

Felix Grimsrud

Tilrettelegging for forskning på data ved et nettbasert forum

Bacheloroppgave i IT med spesialisering i informasjonsbehandling
Veileder: Tore Mallaug
Juni 2020

Felix Grimsrud

Tilrettelegging for forskning på data ved et nettbasert forum

Bacheloroppgave i IT med spesialisering i informasjonsbehandling
Veileder: Tore Mallaug
Juni 2020

Norges teknisk-naturvitenskapelige universitet

Forord

Studieprogrammet heter «Bachelor i informatikk med spesialisering i informasjonsbehandling», og tar for seg en hel masse forskjellige fagfelt. Deriblant objekt-orientert programmering, webutvikling, økonomi, prosjektarbeid, databaser, bruk av hjelpemidler og rammeverk m.m.

Dette prosjektet er påbegynt med den hensikt å utforske hvilket datagrunnlag som finnes i et forum, om disse data kan egne seg til forskning, og om det vil være hensiktsmessig å opprette en ny database som omstrukturerer og tilpasser disse data til forskning.

Strukturen på denne rapporten er slik at man i starten får en presentasjon av prosjektets interesser og krav, samt bakgrunn for prosjektet. Deretter kommer det en evaluering av dagens systemer, og den underliggende databasen. Etter systemet er presentert og databasen utforsket, blir det drøftet grunner til å opprette en ny database, og om datagrunnlaget er av tilstrekkelig kvalitet til forskning. Det dras her opp ulike alternativer til databaseløsninger, samt noen fordeler og ulemper ved dem. Deretter begrunnes det hvorfor den endelige løsningen er den som er blitt valgt. Herfra presenteres så verktøy og teknologier som er tatt i bruk for å implementere løsningen, før selve implementasjonen presenteres.

Etter implementasjonen er presentert, er det dratt frem arbeidsområder for videre arbeid som har presentert seg gjennom prosessen. Før kildene presenteres i siste kapittel, er det lagt ved et refleksjonsnotat om hvordan gjennomføringen av prosjektet er gått, som også inneholder en forenklet timeoversikt.

Takksigelser

Jeg vil takke oppdragsgiver Karl Johan for å være støttende i mitt arbeid, og bistå ressurser som det har vært behov for, etter hvert som behovet har oppstått.

Takk veileder Tore Mallaug for å ha vist et godt innsyn i konstruksjonen til en bachelor-oppgave, og prosessen bak. Han har også vært en god pådriver, og gitt nyttig konstruktiv kritikk, spesielt da med tanke på skriving av oppgaven og dens deler.

Takk til Thorleif Hjeltnes for å ha delt ressurser som har vært nyttige for å finne ut av GDPR reglement og standarder, og god praksis på den fronten.

Sist vil jeg også takke Stig Ålgårdstad for å ha delt sin kunnskap om og sine erfaringer med verktøyet WordPress (WP), og for å ha pekt i en god retning når jeg ikke helt har visst hvor til å begynne.

Det har vært en fryd å få jobbe med dere alle sammen, og jeg er så glad for at jeg har hatt tilgang til slike ressurssterke mennesker gjennom prosessen.

Innhold

Forord.....	1
Takksigelser.....	1
1. Introduksjon.....	4
1.1 Om prosjektet og Mental Helse.....	4
1.2 Interessenter.....	4
1.2.1 Mental Helse – brukere & administrator.....	4
1.2.2 TISIP.....	4
1.2.3 NTNU.....	5
1.2.4 Brukerrepresentanter.....	5
1.3 Oppgave.....	6
2. Brukerkrav.....	7
3. Evaluering av eksisterende løsning.....	8
3.1 Evaluering av databasen.....	8
3.2 Evaluering av personvern i henhold til GDPR.....	11
4. Design.....	12
4.1 NoSQL.....	12
4.2 RDBMS – Relasjonelle databaser.....	13
4.3 Plugins (utvidelser).....	13
4.4 Selvlaget plugin-design.....	14
4.5 Design-beslutning.....	14
4.6 Den endelige løsningen.....	15
5. Teknologi og Verktøy.....	17
5.1 WordPress.....	17
5.2 MySQL og PHP-myadmin.....	17
5.3 WampServer.....	17
5.4 FileZilla.....	17
5.5 jQuery.....	18
5.6 AJAX.....	18
6. Utvikling.....	19
6.1 Sette opp lokalt testmiljø.....	19
6.2 Testing av «Hello Dolly».....	19
6.3 Eksport av sql-spørring gjennom php til .csv fil.....	19

6.4 Forhåndsvisning med AJAX og jQuery	20
7. Videre arbeid.....	22
8. Evaluering av Gjennomføring.....	24
8.1 Timebruk (tabell og kanskje graf/e.l.	24
8.2 Kontakt med ulike personer.....	25
8.3 Refleksjonsnotat	26
8.3.1 Oppgavevalg.....	26
8.3.2 Manglende erfaring med NoSQL.....	26
8.3.3 Undersøke arbeid med GDPR	27
8.3.4 DIFI – retningslinjer	27
8.3.5 Godaddy.....	28
9. Kilder	29
Vedlegg	32
Oppgavetekst	32
Forstudierapport.....	34
Kravspesifikasjon.....	40

1. Introduksjon

1.1 Om prosjektet og Mental Helse

Bacheloroppgaven er en del av et større prosjekt. Hos arrangørene av dette andre prosjektet har det oppstått et behov for å anskaffe og fremstille informasjon om hvordan det går med prosjektet og dets resultat. Dette er det som ligger til grunn for bacheloroppgaven.

Det større prosjektet er et nettsted som utvikles for brukerrepresentanter, med den hensikt å gi dem et forum å diskutere fagstoff og fremgangsmåter, erfaringer og mer. Andre borgere som er interesserte i mental helse, være dette vanlige folk, pårørende, eller folk som sliter med den mentale helsen skal også være velkomne til å delta i forumet. Dette er nært tilknyttet hva Mental Helse som organisasjon jobber for. De ønsker å utbedre den allmenne holdningen til mental helse og problemene det kan medføre seg.

Dersom en besøker Mental Helse sin hjemmeside, kan man få hjelp i noen forskjellige former, og man kan også melde seg opp som medlem. Det som ikke tilbys der er noe forum, eller grensesnitt for å diskutere erfaringer og fagstoff med andre interesserte, pårørende og brukerrepresentanter. Det er nok delvis dette som ligger til grunn for at forumprosjektet er iverksatt.

På stiftelsen Dam sine sider kan man også finne en del informasjon om prosjektet. Her er også rekruttering og kunnskapsdeling dratt frem som viktige momenter¹. De har en veldig oversiktlig presentasjon av prosjektet som helhet.

I fjor var det en bachelor for den samme studielinja som jeg går på, som gikk ut på å sette opp nettstedet. Den dokumenterer deler av systemet og hvordan det er satt opp. Noen mindre endringer av detaljer som er gjort i etterkant av bachelorens slutt, mangler det dokumentasjon for, dette sannsynligvis fordi det er tatt en sommerjobb i etterkant, og at andre har kommet til og hjulpet prosjektet, uten nødvendigvis å dokumentere arbeidet.

1.2 Interessenter

Mental Helse er oppdragsgiver for bacheloren. I det overordnede prosjektet er det et samarbeid mellom medlemsorganisasjonen Mental Helse, fagstiftelsen TISIP og universitetet NTNU. I bacheloroppgaven er det administratoren som er mest relevant. Det er denne administrerende brukeren som skal ha tilgang til det bacheloroppgaven munner ut i.

1.2.1 Mental Helse – brukere & administrator

Mental Helse er en ideel, landsdekkende medlemsorganisasjon som ønsker å spre sosial aksept rundt det å slite mentalt. Som skrevet på deres forside²:

Mental Helse er en medlemsorganisasjon for alle mennesker med psykiske helseproblemer, pårørende og andre interesserte. Vi mener det skal være lov å ha det vanskelig og at mental helse skal være like lett å prate om som fysisk helse.

1.2.2 TISIP

TISIP er en FoU-stiftelse som er knyttet hovedsakelig til byen Trondheim. Deres spesialitet er forskning på IKT-anvendelser. TISIP samarbeider en del med universitetet NTNU, og en del av deres kurs i IT kan

tas opp som fag gjennom universitetet. Om arbeidsområder er det skrevet disse punktene på infosiden deres³:

- FoU knyttet til IKT og e-læring
- Utvikling av nettbasert læremateriell
- Fagskoleutdanning innen IT
- Utgivelse av IKT-bøker i samarbeid med norske og utenlandske forlag
- Nettkurs innen IT som gir studiepoeng
- Portal og studiesenter for nettbaserte utdanningstilbud

1.2.3 NTNU

Forkortelsen NTNU står for Norges Teknisk- Naturvitenskapelige universitet. Det er gjennom NTNU at jeg som elev har fått oppgaven. Alle tre organisasjonene samarbeider om det totale prosjektet.

1.2.4 Brukerrepresentanter

De som skal bruke nettstedet er brukerrepresentanter eller andre folk som er interessert i arbeid med mental helse. Gjerne medlemmer av organisasjonen Mental Helse blir brukere av nettstedet. Også andre foreninger som jobber med psykologiske utfordringer vil være brukere av nettstedet.

