. Flere efterspørger en kopi af scriptet og Thomas lægger det online.

I løbet af en kort periode bliver der bygget forbedringer på scriptet, så man nu kan integrere fra flere kilder. Thomas får forbedringerne tilbage, og beslutter sig for at bruge det som udgangspunkt for en mobilapp. Appen kan også hente informationerne om bøger man har lånt på biblioteket fra FBI platformen, og er således et sted hvor man kan få et samlet overblik over de bøger og film man har læst/set, fra flere kilder.

Thomas kalder appen "BoFiTrack" og den bliver hurtigt populær blandt mange brugere. Et feedback er om man kan bruge appen til at registrere de bøger man læser eller film man ser, når det nu alligevel er i appen man får samlet informationerne? Thomas udbygger appen, så man kan scanne strekkoder bag på de bøger man har læst, og automatisk få titelinformationer, mv. registreret i app'en og overført til FBI platformen. En bruger foreslår om ikke der kan gøres noget, så det bliver nemmere at registrere når man har set en film på fx Netflix. Thomas udbygger appen, så man kan tage et billede af titelinformationerne på en Netflix film, og vha. OCR scanning få registreret filmen. Andre udbygger så funktionaliteten dækker flere streamingtjenester.

Al denne information om de enkelte brugeres læste bøger og set film tilflyder FBI platformen, og kan derfor både bruges i recommendere og i andre systemer der bygges af andre i sektoren.

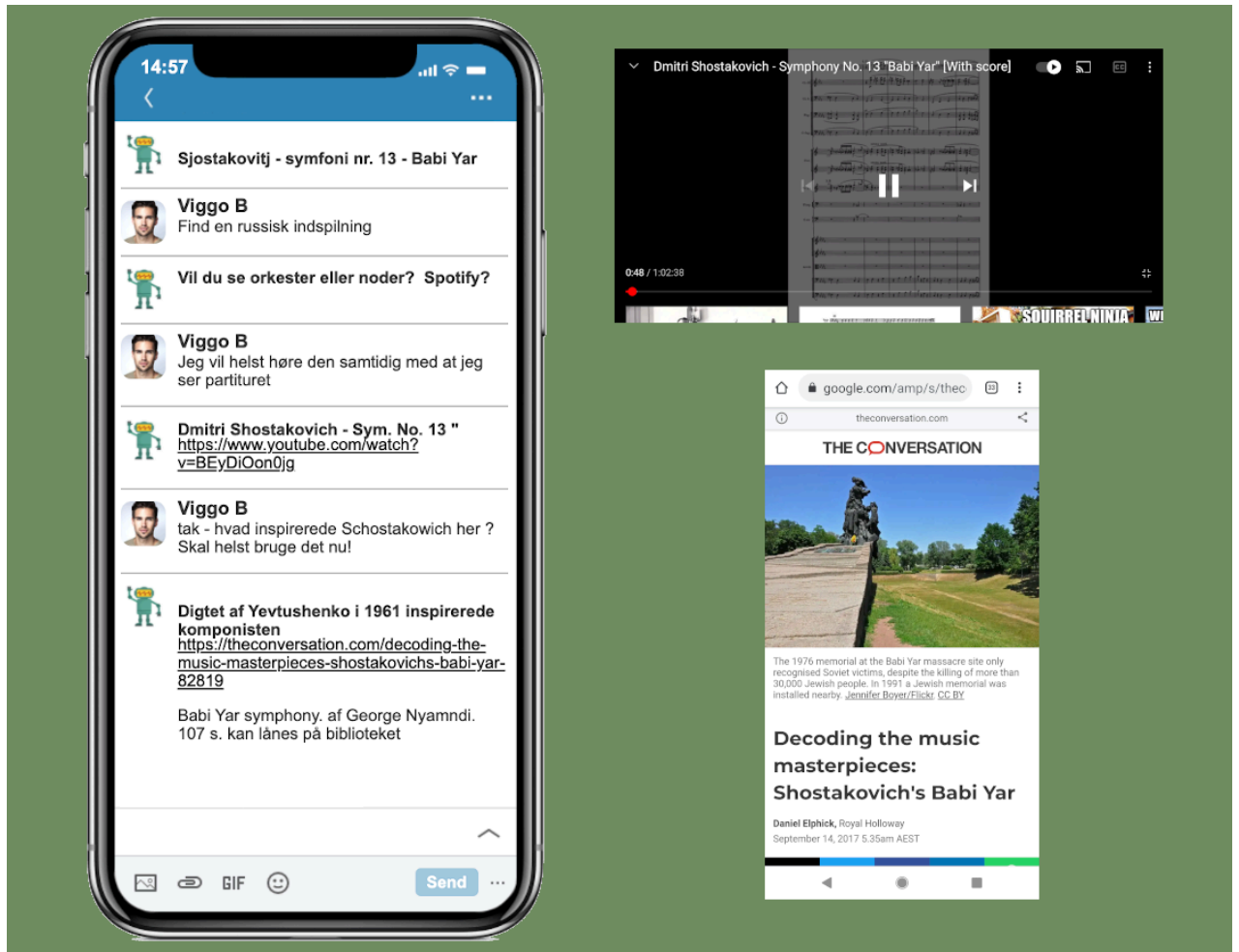
## 2.6 Heterogene data og adgang til resurser uden for sektoren