1.3 Oppgave (som beskrevet av studenten, se vedlegg for original Oppgavetekst)

Oppgavetittel:

Tilrettelegging for forskning på data ved et nettbasert forum

Hensikt:

Det er i våren 2019 opprettet et nettforum for brukerrepresentanter i regi av den ideelle stiftelsen Mental Helse, hvor brukerrepresentanter og andre som er interesserte i mental helse kan dele sine erfaringer og meninger rundt diverse temaer som har tilknytning til dette. Forumet er opprettet av Mental Helse med støtte fra NTNU og stiftelsen TISIP, to organisasjoner som driver med kurs, opplæring og forskning innen blant annet IT. Oppgaven blir et delprosjekt i det større prosjektet som da blir å sette opp og blåse liv i forumet. Oppgavens prosjekt handler om å tilrettelegge for analyse av og forskning på innhold og data fra forumet. Forumet er satt opp i WordPress.org. Det er i tillegg til ønske om forskning, uttrykt et ønske om funksjonalitet som lar en lagre diverse dokumenter i ulike format på nettstedet. Dersom det er mulig er det også ønskelig om en får satt tilleggsfunksjoner som savnes eller mangler.

Det er verdt å merke at de som har satt opp nettstedet ikke jobber med prosjektet lengre, ettersom de selv var to bachelorstudenter som fikk i oppgave å sette opp nettstedet. Dermed er det to verdifulle ressurser som ville visst mye om nettstedet som ikke lenger er tilgjengelig.

2. Brukerkrav

Administratoren av nettstedet ønsker å få tak i informasjon om brukeraktivitet, brukerlokalisering samt innhold i foruminnlegg for bruk i kvalitativ analyse. Tilgangen til disse data må være begrenset til administrerende brukere. Aktuell informasjon, slik foreslått av administrator er: Innleggsinnhold, bruker(id), tidspunkt, tittel, Tema, bakgrunnsdata (er skribent brukerrepresentant, region, hvor da?)

Dataene må være nedlastbare. Akseptable fil-formater er .csv, .txt, .doc, .xls, .pdf.

Det totale systemet må være utformet i henhold til GDPR-reglementet.

Se use-case i Kravspesifikasjon

3. Evaluering av eksisterende løsning

Ettersom det er snakk om en forbedring av et allerede eksisterende system, er det aktuelt å gjennomgå deler av funksjonaliteten som allerede finnes, og evaluere om noe bør endres, og for å kartlegge hva som mangler.

3.1 Evaluering av databasen

Databasen lagrer mye nyttig informasjon. Spesielt interessant for Mental Helse er gjerne brukertabellene og innleggstabellene. Disse er interessante fordi det er snakk om å bedrive analyser og forskning på dem. Andre aktuelle data er ifølge oppgaveteksten metadata, trafikkdata og dokumenter. Den lagrer også mye metadata som gjerne ikke er like interessant fra et analyseperspektiv, men som er viktig for strukturen til nettstedet. Et eksempel er gjerne kolonner for å vise om noe er et bilde. Bilder lar seg ikke tekstanalysere. Databasen har 42 tabeller, hvorav nesten halvparten er noe interessante, halvparten uinteressante, og fire-fem stykk er meget interessante.

Trafikkdata lagres i form av sist aktivitet. Dette måles både på forumsider, innlegg, grupper og individuelle brukere. Databasen har også mange tomme tabeller. Disse fylles nok med informasjon etter hvert som forumet tas i bruk. Det kan kanskje med fordel nevnes at det til nå er registrert mindre enn 100 brukere hvorav en del er testbrukere.

Det finnes ingen egen tabell for dokumenter. Generelt er det få egne tabeller for noe som helst. Dette gjør databasen ganske rotete, og det er tungt arbeid å gå gjennom hva hver tabell inneholder. De som har få rader er relativt greie å få oversikt over, men de lange meta-tabellene er skikkelig tungleste. Ikke alle objektene har definert alle meta-verdiene, og da er det ingen rad for denne udefinerte meta-verdien. Hvis en meta da er sjelden, er det lett å gå glipp av den, selv om den kanskje kan tenkes å være nyttig i de tilfellene den er definert. Grunnen til at den har mange tabeller er fordi hver plugin oppretter sine egne tabeller, og tar ikke særlig hensyn til hva som finnes allerede.

Databasen har kanskje en grad av normalisering på andre normalform. For eksempel har de ikke en egen tabell for innleggstype og har dermed mange rader med samme verdi i kolonnen for innleggstype. Dette skaper også veldig mange kolonner i noen av tabellene, som igjen gjør dem uoversiktlige.

Meta-tabellene har generelt færre kolonner som gjør dem noe greiere å jobbe med. Til gjengjeld har gjerne enda flere rader enn de andre. Flere innskudd i meta-tabeller tilhører gjerne samme objekt, og kobles sammen med en objekt-id, som da er fremmednøkkel.

Noen tabeller som tok lang tid å analysere, til tross for at de inneholder rimelig få kolonner, var «bp_xprofile_meta», «bp_xprofile_data» og «bp_xprofile_fields». Disse måtte utforskes fordi de inneholder informasjon om brukerne og deres profiler, i den ene pluginen vel å merke. Noe som spesifikt er interessant er dataene om hvor brukeren holder til, for forskning og kartlegging av hvilke geografiske områder som har mest aktivitet. Tabellen for profildata inneholder all profilinformasjon som samles gjennom utvidelsen BuddyPress, som er brukt for registrering, grupper og meldinger. Her registreres tilholdsfylke (før fylkessammenslåing⁴), hvilken organisasjon man representerer, hvor man representerer, telefon, navn og interesser.

Meta-tabellen viser kun tall og verdier for visningen av knappene som vises på siden når man er inne, noe som er annerledes fra de vanlige WordPress tabellene. Det er i felt tabellen at man finner verdiene

som ligger inne i datatabellen. Strukturen her er stress å forklare så det gjelder å prøve å ha tunga beint i munnen.

Datatabellen har kolonnen «value» som må hentes i felt-tabellen, hvor kolonne med rette verdier har navnet «name». Disse kobles sammen av fremmednøkkelen field_id i datakolonnen. Field_id kolonnen peker til primærnøkkelen i felttabellen der kalt «id». Raden i felttabellen som field_id peker på inneholder ikke verdien som er i «value»-kolonnen i datatabellen, men er heller en fremmednøkkel til en annen rad i samme tabell som inneholder denne verdien. En kan tenke på det som en select-tag i html; select-taggen sin id benevner selecten, mens det er option-taggene som sier noe om mulige verdier. På databasespråk blir det at primærnøkkelen i en rad i en databasetabell er fremmednøkkel til en annen rad i samme tabell. Enkelt sagt betyr det at tabellen inneholder en fremmednøkkel til seg selv, noe som kan være meget forvirrende om en ikke vet hva en ser på.

	id	group_id	parent_id	type	name
<input type="checkbox"/>	1	1	0	textbox	Navn
<input type="checkbox"/>	39	1	0	checkbox	Interesser
<input type="checkbox"/>	155	1	0	checkbox	Interesser
<input type="checkbox"/>	155	1	0	selectbox	Fylke
<input type="checkbox"/>	169	1	0	checkbox	Godkjenning av vilkårene
<input type="checkbox"/>	172	1	39	option	Medikamenter
<input type="checkbox"/>	173	1	39	option	Brukerrettigheter
<input type="checkbox"/>	174	1	39	option	Brukermedvirkning
<input type="checkbox"/>	175	1	39	option	Brukerorganisering
<input type="checkbox"/>	176	1	39	option	Bedringsprosesser
<input type="checkbox"/>	177	1	39	option	Likepersonsarbeid
<input type="checkbox"/>	178	1	39	option	Tjenestetilbudene
<input type="checkbox"/>	179	1	39	option	Tvang og makt
<input type="checkbox"/>	180	1	39	option	Alternativ behandling
<input type="checkbox"/>	182	1	0	selectbox	Hvem man representerer
<input type="checkbox"/>	191	1	0	selectbox	Hvor man representerer
<input type="checkbox"/>	192	1	191	option	Spesialist helsetjenesten
<input type="checkbox"/>	193	1	191	option	NAV
<input type="checkbox"/>	194	1	191	option	Kommunale tjenester
<input type="checkbox"/>	195	1	191	option	Forskningsprosjekt
<input type="checkbox"/>	196	1	191	option	Privat stiftelse

Figur 1, skjermdump fra felttabellen, merk at rad 172-180 alle har parent_id 39. Disse er altså alle «options» i rad nummer 39, som har navnet interesser. Rad nr. 9 er på registreringssiden en masse sjekkbokser, som her har hver sin tilhørende option.