Sektoren er i dag fragmenteret, når det handler om en række data der ligger uden de metadata, der er repræsenteret i databrønden, der primært reflekterer fysiske og digitale samlinger og kataloger. Dette skal adresseres så sektoren i langt højere grad opfattes som sammenhængende, set fra et brugerperspektiv og muliggør inddragelse af kvalitetskilder, der ikke ligger i brønden eller på anden måde er købt af bibliotekerne, men alligevel vil kunne understøtte brugssituationer der er typiske i bibliotekerne.

Infrastrukturen for søgning skal kunne anvende heterogene data for at understøtte søgningen. Det betyder at det skal være muligt at bruge fx fuldtekster og "ukurante" metadata til søgning uden at data skal konverteres til bibliografiske poster. Dette skal være med til at forbedre søgning, men også understøtte muligheden for at gøre større dele af bibliotekernes samlinger mere tilgængelige, fx artikelsamlinger og andre elektroniske resurser. Det skal også være muligt for lokale aktører at registrere/kuratere arrangementer og websider, andet der i bredere forstand kan opfattes som materiale, og derved understøtte den bredere formidlingsopgave. Dette vil også gøre det nemmere

at integrere søgning på resurser der ikke er traditionelle materialer med bibliografisk information, fx kulturelle arrangementer, statuer, hjemmesider eller musik.

En særlig problemstilling inden for området, handler om rettigheder til adgang til fx fuldtekster. Altså, ikke kun slutbrugerens adgang til at forbruge en fuldtekst, men fx FBI platformens ret til at indeksere en fuldtekst. Tilsvarende for metadata for mange digitale kilder. En sømløs integration vil muligvis kræve politiske initiativer, som naturligvis ligger uden for rammen af denne strategi.



### 3 Projektoverblik

De ovenstående strategiske pejlemærker kan konkretiseres i en række projekter. Projekterne skal ses som et redskab til at komme i gang med at realisere strategien og vi forventer at både antallet og indholdet vil ændre sig. Afhængighederne er vejledende.