			id	field_id	user_id	value	last_updated
<input type="checkbox"/>			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
<input type="checkbox"/>			4	2	1	Trøndelag	2019-04-04 14:42:12
<input type="checkbox"/>			5	39	1	a:5:{i:0;s:17:"Brukerrettigheter";i:1;s:18:"Bruker...	2019-06-21 07:11:04
<input type="checkbox"/>			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
<input type="checkbox"/>			95	155	24	Trøndelag	2019-05-16 11:54:08
<input type="checkbox"/>			96	52	24	95082226	2019-05-16 11:54:08
<input type="checkbox"/>			97	169	24	a:1:{i:0;s:25:"Jeg godkjenner vilkårene";}	2019-05-16 11:54:08
<input type="checkbox"/>			128	169	1	a:1:{i:0;s:25:"Jeg godkjenner vilkårene";}	2019-06-21 07:11:04
<input type="checkbox"/>			[redacted]	[redacted]	[redacted]	[redacted]	[redacted]
<input type="checkbox"/>			130	155	31	Vestlandet	2019-06-17 17:01:30
<input type="checkbox"/>			131	39	31	a:8:{i:0;s:12:"Medikamenter";i:1;s:17:"Brukerretti...	2019-06-17 17:01:30
<input type="checkbox"/>			132	52	31	92804046	2019-06-17 17:01:30
<input type="checkbox"/>			133	169	31	a:1:{i:0;s:25:"Jeg godkjenner vilkårene";}	2019-06-17 17:01:30
<input type="checkbox"/>			139	181	1	11111	2019-06-18 14:55:42
<input type="checkbox"/>			146	1	34	test14	2019-06-19 12:06:36
<input type="checkbox"/>			147	182	34	Mental helse	2019-06-19 08:24:05
<input type="checkbox"/>			148	155	34	Østlandet	2019-06-19 08:24:05
<input type="checkbox"/>			149	39	34	a:3:{i:0;s:17:"Brukerrettigheter";i:1;s:17:"Bruker...	2019-06-19 08:24:05
<input type="checkbox"/>			150	52	34	98081867	2019-06-19 08:24:05
<input type="checkbox"/>			151	169	34	a:1:{i:0;s:25:"Jeg godkjenner vilkårene";}	2019-06-19 08:24:05
<input type="checkbox"/>			152	1	35	test15	2019-06-19 12:45:21
<input type="checkbox"/>			153	155	35	Østlandet	2019-06-19 12:45:21
<input type="checkbox"/>			154	39	35	a:0:[]	2019-06-19 12:45:21
<input type="checkbox"/>			155	169	35	a:0:[]	2019-06-19 12:45:21
<input type="checkbox"/>			156	52	35	98081867	2019-06-19 12:45:21
<input type="checkbox"/>			157	1	36	Test16	2019-06-26 11:44:10
<input type="checkbox"/>			158	182	36	Mental helse	2019-06-19 12:50:51
<input type="checkbox"/>			159	155	36	Vestlandet	2019-06-19 12:50:51
<input type="checkbox"/>			160	39	36	a:3:{i:0;s:17:"Brukerrettigheter";i:1;s:17:"Bruker...	2019-06-19 12:50:51
<input type="checkbox"/>			161	52	36	44444444	2019-06-19 12:50:51

Figur 2 profildatatabellen. Her kan man sammenligne med figur 1. Man ser at field_id = 182 er sjekkboks for hvem man er representant for. Mental helse kommer lengre nede, men andre alternativer til denne er f.eks. NAV, som man kan finne i figur 1.

Noe annet jeg fant finurlig i databasen var at det var en meta i tabellen «usermeta», som har navn Show-admin-bar-front og har verdi true for alle brukere. Først var det mistanke om at dette var en sikkerhetsrisiko, men det virker å være standard i WordPress. Dersom en ser på template-filene til den forrige bachelorgruppen, kan man se at de ordner dette med et filter-hook på server-siden. Dette ser jo ut til å virke helt fint. Det kan jo nevnes at akkurat dette kunne vært løst i databasen også, men løses i kode utenfor i stedet, da dette er standard for WordPress og det tilrettelegges for denne fremgangsmåten.

3.2 Evaluering av personvern i henhold til GDPR

De nye GDPR-reglene som er trådte i kraft i juli 2018 er viktige å følge. Dette blir ekstremt tydelig ved hendelser som den i Bergen Kommune, der det ble millionbot for brudd på reglementet⁵. Fordi det viste seg at det ikke trengtes en helt ny database besluttet jeg å utforske om nettstedet fulgte EUs GDPR-direktiv. WordPress.org følger GDPR dersom innstillingene til utvidelser er satt til å gjøre det. I tillegg er det et krav at brukerne av nettstedet samtykker til bruken av deres data, samt at de kan gå inn og se alle data som er lagret om dem, og sist at de også da har mulighet til å klage på, slette eller endre data som er lagret om dem.

På nettstedet kan man ikke endre navn, men det meste andre av informasjon kan endres av brukeren selv. Det hadde vært uhensiktsmessig om brukerne kunne endre navn, ettersom sammenhengen til foruminnlegg kunne blitt fortapt. Informasjon som interesser, fylke, telefon, e-post, passord kan alle endres. All informasjonen som lagres om brukeren kan også fremhentes. Det meste av disse finner man på profilsiden som man får tilgang til gjennom profil-menyen. Der er det kun e-post som ikke vises. Denne vises under innstillinger på samme meny.

Det eneste som står ut som relativt aktuelt å endre er vilkårene. Ettersom det er en offentlig aktør, er det gjerne på sin plass å ha en personvernerklæring, heller enn bare vilkår.

4. Design

4.1 NoSQL

NoSQL-databasesystemer har flere fordeler i forhold til tradisjonelle relasjonelle databaser. Dette fordi at NoSQL er lettere å skalere, i mange tilfeller har bedre ytelse, kan lagre flere ulike typer data, lagrer data uavhengig av datatype, og ofte strukturerer dataene for deg. De kan også generelt støtte mer fleksible spørringer, selv om dette vanligvis gjøres på applikasjonslaget. De har også mulighet for null nedetid med «masterless» struktur. Objekter/data kan også lagres slik at de støtter flere datamodelltyper, slik at de lettere tilpasser seg det de skal brukes til.

Populære datamodeller er grafmodellen, dokumentmodellen, vid-kolonnemodellen og nøkkelverdmodellen. Disse har alle egne fordeler og ulemper:

Grafmodellen opererer objektorientert, er meget rask og har svært god skalering. Dette fordi den ikke har noen global indeksering av innholdet. Hvert objekt i databasen har kun lagret pekere til naboobjektet i databasen. Dette gjør naturligvis arbeidet med overhead mindre. De kan også kjøre flere spørringer samtidig, noe som kan være nyttig i noen tilfeller, spesielt på applikasjonslaget. For å kunne ta i bruk alle de beste funksjonene i en grafisk modellert database, er det vanligvis nyttig å kunne programmere på lav-nivå språk slik som C++ og Java⁶. Populære grafdatabaser er Neo4j og Titan.

Dokumentmodellen lagrer innslag i databasen gjerne i form av dokumenter i f.eks. JSON-format. Har man et innslag av en bruker, vil alt om denne brukeren være lagret i samme dokument/objekt. Dokumentmodellen bruker ikke ID-er slik man gjør i relasjonelle databaser, og støtter ofte ikke JOIN funksjonalitet. De krever da ofte at om man vil bruke JOIN funksjonalitet, så må en gjøre det på applikasjonslaget. Det blir også vanskeligere å oppdatere f.eks. et bynavn, ettersom dette er lagret slik vi ville lest det i JSON dokumentet. Dermed må en oppdatere alle dokumenter med sagt bynavn om det skal endres. En fordel er at man ikke trenger å bruke like mange JOIN-spørringer; all data om et objekt ligger i dets dokument, og dermed trenger man kun en spørring for å få all info om dokumentet; det er litt som om relasjonene er lagret i samme tabell⁷. Eksempler på slike databaser er MongoDB og CouchDB.

Nøkkelverdmodellen er den kjappeste databasetypen for spørringer. Det kommer av at informasjonen i databasen lagres som en såkalt «dictionary» eller «hash table», som er en lagringsmetode som brukes til å lagre assosiative matriser. Disse databasene er mer fleksible og de kan i rett tilfelle ha ekstremt store ytelsesfordeler⁸. Minnet de bruker i visse tilfeller kan være mye bedre utnyttet og da kan man få mer ytelse ut av maskinvaren en har, som jo er blitt viktig når datamengdene blir så store. Noen instanser av grafmodeller kjører over en nøkkelverdmodell for lagring i bunn⁹. Gode eksempler på nøkkelverdi-databaser er Cassandra, DynamoDB, Azure Table Store (ATS), Riak og BerkeleyDB.

Vid-kolonnemodellen bruker enda en spesiell type dataarkitektur. De lagrer data i kolonner i motsetning til rader slik de mer vanlige RDBMS-systemene gjør. Kolonnefamilie inneholder vanligvis data som har et tydelig forhold til hverandre. Inni en kolonnefamilie lagres alle underliggende data ytterligere i kolonner. Dvs. selve verdien som en ville hatt i en celle i et vanlig RDBMS, er lagret sammen med kolonnenavnet som den tilhører. Dette betyr omtrent at istedenfor at kolonnen peker på alle verdiene som finnes i kolonnen, peker alle verdiene til samme kolonne. I kolonnefamilien adresse har man gjerne fire underlagte kolonner, en for by, fylke, husnummer og gate. Ovenfor kolonnenfamilien adresse har

man gjerne en superkolonnefamilie kalt selskap, som inneholder diverse kolonner som inneholder informasjon om selskapet. Denne siden har en veldig fin visualisering¹⁰. Typiske vid-kolonnedatabaser er HBase, BigTable og HyperTable.

På grunn av oppgavens natur ble redskapet Riak vurdert, med den hensikt å undersøke om det var hensiktsmessig å sette opp en slik NoSQL database. Dette er en nøkkelverdimodellert database, som tilbys som en tjeneste fra dette nettstedet¹¹. Denne teknologien garanterer lese- og skriveoperasjoner til tross for feil på maskinvaren og nettverksnoder, ved å replisere dataen og distribuere den. Dette betyr at databasen har en nær hundre prosent garantert tilgjengelighet. En annen attraktiv egenskap er at den er enkel å skalere. Dette kommer av at databasen har en nøkkelverdi-struktur. De har funksjonalitet for automatisk sletting av utdatert data. Dette er ganske verdifult med tanke på hvor mye informasjon en vil filtrere vekk i WordPress-databasen. At det er lett å bygge en big-data-applikasjon under databasen er trukket frem som en ekstra verdi. Språk man kan skrive applikasjonen på inkluderer Java, Ruby, Python, C#, Erlang, Node.js og .NET-rammeverket.

4.2 RDBMS – Relasjonelle databaser

Mer vanlige relasjonelle databaser er veldig godt utprøvd og dokumenterte. De er derfor lette å ta i bruk. Hastigheten deres er noe dårligere enn andre designtyper, gjerne spesielt NoSQL, men til gjengjeld får man ofte en veldig fin og lettlest oversikt over tabellene. Språket man kommunisere med databasen på heter SQL. Dette språket tar i bruk mange dagligdagse engelske ord, og har relativt enkel syntaks. Dermed er det et språk som er rimelig lett å lære seg. Dette kan naturligvis være en fordel for dem som skal ta i bruk systemet, dersom de ikke kan så mye om slike ting fra før.

MySQL er globalt kjent for å være en av de sikreste DBMS-systemene. De egner seg svært godt for transaksjonsoperasjoner, ettersom relasjonelle databaser ble utviklet nettopp for å holde styr på ting som penger, transaksjoner og aksjer. Skalerbarheten for høyt spesialiserte systemer kan i noen tilfeller være bedre i MySQL enn andre løsninger. Med diverse programtillegg kan man sette opp MySQL til å ha 24/7 opptid, samt god responstid til tross for store datamengder. Dataintegriteten på slike databaser er også en del bedre enn andre typer. I tillegg er det lett å sette opp MySQL-databaser, og de har svært mangfoldige opensource ressurser som en kan ta i bruk, som kan øke funksjonaliteten, ytelse etc. avhengig av hva det er behov for.

4.3 Plugins (utvidelser)

Plugins er programtillegg som hjelper deg å få lagt inn funksjonalitet på WordPress-siden du administrerer. Disse utvidelsene er skrevet av andre programmerere som har utviklet funksjonalitet som de savnet, og deretter har delt koden sin med WordPress-samfunnet.

Det er blitt undersøkt en del ulike plugins som kan kommunisere med databasen. De fleste av disse krever en viss forståelse av SQL for å kunne tas i bruk effektivt. Det finnes ikke utvidelser som tar i bruk NoSQL-databasestrukturer fordi WordPress kun har støtte for MySQL¹².

Det er blitt tatt i bruk en utvidelse lar brukere laste opp dokumenter allerede. Disse kan så lastes ned igjen av andre brukere, så det eneste som egentlig mangler her er funksjonalitet for å hente ut eventuell

tekst til tekstlig og innholdsmessig analyse. Måten utvidelsen henter dokumentene på er via en sti som lagres i databasen, som peker på hvor dokumentet ligger på tjeneren. Dermed forsvinner litt av grunnen til å utvikle en egen database som kan lagre disse dokumentene. En mulig forbedring en kunne tenke seg å gjøre kunne da være å hente ut all tekst fra slike dokumenter i bulk, slik at administrator slipper å laste ned og analysere hvert dokument hver for seg.

4.4 Selvlaget plugin-design

Et annet alternativ er å lage en utvidelse selv. Denne trengs da ikke publiseres offentlig til WordPress sine sider. Dette gir full frihet mtp utforming av brukergrensesnittet. I tillegg er det da mye lettere å tilrettelegge for den spesifikke sluttbrukeren, altså administrator, sitt behov. Dersom det skulle komme et nytt behov senere, kan man alltid skrive en ny utvidelse, eller forbedre den som finnes. Dette krever at noen skriver utvidelsen, noe som kan være vanskelig og ta tid. Dersom en ikke har ferdighetene til å lage den bra nok vil det være tapt arbeid.

4.5 Design-beslutning

Den endelige løsningen som ble valgt var en selvlaget utvidelse/plugin som baserer seg på informasjonen som allerede finnes i den relasjonelle databasen som allerede er satt i stand ved hjelp av WordPress og de utvidelsene som allerede er tatt i bruk. Dette ble valgt fordi det å tilrettelegge for analyse av dokumenter og andre filer virker å være en større jobb enn det er ressurser til, og utenom dokumentene finnes den data som ønskes å tas i bruk i databasen allerede. Dersom dataene i databasen ønskes å utvides og man vil tillegge mer forskjellige data, finnes det støtte for dette i WordPress-løsningen slik den er, dette må i så fall tilrettelegges for med kode eller via en utvidelse.

Mange av fordelene, nevnt i delkapitlene om NoSQL, MySQL og eksisterende plugins, vil være vanskelig å ta i bruk i vår løsning. En NoSQL-løsning krever nok mer ressurser enn det som er til råde. Ettersom nettstedet er relativt lite fortsatt, og datamengdene overkommelige med andre løsninger, er det ikke er like mange fordeler som en kan dra nytte av ved bruk av en NoSQL-løsning. Derfor er beslutningen at dette ikke vil være verdt kostnaden for øyeblikket. Riak virker som et redskap man skal bruke til større bedrifter som har egne nettverk satt opp, gjerne med intranett etc. noe vårt forum ikke har støtte for enda. I tillegg er det ikke underliggende støtte fra WordPress for NoSQL-databaser, så en måtte i så fall laget en egen applikasjon i et av de støttede Riak-språkene, noe det ikke er ressurser til å gjøre.

MySQL kunne vært tatt i bruk. Dette ville bragt med seg mer arbeid, men også noen fordeler. Blant annet ville man hatt mindre rot i databasen. Det som er det største argumentet mot å ta i bruk en ny MySQL-database er at en uansett må manuelt rydde i den nåværende databasen. I tillegg må man da uansett gjøre endringer på den nye MySQL-databasen dersom en vil legge til nye datatyper, eller også dersom nye utvidelser legger til nye data. Dvs. en vil ha mer arbeid å gjøre dersom en skal sette opp en ny database, fordi en må designe den. Dette slipper en hvis en tar i bruk databasen som finnes. En må uansett utforske databasen som finnes for å finne ut hva en vil lagre i den nye databasen, ergo så skaper en mer arbeid. Det som hadde vært en fordel med en slik database hadde vært at spørringer mot den sannsynligvis ville vært litt raskere, men dette er ikke et problem i dagens system fordi datamengden sannsynligvis ikke vil bli så stor at spørringene vil gå uakseptabelt treigt. Ved utforsking av om en kan lagre dokumenter i MySQL-databaser finner man at det er kjappere å lagre stier til dokumenter på

serveren enn å lagre dem i selve databasen¹³. Det er dette WordPress-utvidelsen for lagring av dokumenter allerede gjør, derfor blir ikke dette en fordel her heller og en vil skape merarbeid igjen. Disse kan uansett hentes ut i bulk via FTP.

Det fantes ikke særlig til forhåndsprogrammerte utvidelser som henter akkurat de data administrator syntes var interessante i vårt tilfelle, uten at en må kunne SQL. For å lette arbeidsoppgaven til administrator er det en stor bonus å kunne tilrettelegge for at vedkomne trenger å lære så lite SQL som mulig.

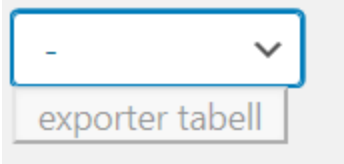
4.6 Den endelige løsningen

Designet som er blitt valgt er tenkt frem med brukervennlighet som viktigste fokus. Administrator ønsker å få fatt i den aktuelle informasjonen fra databasen i form av nedlastbare filer som lett lar seg analysere. For å gjøre dette så enkelt som mulig, er det skrevet spørringer til databasen på forhånd, som en da velger i en nedtrekksmeny. Når en velger en spørring, kommer det opp en forhåndsvisning av de første 30 resultatene av spørringen. Slik kan en da forhåpentligvis tydelig se at tabellen er den man faktisk tenkte å fremskaffe. Om det er feil er det bare å skifte valget i nedtrekksmenyen igjen, så vil det nye valget forhåndsvises. Kolonnene har fått noe endrede navn for å gjøre de mer lettlesbare for administrator og andre brukere av utvidelsen. Forhåndsvisningstabellen har også fått forskjellige toner av farge for å bedre leseligheten. Det hadde nok vært et pluss om en kunne skrive egne SQL-spørringer i utvidelsesgrensesnittet, ettersom en kan forestille seg at noen i fremtiden vil ha interesse av det, men slik situasjonen er i dag er det ikke behov for det.