	id	Afh. af	Titel	Beskrivelse
Etablering af målarkitektur	a.1	NA	Etablering af målarkitektur	<p>En målarkitektur/systemarkitektur, som fastlægger hvordan FBI arkitekturen, fsva. det brugerrettede område, skal se ud for at understøtte visionen, fastlægges.</p> <p>Arkitekturen fastlægger principper for denne del af platformen, og vil dels fungere som overordnet rettesnor for arkitekturen i det der udvikles, dels fungere som rettesnor ved beslutningsanledninger. Overblik over services, sammenhænge, mv. sikres ved dette arbejde.</p> <p>Arkitekturen vil naturligt kunne påvirke de projekter der foreløbigt er identificeret nedenfor.</p>
Brugerrettet model for data	d.1	NA	Definition af brugerrettet model for data	<p>Fastlæggelse af en første revision af den brugerrettede model for data. Dette skal baseres på eksisterende erfaringer med behov i grænseflader og services (DBC, DDF, udviklere fra sektoren, leverandører, mv.).</p> <ul style="list-style-type: none"> <li>• Analyse af schema.org som udgangspunkt for en model</li> <li>• Afdækning af behov/krav i services, grænseflader, mv.</li> <li>• Fastlæggelse af hoveddataentiteter og deres relationer - logisk model for data</li> <li>• Definition af modeller for værk, serier, mv. der er "first class", unikke og persistent</li> <li>• Dokumentation til interessenter, mv.</li> </ul>
	d.2	(d.1)	Implementation af brugerrettet model for data	<p>Implementation af modellen for data i en faktisk database/datastore</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Fysisk model for data</li> <li>• Implementation af model for data</li> <li>• Syntetiske data for testformål, evaluering, mv (fx forsimplet engangstræk af alle data)</li> </ul>
	d.3	(d.2)	Løbende synkronisering af bibliografiske data til brugerrettet model for data	<p>De eksisterende bibliografiske data, som bevares, skal synkroniseres</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur til synkronisering</li> <li>• Definition af model for konvertering, herunder værk og serie begreber, mv.</li> <li>• Implementation af konverteringer og synkroniseringer</li> </ul>
	d.4	d.3	Løbende synkronisering af holdingsdata, mv. til brugerrettet model for data	<p>Holdingsinformationer, mv. gøres tilgængelige i den brugerrettede model for data</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur til synkronisering</li> <li>• Implementation af synkronisering.</li> </ul>
	d.5	d.2/d.3	Udstilling af data i dominerende søgeplatforme	<p>Data fra den brugerrettede datamodel udstilles i formater som er kompatible med de dominerende søgeplatforme (Google)</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur til udstilling af data via Google (schema.org), evt. andre.</li> <li>• Evt. mindre konverteringer/tilpasninger til de datamodeller der anvendes i de dominerende søgeplatforme.</li> </ul>

	id	Afh. af	Titel	Beskrivelse
	d.6	d.2	Søgesystem baseret på den brugerrettede model for data	<p>Implementation af et søgesystem (API, service, mv) baseret på de brugerrettede data. (Her skal søgning forstås som svarende til SimpleSearch/OpenSearch, altså selve søgesystemet, ikke de omkringliggende services.)</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Migrering/baseret på erfaringer fra SimpleSearch/Enboks-værksøgning</li> </ul> <p>I praksis er dette en justering af datagrundlaget til værkvisningsservicen udarbejdet i FBI-46 og værksøgeservicen i FBI-44.</p>
	d.7	d.2	Diverse services forenkles	<p>Uspecificerede projekter inden for holdings, update, bestilling, suggerer, recommender, mv, der tilrettes/ændres/genopfindes med erfaringerne fra den brugerrettede datamodel, samt tilrettet REST/JSON, en global/fint granuleret autentificeringsmekanisme (forventet baseret på DÅP), mv.</p> <p>Formål inkluderer, som beskrevet ovenfor,</p> <ul style="list-style-type: none"> <li>• At gøre services mere tilgængelige for øvrige aktører</li> <li>• At gøre services kompositionerbare i andre systemer</li> <li>• At reducere drift og udviklingsomkostninger</li> </ul> <p>De endelige projekter vil i vid udstrækning skulle defineres som en konsekvens af målarkitekturen.</p> <p>Projekterne vil i vid udstrækning bygge videre på projekter som FBI-47, FBI-48, mv.</p>
Kvalitet og evaluering	k.1	d.6	Søgeevalueringsystem	<p>Der skal etableres et system til, ud fra en fælles model, objektivt at vurdere kvaliteten af søgeresultater.</p> <ul style="list-style-type: none"> <li>• Fastlæggelse af interessenter (interne og eksterne) og styregruppe for den løbende evaluering af søgekvaliteten</li> <li>• Inddragelse af DDF, biblioteker, udviklere fra sektoren, leverandører, mv</li> <li>• Udvælgelse af korpus</li> <li>• Definition/udvælgelse af kuraterede søgninger og deres optimale resultater</li> <li>• Fastlæggelse af scoremodel</li> <li>• Teknisk platform til at udregne og udstille scoren/afrapportering</li> <li>• Duplikere evalueringen over alle relevante søgemaskiner og testsæts</li> </ul>
	k.2	k.1	Tuning af søgemaskine	<p>Tuning af søgemaskine med udgangspunkt i søgeevalueringsystem, med inddragelse af data om brugeradfærd.</p> <ul style="list-style-type: none"> <li>• Tuning af søgning ud fra evalueringssystemet, fx learning to rank</li> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Implementation</li> <li>• Evaluering</li> </ul>