aktivitets_id	type_aktivitet	innhold	dato
53	bbp_topic_create	Hvordan best tilnærme seg nye medlemmer og invitere de til forumet? Mener den enkleste måten vil være å gi en forklarende setning og en lenke til registreringssiden. Hva tenker dere?	2019-06-11 08:48:17
54	bbp_topic_create	Hvilke er deres beste erfaringer? Ønsker å samle inn en liste over potensielle gode erfaringer for å få en nærmere innsikt i hva som betegnes som gode erfaringer.	2019-04-05 19:52:20
55	bbp_topic_create	I et nytt innlegg på Psykologgen forteller Eline Skår om opplevelsen fra å være innlagt på døgnavdeling til å komme hjem og bo alene. Dette er det nok flere som føler på. Innlegget kan leses her Hvilke erfaringer har dere rundt dette, er det noe vi kunne gjort for å forbedre overgangen fra å være innlagt til å komme hjem?	2019-06-11 09:32:47
56	bbp_topic_create	denne artikkelen på forskning.no, legger de fram resultatet fra et lavterskeltidbud holdt for ungdom mellom 14 og 20 år. Resultatene viser at mestringkurset har hatt effekt mot depresjon blant de som har deltatt. Likevel ser de et behov for mer aktiv rekruttering og nevner at skolen bør benyttes som en større arena til å gjennomføre rekruttering. Tenker dere at skolen bør benyttes i større grad til å rekruttere? Og hvordan kan dette eventuelt bli gjort?	2019-06-11 11:02:03
57	bbp_topic_create	Under filer finner man mapper for dokumentasjon, bilder og webinarer. Hvis det er ønskelig vil vi kunne legge til flere hovedmapper. Kom gjerne med forslag til mapper eller strukturering i filsystemet.	2019-06-11 12:12:47
61	bbp_topic_create	Det kan være utfordrende å være representant i brukerutvalg. Hvem skal man egentlig representere? Skal man representere bare seg selv eller skal man uttale seg på vegne av andre? Dersom man kun legger vekt på egne erfaringer kan dette bli lite representativt. Man bør være i kontakt med andre brukere og vite om deres erfaringer slik at man har et bredere grunnlag å uttale seg ut fra.	2019-06-12 10:19:42

Figur 3 nedtrekksmeny plus export-knapp, samt forhåndsvisning tabellen

Når en spørring er valgt og presentert i brukergrensesnittet, kan man trykke på «eksporter», også vil hele resultatet av spørringen eksporteres til en .csv fil. Denne får da filnavn tilsvarende valget i nedtrekksfeltet pluss .csv-ending. Dersom man ikke velger en spørring, skal eksporter-knappen deaktiveres. Dette gjøres i jQuery, altså på klient-siden. Dette fungerer ikke i live-nettstedet, og kun på test-nettstedet, selv om koden er den samme. Dette kan nok skyldes at TLS-sertifikatet på live-nettstedet er utløpt, som igjen skyldes at det skal foretas en migrering av nettstedet til et annet webhotell over sommeren.



Figur 4, nedtrekksmeny med deaktivert knapp

Det er tatt en bevisst beslutning om å ikke sortere informasjonen eller filtrere resultat-tabellene med GROUP BY eller ORDER BY, fordi eksempelanalysene som er kommet fra administratoren, har tydelige sorteringer og filtreringer. Analyseprogrammet ser ut til å takle pent de brede spørringene som har ganske forskjellig informasjon fra rad til rad. Derfor er det satt fokus på å hente ut informasjonen på enklest mulig måte og med et utbredt og vanlig format, slik at administrerende bruker kan få tilgang til den, og deretter gjøre med den hva vedkomne vil.

```
case 'user-activity':
  $query = "SELECT id, type, content, date_recorded
FROM [redacted]_activity`
WHERE [redacted]_activity.type != 'last_activity'
&& [redacted]_activity.type != 'new_member'
&& [redacted]_activity.type != 'new_avatar'
&& [redacted]_activity.type != 'created_group'
&& [redacted]_activity.type != 'joined_group';";
break;
```

Figur 5, utklipp fra Atom teksteditor som viser den enklere spørringen uten noe strukturering av innholdet. (Se vedleggsmappe om tilgjengelig for full kildekode)

5. Teknologi og Verktøy

5.1 WordPress

WordPress er et verktøy for å lage nettsider. Det gjør det lettere for folk som ikke nødvendigvis kan kode selv, å kunne stelle i stand et webområde. WordPress.org gir flere muligheter enn WordPress.com, men i .org-versjonen må en stille med server selv. I WordPress.com fikser de webtjener for deg, men du får mindre frihet.

En interessant ting ved WordPress er at det ikke finnes en spesifikk html-side for hver enkel side. Sidene er ikke lagret i en egen fil. Dvs. siden du ser på er konstruert veldig dynamisk, vha. tabeller i databasen og templatefiler, som bestemmer hvordan utseende, innhold og funksjonalitet skal være.

WordPress.org bruker noe som heter plugins for å endre på funksjonaliteten. Disse bygger på wordpress sitt underliggende rammeverk, som hovedsakelig består av modifisert PHP-kode, jQuery, css, og AJAX. WordPress har noen egendefinerte globale og superglobale variabler og objekter. \$wpdb er en av dem, og denne styrer kommunikasjon mot databasen. Da slipper man å sette opp en \$conn tilkoblingsvariabel som kobler PHP-scriptet til databasen. Ellers har man også de vanlige globale PHP-variablene, som \$_POST arrayet.

En svært nyttig ressurs har vært den såkalte «plugin handbook» dokumentasjonen. Her får man eksempler på hvordan sette opp en plugin i WordPress, samt «best practices». Den tar en god stund å lese gjennom, og ikke minst å forstå alle kodesnuttene som ligger der inne. Mye av det den tar opp var til stor hjelp i prosessen med å lage en ny spesialisert plugin.

5.2 MySQL og PHP-myadmin

Man får gjennom webhotellet GoDaddy tilgang til php-myadmin, som et grensesnitt til mot databasen. Dette har jeg tatt i bruk for å evaluere databasen, kartlegge hvilken informasjon som finnes, og for å teste meg frem til hvordan jeg skulle skrive spørringene som ville gi resultatene jeg ville bruke i grensesnittet jeg ville lage.

5.3 WampServer

WampServer er et redskap som tillater en å enkelt sette opp en lokal tjener på maskinen sin. Dette ble aktuelt for å kunne teste kode uten at det skulle få konsekvenser for nettstedet og dets tilgjengelighet. På denne måten kan man også få advarsler fra php uten at de trenger å vises for brukerne som vil bruke nettstedet. WampServeren kjører php 7.2.4, som er den seneste versjonen innstillingene tillater. Det faktiske nettstedet kjører på php 7.3, oppdatert i februar. Dette har ikke ført til noen oppdagede feil. WordPress-versjonen er den samme på lokal server så vel som live-server.

5.4 FileZilla

FileZilla er en FTP-tjeneste, som er blitt brukt i kombinasjon med WampServer for å få innholdet fra nettstedet til den lokale testversjonen. Dette redskapet ble også brukt til å laste opp den ferdige utvidelsen som ble produsert i løpet av bacheloroppgaven.

5.5 jQuery

jQuery er et bibliotek som tillater en å skrive mer oversiktlig og tilregnelig Javascript kode. `$()`-funksjonen, med alternativ syntax `jQuery()`, er svært nyttig til navigasjon av HTML sin «Document Object Model». I oppgaven har det vært nyttig å kunne ta i bruk `on.change()` jQuery syntaks, for å kjøre en AJAX funksjon som kan forhåndsviser tabellen en holder på å laste ned før en laster den ned.

5.6 AJAX

Ajax er et redskap for å kunne kommunisere asynkront med serveren. Dvs. at man kan kommunisere med tjeneren etter at en side er lastet inn. Dette gjøres gjerne ved hjelp av javascript eller jQuery hendelser (events). Et eksempel på dette er når google prøver å gjette hva du skal søke på i søkefeltet på google.no. En har nemlig ikke bevisst sendt inn noen ny http-forespørsel fra klienten enda, men det er satt inn et javascript/jquery skript som kjøres hver gang en setter inn et nytt tegn. Denne sender så inn en ny http-forespørsel på vegne av klienten, uten at brukeren egentlig trenger å vite at den har gjort det, og svaret ser du som forslagene de kommer med. Dette skal ideelt sett forbedre brukeropplevelsen.

6. Utvikling

6.1 Sette opp lokalt testmiljø

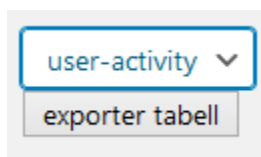
Før utviklingen kunne skje, var det et behov for å sette opp et lokalt testmiljø, for å teste eventuell kode, uten å påvirke serveren. I tillegg slipper man å laste opp filen til en server hver gang man gjør en endring. Dette ble utført ved hjelp av WampServer. Filezilla ble brukt til å hente filer og mapper fra nettstedet, og oppkoblingsinformasjonen til denne ftp-tilkoblingen ble hentet fra webhotellets sider. Det er anbefalt å unngå å laste ned alle utvidelser fra live-nettstedet, fordi det kan være semantikk som ikke er overførbart i alle tilfeller, og da skaper uventede feil. Derfor ble det installert en frisk ny WordPress installasjon lokalt, der en manuelt lastet ned alle de samme utvidelsene, og stilte dem inn så likt som mulig som live-nettstedet.

6.2 Testing av «Hello Dolly»

Hello Dolly er en utvidelse som det refereres til i WordPress-plugin-håndboken. Denne utvidelsen er meget enkel. Koden dens ble brukt til å finne ut hvordan en plugin settes opp i praksis, og kan finnes her¹⁴. Oppsettet av denne pluginen manuelt gav god innsikt i hvordan en skal deklareere plugin-navn i toppen av koden, samt hvordan såkalte hooks fungerer i WordPress. Med utgangspunkt i denne utvidelsens kildekode og plugin-håndboken, ble det satt opp et php-skript, en admin-side, og en meny-knapp for pluginen.

6.3 Eksport av sql-spørring gjennom php til .csv fil

Det aller viktigste i utvidelsen var å få informasjonen fra databasetabellene inn i en fil, så jeg begynte med å programmere denne funksjonaliteten. Grensesnittet er bare en nedtrekksliste med en «eksporter»-knapp. En velger altså en av de tre forhåndsbestemte spørringene via nedtrekkslisten. Når en trykker eksporter sendes spørringen til tjeneren via submit-knappen, og svaret en får settes i en fil via fopen()-funksjonen i php-språket.



Figur 6, som en ser er brukergrensesnittet meget primitivt

```

58 //hook for function that exports the requested data
59 add_action('admin_post_submit-form', 'mh_db_handle_form_action'); // If the user is logged in
60
61 //function to fetch data from database and export it based on select tag in UI
62 function mh_db_handle_form_action(){
63
64     //gets query from function with a switch
65     $query = mh_db_get_query($_POST['fetch-data-select']);
66     global $wpdb;
67
68     if($query === ""){
69         return;
70     }
71     else {
72         $rows=$wpdb->get_results( $query );
73
74         header('Content-Type: text/csv; charset=utf-8');
75         header('Content-Disposition: attachment; filename=' . $_POST['fetch-data-select'] . '.csv');
76         $output = fopen('php://output', 'w');
77         //gets an array of human-friendly table-heads to put in the tableheads of the .csv- file
78         $th_array = mh_db_get_th_array($_POST['fetch-data-select']);
79         fputcsv($output, $th_array);
80         foreach ($rows as $row) {
81             $lead_array = (array)$row;
82             fputcsv($output, $lead_array);
83         }
84     }
85     die();
86 }

```

Figur 7 viser koden som brukes til å hente og eksportere resultatet av spørringen. mh_db_ er den unike prefiksen for funksjoner etc. i utvidelsen, slik at andre utvidelser ikke kaller disse funksjonene. Dette er god standard i WP¹⁵.

6.4 Forhåndsvisning med AJAX og jQuery

For å få til en forhåndsvisning uten å måtte trykke submit og laste inn siden på nytt, så var det en ide å bruke AJAX sammen med jQuery for å få til dette, ettersom det allerede er støtte for disse redskapene gjennom WordPress. Det var hensiktsmessig å ta i bruk jQuery event-funksjonaliteten on.change() til å lytte etter endringer av valg i nedtrekkslisten. Da valget endres, kjøres da en jQuery funksjon som først sjekker om valget er tomt (det finnes et tomt valg) som i så tilfelle returnerer ingenting og deaktiverer eksporter-knappen. Hvis valget ikke er tomt, kjører funksjonen videre og sender en AJAX http-forespørsel, som da ber tjeneren kjøre en funksjon som igjen henter de første 30 radene av resultatet av en sql-spørring tilsvarende den til eksport, og overfører dem til JSON-format, før de returneres til jQuery-funksjonen som sendte forespørselen, som videre oppretter en html-tabell som viser de 30 resulterende radene. Se figur 3 for et eksempel på hvordan dette ser ut. Eksporter-knappen gjenaktiveres i denne funksjonen.


```

14 //ekstra sikkerhetssjekk WP-funksjonalitet
15 var nonce = jQuery(this).attr("data-nonce");
16 jQuery.ajax({
17     type : "post",
18     dataType : "json",
19     url : myAjax.ajaxurl,
20     data : {action: "mh_db_preview_data", query_type : query_type, nonce: nonce},
21     success: function(response) {

```

Figur 8, viser AJAX funksjonen som da vil sende en http-forespørsel til serveren asynkront.

```

38 //funksjon looper gjennom resultatet av spørringen
39 //spørringen returnerer maks 30 linjer, dette er satt i funksjonen som avfyres i PHP i AJAX-kallet
40 jQuery.each(response, function(i, row) {
41     var tr_color = "#DDDDDD";
42     //for styling/farging av td elements
43     if (i%2 == 0){
44         tr_color = "#DDDDDD";
45     }
46     else {
47         tr_color = "#FFFFFF";
48     }
49     //konstruerer radene som tilhører kolonnene satt i tidligere array
50     const query_values = Object.values(row);
51     jQuery("#myTable")
52         .append(jQuery('<tr id=' + query_values[0] + '>'))
53         for (var j = 0; j < query_values.length; j++) {
54             jQuery("#" + query_values[0])
55                 .append(jQuery('<td bgcolor="' + tr_color + '>'))
56                 .append(jQuery('<p>'))
57                 .text(query_values[j])
58             )
59         }
60     } //for
61 } //each
62 } //successresponse
63 } //ajaxfunction

```

Figur 9, viser hvordan tabellen opprettes i HTML-DOM, i AJAX-funksjonen. (Se vedleggsmappe om tilgjengelig for full kildekode)

7. Videre arbeid

- Engasjere brukere og tilrettelegge for økt bruk
 - Administrator ønsker seg flere aktive brukere. For å få dette må det nok gjøres en innsats for å aktivisere brukerne som finnes, samt å innhente nye brukere. Det må skapes en dialog i forumet som er relevant. Kanskje en bør kontakte brukere og delegere ansvar for å skape dialog i en periode, slik at man får gang på debatten. Kanskje det er en aktuelt å starte et samarbeid med NAV eller lignende institusjoner?
- Spillifisering
 - Spillifisering er et konsept som er hyppig brukt i nettbasert brukerforum. Under det inngår ting som «likes» eller poeng og titler som en som bruker kan få rett til å ha/visse på bakgrunn av hvor mye en deltar. Dette har vist seg meget effektivt for å engasjere brukere, og brukes hyppig i forum. Ett eksempel er Facebook sin like funksjon. Et annet er StackOverflow¹⁶ sine titler/badges, poeng, «up-» og «downvotes» samt deres medaljer. Det engelske ordet er «gamification».
- DIFI retningslinjer bør oppfølges ordentlig – knapper i noen plugins er ikke aktive ved keyboard-valg.
 - DIFI har retningslinjer som ikke alltid følges 100% av alle utvidelser, ettersom disse er av internasjonal standard. Dette gjør at de gjerne ikke passer til den offentlige høye norske standarden. Dette betyr at disse nok må endres manuelt. Om dette er mulig å gjøre gjennom eventuelle utvidelser eller om en må inn i «template»-filene og endre utseende på den måten, har jeg ikke fått undersøkt ordentlig enda, ettersom dette er noe jeg relativt nylig fikk høre om fra en ressursperson.
- Aktivere gruppe-funksjonaliteten
 - Det er tatt i bruk en utvidelse som bygger opp gruppe-funksjonalitet. Denne har status aktiv, men den er ikke enda blitt satt opp ordentlig slik at det tilrettelegger for bruk av disse gruppene. Dette fordi gruppene er private, og man har ingen måte å finne dem på utenom å bli invitert eller ved å være moderator.
- Rapporter innhold i innlegg
 - Det bør lages en knapp som gjør at en kan rapportere upassende innhold. Dette vil gjøre jobben for eventuelle moderatorer mye lettere.
- Tilrettelegge for analyse av formaterte tekstdokumenter
 - Noen tekstformater er lettere å hente teksten ut fra enn andre tekstdokumenttyper. .doc og .docx typene er litt vanskelig å uthente teksten fra. Grunnen til at dette er vanskelig er at de er formaterte, som betyr at maskinen må ha en slags oversetter for å kunne filtrere ut metadata om teksten. Dette er også mulig å gjøre manuelt i for

eksempel PHP, men da må man gjøre det manuelt for alle forskjellige filtyper, noe som vil være svært tidkrevende. Eventuelt kunne en stilt det inn slik at det kun går an å laste opp dokumenter av visse formater. Dette kunne jo betraktelig senket arbeidsmengden om en ville gjøre dette manuelt. Dette er da mindre brukervennlig, men gjerne mer hensiktsmessig.

- Sette opp nyhetsbrevet slik at det fungerer ordentlig
 - Det er satt opp en utvidelse for nyhetsbrev, men på grunn av naturen til wordpress sine tidsbestemte hendelser, som kun kjøres når setup-hook/filter kjøres, så egner slike tidsbestemte hendelser seg svært dårlig til små websider med lite trafikk, fordi koden man vil skal kjøres hvert femte minutt ikke blir kjørt fordi den koden kun kjøres når noen laster inn siden. Hvis da ingen laster siden på en time, blir ikke funksjonen med koden utført og tidsbestemte hendelser skjer ikke og har en tendens til å skape problemer. Dette ville løse seg selv ved opptur i antall brukere og i aktivitet, men til den tid er det nok lurt å finne en bedre løsning.
- TLS-sertifikatet er utløpt
 - Vi holder på for tiden å migrere til et annet netthotell, som gjør at det ikke er betalt for ny TLS gjennom den nåværende tjenesten. Dette kan skape advarsler hos brukerne som sier at nettsiden ikke er sikker. Dette burde ordnes snarest mulig.