	id	Afh. af	Titel	Beskrivelse
Integration af brugerdata og bibliotekernes kuratering	i.1	d.7	Integration af kuraterede data fra lokale biblioteker i recommender	<p>Der skal bygges et system hvor lokalt registrerede data kan integreres med recommender i den brugerrettede datamodel.</p> <p>Eksempel: A, B og C burde flere læse. A ligner B. A er beskrevet i artikel B. Tekst til at understøtte relation, fx "Moby Dick baseret på Havets hjerte". Det vil resultere i at en visning af Moby Dick vil føre til en anbefaling af "Havets hjerte".</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Integration i recommender</li> <li>• Service der kan registrere disse data relateret til den brugerrettede datamodel</li> <li>• Grænseflade til brug for lokale aktører</li> </ul>
	i.2	d.7	Systemer til integration af brugerspor/data	<p>En platform for opsamling og udlevering af brugerspor og BI data på tværs af alle FBI systemer, hvortil eksterne data nemt kan integreres. (Naturligvis med respekt for brugernes ønsker i relation til fx privacy)</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Datamodeller, mv.</li> <li>• API og services</li> <li>• Grænseflade til brug for brugere</li> </ul>
	i.3	i.2	Recommendere til slutbrugerpersonalisering	<p>Recommendere der kan integrere data fra i.1 og i.2, så bibliotekernes kuratering afspejles og brugerne fx ikke oplever at blive anbefalet materialer de allerede har forbrugt, fravalgt, eller lignende.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Integration i recommender</li> </ul>
Heterogene data og adgang til resurser uden for sektoren	h.1		Integration i søgning - arrangementer og hjemmesider	<p>Der skal bygges et system hvor informationer om arrangementer eller relevante hjemmesider kan integreres i søgesystem.</p> <p>Eksempel: Et lokalt bibliotek laver et arrangement om H.C. Andersen. Dette vil de gerne have komme op i den nationale infrastruktur, hvis en bruger i deres opland, fremsøger H.C. Andersen eller emner relateret til dette. Et andet bibliotek ved at der findes en hjemmeside der indeholder information om H.C. Andersens gøren og laden i den lokale kommune, og ønsker at denne fremkommer ved en søgning.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Service der kan registrere disse data relateret til søgesystemerne</li> <li>• Grænseflade til brug for kuratering, fx på lokalt bibliotek eller centralt udvalg af relevante hjemmesider.</li> </ul>
	h.2		Integration i søgning - fuldtekster	<p>Der skal bygges et system hvor fuldtekster kan bruges til at forbedre søgningen</p> <p>fx at Infomedia artikler bruges til at producere data der kan indgå i søgemaskinen i forhold til at finde artiklerne</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Udvidelse af søgemaskine til at kunne understøtte fuldtekstafledte metadata eller embeddings</li> </ul>