8. Evaluering av Gjennomføring

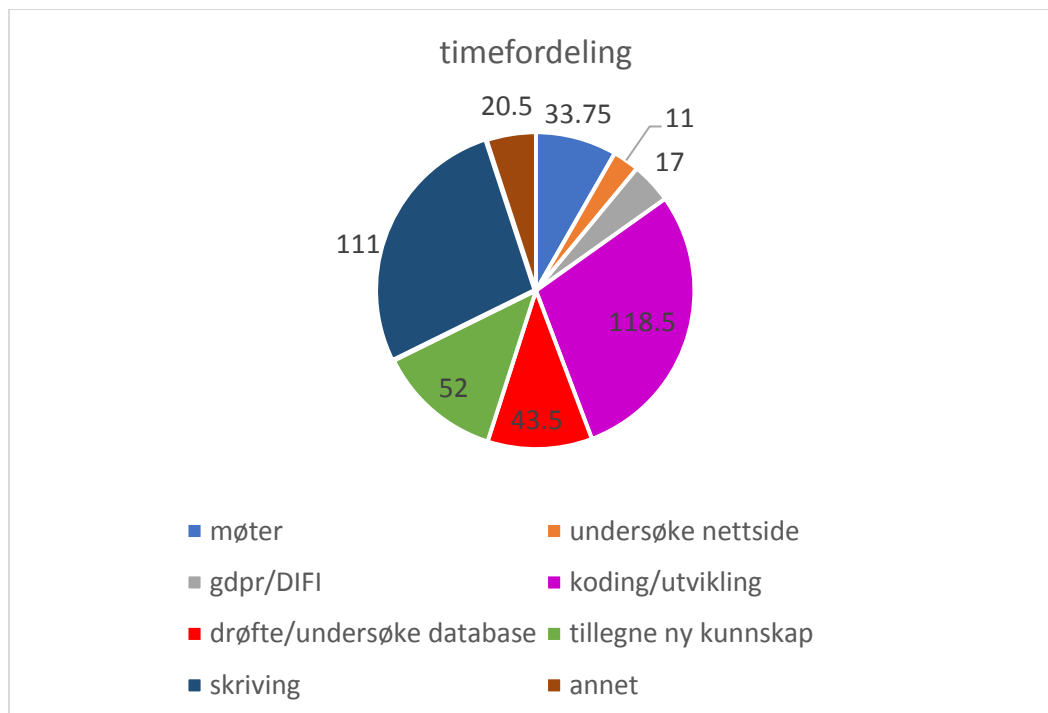
Prosjektet har et klart fokus på databaser, ettersom det i oppgaveteksten er snakk om å opprette en database. Det står ikke i oppgaveteksten, men prosjektet er i en WordPress-løsning. Dette er da et ganske kjent hjelpemiddel, som ofte brukes for å gjøre jobben med å lage en webside lettere. Det er tatt i bruk WordPress.org, som er noe mer fleksibel enn WordPress.com-løsninger, men samtidig krever det litt mer av brukeren. WordPress brukes generelt for at brukere av verktøyet (administratorer), skal ha mindre terskel for oppsett av nettsted og for å lette på hvor mye programmering en må kunne. Studenten har ikke hatt erfaring med akkurat dette redskapet før.

Fordi WordPress er et ganske stort redskap med mye dokumentasjon, har det tatt tid å lære hvordan en bruker det. WordPress-syntaksen er generelt en justert PHP-syntaks, med en hook- og filterstruktur, med bruk av forhåndsdefinerte objekter, samt templates.

WordPress har også støtte for både jQuery og AJAX og har en innebygd MySQL-database med lav grad av normalisering. Dette har også vært tidkrevende å sette seg inn i. Det tar alltid litt ekstra tid å finne ut av hvordan slike redskaper skal brukes til hvert enkelt tilfelle når en ikke har veldig mye erfaring med dem. Attpå kommer en del begreper eller uvante syntakser som en gjerne ikke kjenner til, som en må vende seg til. Utenom disse tre redskapene finnes det også en enorm mengde plugins som er programmert, som kan tilføye ekstra funksjonalitet.

8.1 Timebruk (tabell og kanskje graf/e.l.

Kakediagram over timefordeling:



Annet inkluderer blant annet arbeid med å fikse grunnleggende funksjonalitet på nettstedet og lesing av bakgrunnsprosjektets tilhørende bachelor.

Oversikt over møter:

29.01.2020	Oppstartsmøte med oppdragsgiver og veileder
05.02.2020	Møte med oppdragsgiver
13.02.2020	Møte med oppdragsgiver om nettstedet
13.02.2020	Møte med veileder om progresjon
18.02.2020	delta på zoom møte for å demonstrere zoom for MH ansatte/interessenter
20.02.2020	Møte med oppdragsgiver om krav og ansvarsområde
24.02.2020	Møte om hvor WP-config.php fila er, og om å få satt opp cron-jobs til nyhetsbrevsfunksjonalitet
28.02.2020	møte med veileder om status
03.03.2020	møte med TISIP om gdpr
11.03.2020	stormøte med samtlige interessenter m. fl. Gjennomgang av rapport for søkt økonomisk støtte til det større prosjektet
18.03.2020	oppdateringsmøte med veileder og oppdragsgiver
20.03.2020	møte med oppdragsgiver om .csv-filer og formater, samt hvilke data som er interessante
24.03.2020	Møte med oppdragsgiver og ressursperson om flytting til one.com eller ikke
02.04.2020	møte om gjennomgang innsendt kravspesifikasjon med veileder
23.04.2020	veiledningsmøte med veileder
24.04.2020	kjapt oppdateringsmøte med oppdragsgiver, vi tok også opp gdpr igjen
28.04.2020	møte med oppdragsgiver
30.04.2020	møte med veileder
01.05.2020	Gjennomgå ny personvernerklæring sammen med oppdragsgiver
20.05.2020	Møte med veileder og diskutere videre arbeid etter implementert data-uthentings-grensesnitt
25.05.2020	møte med oppdragsgiver om overføring til one.com, snakk om hvordan man tar backup for å forberede oss til migrasjon

8.2 Kontakt med ulike personer

Oppdragsgiver har gitt studenten tilgang til alle de ressurser som har vært nødvendig for at studenten skulle klare å gjennomføre oppgaven. De har møttes relativt ofte, ca. 1 gang i uken til vanlig.

Veilederen har vært nært tilknyttet skrivingen av diverse rapporter, og har veiledet på hvor ting bør ligge, hva som bør være med, hva som gjerne kan utelates, og ikke minst hvordan formulere ting på en mer hensiktsmessig måte. Det har vært møter mellom veileder og student omtrent annenhver uke.

Stig er en ressursperson som har bistått informasjon og kunnskap om WordPress og hvordan det fungerer. Disse har for det meste hatt kontakt gjennom e-post, gjerne en e-post-utveksling annenhver uke eller noe slikt. De har deltatt i samme møte ett par ganger.

TISIP har bistått med informasjon og anbefalinger når det kommer til gransking av gdpr og om nettstedet følger disse retningslinjene. Det har kun vært ett par møter med TISIP gjennom prosessen.

Irene er en sluttbruker som påpekte at det var et problem med forumsiden. Problemet var at sidene var private, som gjorde at alle uten tilstrekkelig tilgang på brukeren sin, ikke fikk delta. Dette ble løst i innstillingene for forumene. Hun har hatt indirekte kontakt med studenten gjennom oppdragsgiver.

Kurt er en sluttbruker som har vært med og testet nettstedet, og gitt noen tilbakemeldinger på hva som opplevdes bra. Kurt har deltatt på to møter med student og oppdragsgiver for å gi en brukers perspektiv på ting.

8.3 Refleksjonsnotat

Fra oppgavestart til slutføringen har jeg gjort meg en del tanker rundt prosjektet som helhet, og delprosjektet som kommer med bacheloroppgaven. Jeg velger å utdype disse her ettersom de vil gi et innblikk i hvilke tanker og kjenninger jeg har gjort gjennom arbeidet med oppgaven, og forhåpentligvis gi litt innsikt i hvorfor noen valg er tatt. Det er også viktig å gjennomgå en prosess etter den er utført for å kunne forbedre ens beredskap til lignende situasjoner i fremtiden. Dette er ikke bare en tanke, men kommer tydelig frem i mange modeller, bøker og artikler om og for prosessforbedring. For å nevne noen:

- ITIL sine retningslinjer for livssyklusfasen kontinuerlig forbedring for tjenester
- Workflow modeling (Sharp & McDermott 2009)
- Virksomheters informasjonssystemer (Trond R Braadland 2002)

I tillegg er det et eget fagfelt, BPM, som står for «business process management», som går på dette å forbedre prosesser, hvor en evaluering av en prosess er uunngåelig. Jeg kommer ikke til å bruke modeller, men heller konsepter fra prosessforbedring i refleksjonen.

8.3.1 Oppgavevalg

Ved valg av oppgaven tenkte jeg som student, at det ville være svært utfyllende for graden min å ha jobbet med videreutvikling av et allerede etablert og fungerende nettsted. Da jeg valgte oppgaven hadde det oppstått en ide om at jeg overtok noen andre sin kode, og skulle lese gjennom, kanskje forbedre, og sannsynligvis videreutvikle. Det var først ved oppstartsmøtet at det ble klargjort at det var snakk om en WordPress-løsning. Dette var en smule uventet, men burde ikke være en kilde til særlige problemer. Det kjentes litt skuffende, men det er naturlig når en har forventninger som ikke stemmer. Derfor tenker jeg at neste gang jeg har et lignende prosjekt, bør jeg spørre om språk og rammeverk som er tatt i bruk.