	id	Afh. af	Titel	Beskrivelse
	h.3		Forsider	<p>Der skal bygges et system med understøttelse af nem integration og deling af forsider. Uanset om det er DBC, lokale biblioteker eller andre der fremskaffer forsider, skal disse være teknisk mulige at bruge for alle. Derudover bør der undersøges muligheder for at lokalisere populære materialer og fremskaffe forsider, såvel som muligheder for lave forsider for alle materialer der vises.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Service der kan registrere disse data relateret til søgesystemerne</li> <li>• Grænseflade til brug for lokale aktører</li> <li>• Afsøgning af at lave syntetiske forsider</li> <li>• Rettighedsclearing i forhold til anvendelse af forskellige typer forsider</li> </ul>
	h.4		Integration i søgning - rettigheder til at katalogisere og indekserer fuldtekster	<p>FBI skal have rettigheder til at bruge fuldtekster fra fx e-resurser til indeksering i søgemaskiner. Dette er en politisk proces, der med fordel kunne drives igennem KOMBIT.</p>
	h.5		Relationer til eksterne kilder.	<p>Der skal bygges infrastruktur til at understøtte relationer fra biblioteksmaterialer til eksterne kilder.</p> <p>Det kan være relationer fra et konkret materialer til VIAF, Wikidata, etc. Hvor målet er at kunne vise andre relevante resurser relateret til et konkret materiale.</p> <p>Eksempel: visning af et konkret materialer på en biblioteksgrænseflade giver adgang til eksterne kilder er relateret til det konkrete værk eller forfatteren.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Service der kan relatere elementer i den brugerrettede datamodel til eksterne kilder</li> </ul>
Søgning via naturligt sprog, dialog og stemmestyring	n.1	d.6	Natursprogssøgning	<p>Søgning (snæver forstand) hvor en sætnings semantik i forhold til søgning bruges til at forbedre søgningen.</p> <p>Eksempler på relevante cases (søgninger):</p> <ul style="list-style-type: none"> <li>• "Hvordan isolerer man gulvet i et træskur"</li> <li>• "Flere bøger af ham der har skrevet ringenes herre"</li> <li>• "Camilla Lækbergs nyeste"</li> <li>• Problemformulering fra gymnasieopgave indsættes i søgefelt</li> <li>• Tekst fra nyhedsartikel kopieres ind i søgefelt</li> </ul> <p>ML teknikker kan bruges til at lave semantiske repræsentationer af fx søgninger og også metadata/flaptekster fra søgbare materialer. Dette giver muligheder for at lave semantiske søgninger (hvor der ikke er tale om direkte match af søgeord).</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Service der kan konstruere semantisk repræsentation</li> <li>• Integration i søgemaskine infrastruktur</li> </ul>



	id	Afh. af	Titel	Beskrivelse
	n.2	d.6	Dialogbaseret søgning	<p>Dialogbaseret søgning i bred forstand, hvor brugeren ved interaktion kan tilpasse søgningen. Ønsket er at grænsefladen stiller flere muligheder til rådighed for brugeren for at brede søgningen ud eller indsnævre denne, og disse muligheder tilpasses løbende på baggrund af brugerens interaktion.</p> <p>Eksempler:</p> <ul style="list-style-type: none"> <li>• Søgning kan initeres vha. "smarte facetter", hvor man fx inden indtastning trykker på <i>krimi</i> og trækker udgivelsesliden frem til i år, og derved har lavet en søgning på nye krimier.</li> <li>• Emnesøgninger giver mulighed for at se relaterede emner som enten breder ud eller forfiner vha fx et emnehierarki</li> <li>• Søgning på "turen går til Rom" resulterer (ud over hits) i mulighed for at, vha. link, at foretage søgningen "rejseførere rom"</li> </ul> <p>Opgaver inkluderer:</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Integration i søgemaskine infrastruktur</li> </ul>
	n.3	d.6	Søgning gennem chatbot	<p>Understøttelse af hele processen fra søgning med dialog-elementer og til bestil i en chatbot.</p> <p>Eksempel:</p> <ul style="list-style-type: none"> <li>• Krimi-eksempel i visionen</li> </ul> <p>Stemmegenkendelse ligger uden for dette projekt, men kan implementeres i grænseflader, mv.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Integration i søgemaskine infrastruktur</li> </ul>
	n.4	d.6	Søgning med oversættelse af søgning	<p>Understøtte muligheden for at søge på dansk og også få resultater med engelsk metadata, fx elektroniske resurser. Skal tage højde for ranking, således at fx brugerens søgning bliver rankeret højere i resultatet end den oversatte søgning.</p> <p>Eksempel:</p> <ul style="list-style-type: none"> <li>• Brugeren søger på nanoteknologi, hvor der er meget lidt på dansk, men får også engelske materialer der handler om nano technology</li> </ul> <p>Projektet skal trække på eksisterende arbejde og erfaringer, fx fra Ballerup bibliotek.</p> <ul style="list-style-type: none"> <li>• Teknisk infrastruktur</li> <li>• Eksperimenter, R&amp;D</li> <li>• Integration i søgemaskine infrastruktur</li> </ul>

	id	Afh. af	Titel	Beskrivelse
	n.5	d.6	Søgning med synonymer	<p>Bedre understøttelse af muligheden for at søge på specifikt søgeord og også få resultater med andre synonyme søgeord .</p> <p>Skal tage højde for ranking, således at fx brugerens søgning bliver rankeret højere i resultatet end søgning på synonyme søgeord.</p> <p>Eksempel:</p> <ul style="list-style-type: none"><li>• Bruger der søger på "skraldemand" får også resultater for "renovationsmedarbejder".</li></ul> <p>Opgaver inkluderer:</p> <ul style="list-style-type: none"><li>• Teknisk infrastruktur</li><li>• Skabe synonympar</li><li>• Integration i søgemaskine infrastruktur</li></ul>

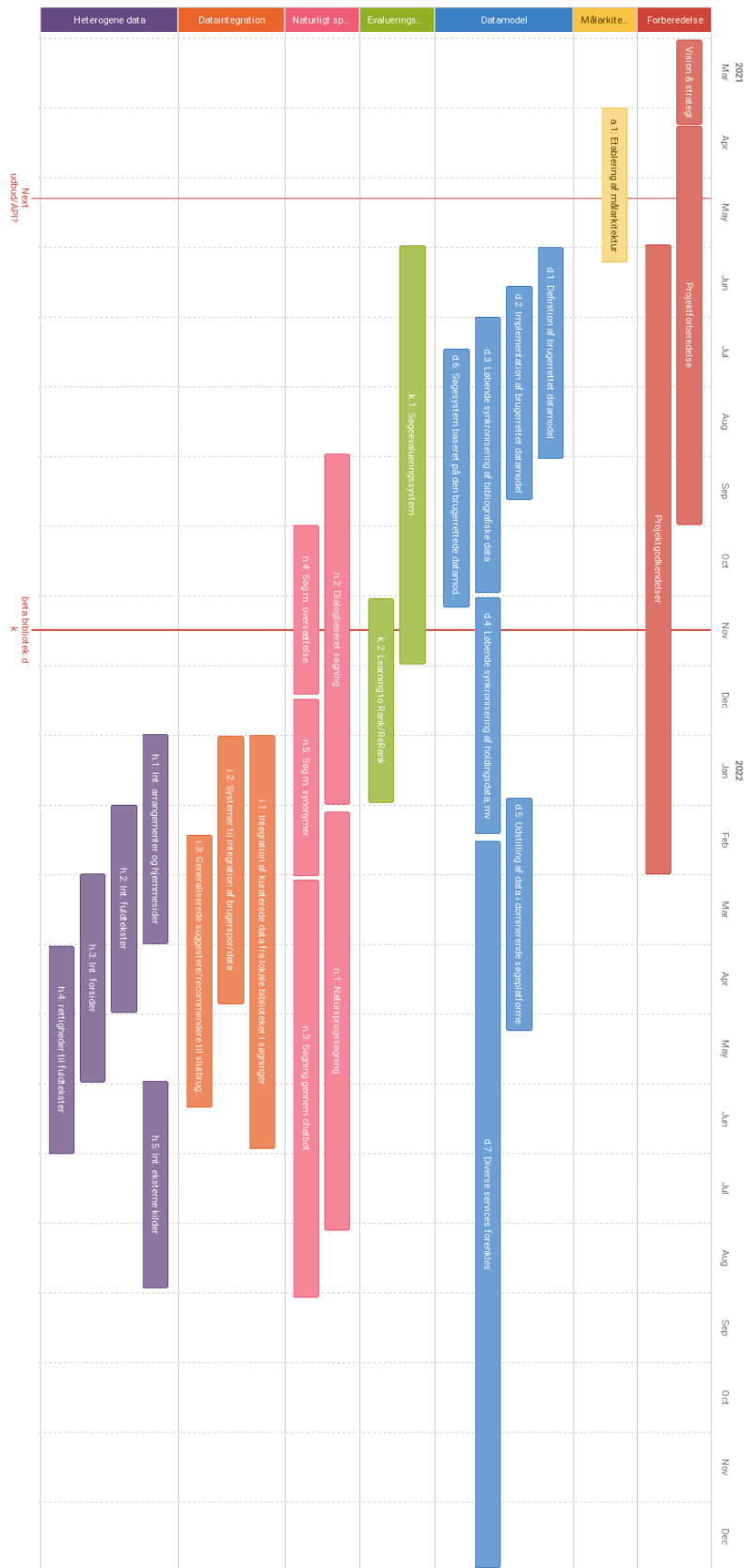
## 4 Roadmap

Målarkitekturen er med til at sikre rammerne for en succesfuld implementering af strategien, og bør derfor igangsættes meget tidligt. Samtidig er en ny brugerrettet model for data det bærende element i strategien – fundamentet for visionens realisering. Den må derfor nødvendigvis komme meget tidligt i forløbet.

Roadmap skal i øvrigt løbende tilpasses så indsatserne understøtter de større projekter der i øvrigt arbejdes med i FBI og i sektoren, herunder især nyt bibliotek.dk og DDF's behov.

En tilgang kan være at udtrække delmængder af flere projekter fra projektkataloget til "indsatsprojekter", fx et behov for serievisninger, el. lignende, der udmøntes i et fokuseret projekt på dette.

Bemærk at usikkerhederne ved roadmappet er betydelige, og det skal kun opfattes som skønsmæssigt vejledende.



## 5 Udvikling af eksisterende søgeinfrastruktur

En optimal søgeplatform for den professionelle og avancerede bruger i FBI skal fortsat understøttes. Grænsen mellem den "almindelige bruger" og den mere "avancerede bruger" er ikke hårfin. Men forenklingerne i API vil også gøre det muligt for en leverandør af en grænseflade at anvende både elementer fra den meget stærke forenkling og de services og data, som den eksisterende søgeplatform – API, brøndsystem, søgeplatform – leverer. Sidstnævnte skal derfor løbende skal udbygges og forbedres, ligesom den trinvist vil skulle integreres med resultaterne af den nye søgevision. Dette ses som en FBI-vedligeholdelsesopgave og en forudsætning for at udbygge konkurrencedygtigheden for FBI.

Platformen skal

- Være den foretrukne søgeplatform rettet imod biblioteksindhold for professionelle og avancerede brugere.
- Understøtte brugerne i deres arbejdsgange således at søgeinfrastrukturen er målrettet de processer, der understøtter materialevalg, accession, udlån, katalogisering, litteraturlister, mv.
- Understøtte brugernes formidling overfor lånerne og bibliotekets øvrige brugere.
- Understøtte fjernlån og de processer i det enkelte bibliotek og mellem biblioteker, der er en del af fjernlånsarbejdet.
- Give mulighed for gennemskue hvordan data, indekser og visningsformater hænger sammen.
- Løbende rammer de behov for indekser, visninger og dataudvekslinger som opstår i den professionelle brug.
- Overholde standarder på biblioteksområdet.

Målgruppen, de professionelle brugere, skal

- Tilbydes et søgesprog der giver dem mulighed for at udforme komplekse kombinatoriske søgninger.
- Tilbydes visningsformater der afspejler deres behov.
- Inddrages med henblik på løbende forbedringer af søgning.

Da den nye søgeplatform forventes at trække hovedparten af søgninger i lokale og nationale borgergrænseflader vil det samtidig gøre presset på den eksisterende søgeinfrastruktur mindre. Der vil dog fortsat være fokus på:

- Svartidsforbedringer og anden form for performanceoptimering.
- Sikkerhed
- Tilgængelighed

I takt med at den nye søgeinfrastruktur realiseres vil det være en væsentlig vedligeholdelsesopgave i FBI at konsolidere og migrere ud fra både økonomiske som funktionelle mål.