8.3.2 Manglende erfaring med NoSQL

Jeg har ikke jobbet med NoSQL før, og kjenner derfor ikke til alle fordelene i praksis. Dette begrenser min forståelse for nytten det kunne brakt til løsningen. Likevel, etter å ha lest om det flere steder, ser det ut som om en relasjonell database er forsvarlig å ta i bruk i dette tilfellet. I det større prosjektets

skriv, er det skrevet at det kun antas å være 2000 potensielle brukere i første omgang. Dette er ikke en mengde brukere og informasjon som er stor nok til at alle fordelene ved NoSQL blir særdeles attraktive. Det man gjerne vil savne mest er fleksibiliteten til å lagre mange flere forskjellige datatyper, men dette lar seg gjøre på andre måter på tjeneren. Dermed tør jeg si at det en kommer til å savne mest, er et felles grensesnitt for uthenting av samtlige tekstdata, og gjerne video- og audio-data, som kunne vært interessant å analysere. Det som gjør dette vanskelig uavhengig av valg av database, er at f.eks. word-dokumenter er formaterte tekst-dokumenter, som gjør dem litt vanskelige å tolke for maskiner som ikke har et tolkningsprogram, en såkalt «parser»¹⁷. Dermed blir det uansett vanskelig å uthente teksten fra f.eks. et .doc- eller .docx-dokument. Akkurat disse to filtypene er gjerne særegent vanskelig fordi formateringen til Microsoft har vært skjult så lenge, og gode tolkningsalternativer er derfor blitt neglisjert¹⁸. Et alternativ her kunne være å kun tillate opplasting av noen typer dokumenter/filformater, og laget støtte for å kun tolke disse.

Det var vanskelig å formulere krav. I møter følte det av og til vanskelig å få konkrete svar på data som var ønskelige å samle. Jeg tror at dette delvis kan ha noe med at det egentlig ikke var noen som visste hvilke alternativer som fantes, og jeg kjente det som om jeg egentlig ikke visste nok selv til å være en veldig aktiv pådriver for endringer og krav. Fordi nettstedet har blitt opprettet av to andre studenter som gjorde seg ferdige sist halvår, var de to som sannsynligvis visste best hva som ble lagret og hva som kanskje kunne komplimentere den dataen som allerede ble lagret, ikke til stede. Heldigvis fantes det noe dokumentasjon på det de har gjort. På side 47 i deres egen bachelorgrad¹⁹, har de kommet med en konklusjon om at en NoSQL-database hadde vært god å ta i bruk. Dette stiller jeg meg imot fordi det er kommet frem at administrator vil bruke en spesifikk analyse-programvare til å strukturere og analysere dataene. Dermed er det ikke et behov for at NoSQL-databasen skal gjøre dette, en må simpelthen kunne fremhente dataene i et format som administrators programvare kan jobbe med.

En annen grunn til at en separat database hadde vært fint er at den hadde blitt mer ryddig, men en måtte da laget et eget grensesnitt mot denne databasen for å uthente dataene, ettersom WordPress ikke har støtte for dette. I tillegg blir det et spørsmål om hvor en skal lagre sagt database. Noen webhotell har støtte for dette, blant annet det vi bruker (GoDaddy), dersom en har rett abonnement/betalingsplan.

8.3.3 Undersøke arbeid med GDPR

Etter jeg så at oppgaven gjerne enten ble mindre enn først tenkt eller mye større, så jeg det aktuelt å lete etter andre ting å holde på med slik at det ble noe å skrive om og vise til. På grunn av det nye GDPR-reglementet utstedt av EU fra 2018 drøftet en del av interessentene dette sammen, og jeg påtok meg oppgaven om å sjekke at nettstedet dekket disse kravene. Dette er også i problemstillingen i den forrige bacheloroppgaven, men – det tas ikke stilling til i særlig detalj. ([Se side 52 av den andre oppgaven](#)) Derfor var det en del å undersøke her, men beklageligvis, eller heldigvis, ikke så mye mer å gjøre heller. Kildene jeg brukte til å undersøke dette var UNIT sine diverse maler og veiledninger²⁰.

8.3.4 DIFI – retningslinjer

Jeg har også kjapt undersøkt om DIFI sine retningslinjer blir fulgt. De blir i all hovedsak dette, men jeg må gå inn og sjekke hvor stor kontrasten er noen steder for å bli sikker. Det er noen knapper man kan navigere til som ikke er spesielt tastatur-vennlige, ved at de ikke markeres visuelt når en bruker f.eks.

TAB-knappen til å navigere. Jeg burde også sjekke om disse lenkene har en ALT-tekst, som gjør det lettere for folk som har støtte fra lese-programvare å forstå hva lenken peker til. Dette satt jeg meg inn i såpass seint at jeg ikke har fått gått inn og sjekket detaljene, men vet at det sannsynligvis ligger noen småting å rydde opp i. Disse retningslinjene skal helst følges fordi det er et offentlig nettsted. Noen av disse detaljene vil ta litt tid å finne ut av fordi elementene det gjelder gjerne er skapt av programutvidelser, som da naturligvis ikke nødvendigvis tar hensyn til DIFI sine retningslinjer.

Noe annet som er interessant å merke er at det kommer et nytt utformingsdirektiv fra EU i 2022 som trer i kraft 2025. Gjerne også disse kravene da blir nærere oppfulgt, og at en da ser bøter eller lignende som gjerne blir høyere enn vi er vant til, slik vi har sett med GDPR 2018.

8.3.5 Godaddy

GoDaddy som har vår tjener på deres webhotell, har det vært en uventet utfordring å ta i bruk. Dette fordi nettstedet deres kan oppleves som noe uoversiktlig. De har nylig byttet utseende, og da forsvant også alle forum-innleggene deres, slik at det ikke fantes noe oppdatert forum som kunne bistå hjelp når en selv ville løse problemer²¹. Altså har dette webhotellet vært til uhyre frustrasjon, med mange kryptiske menyer med en merkverdig navigasjon.

9. Kilder

Lenke til prosjektet på stiftelsen Dam sine sider:

<https://www.dam.no/prosjekter/forum-for-brukerrepresentanter/> nedlastet 01.02.2020

lenke til Mental Helse sitt nettsted:

<https://mentalhelse.no/> nedlastet 22.05.2020

lenke til TISIP sitt nettsted:

http://tisip.no/faces/public/om_tisip/index.jsp nedlastet 22.05.2020

lenke til regjeringen.no som viser fylkessammenslåingen:

<https://www.regjeringen.no/no/tema/kommuner-og-regioner/regionreform/regionreform/nye-fylker/id2548426/> nedlastet 25.05.2020

artikkel om Bergen kommunes brudd på GDPR med tilhørende bot:

<https://www.dn.no/teknologi/personvern/gdpr/bergen-kommune/datatilsynet-varsler-millionbot-til-bergen-kommune/2-1-503797> nedlastet 03.03.2020

artikkel om bruk av grafisk datamodellering for databaser:

<https://dzone.com/articles/what-are-the-pros-and-cons-of-using-a-graph-databa> nedlastet 06.03.2020

artikkel som sammenligner bruk av relasjonell datamodellering opp mot dokumentmodellering:

<https://oracle-patches.com/en/databases/3978-relational-model-versus-document-model> nedlastet 06.03.2020

artikkel om nøkkelverdidatamodellering:

<https://www.aerospike.com/what-is-a-key-value-store/> nedlastet 06.03.2020

artikkel om nøkkelverdidatamodellering:

https://en.wikipedia.org/wiki/Key-value_database nedlastet 06.03.2020

artikkel om vid-kolonne datamodellering:

<https://bi-insider.com/posts/wide-column-column-family-nosql-database/> nedlastet 06.03.2020

lenke til Riak – NoSQL redskap:

<https://riak.com/products/riak-kv/index.html> nedlastet 06.03.2020

lenke til stackoverflow tråd om bruk av NoSQL i kombinasjon med WordPress:

<https://stackoverflow.com/questions/14371819/can-i-use-nosql-database-in-wordpress> nedlastet 06.03.2020

lenke til stackoverflow tråd om lagring av dokumenter i MySQL:

<https://stackoverflow.com/questions/3801438/store-and-retrieve-word-documents-with-mysql>

nedlastet 07.03.2020

lenke til «hello dolly» sin utvidelseskode:

<https://plugins.trac.wordpress.org/browser/hello-dolly/trunk/hello.php> nedlastet 29.04.2020

lenke til prefix-anbefalingene i plugin-håndboka:

<https://developer.wordpress.org/plugins/plugin-basics/best-practices/#prefix-everything> nedlastet

30.04.2020

lenke til badges siden til stackoverflow:

<https://stackoverflow.com/help/badges> nedlastet 28.05.2020

lenke til forumspørsmål om å hente tekst fra word-dokumenter gjennom php:

<https://stackoverflow.com/questions/10646445/read-word-document-in-php> nedlastet 06.03.2020

lenke til forumspørsmål om å hente tekst fra word-dokumenter gjennom php:

<https://www.sitepoint.com/community/t/php-get-text-from-word-document-doc/6922> nedlastet

27.05.2020

lenke til bacheloroppgaven som munnet ut i nettstedet som denne oppgaven er basert på:

<https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2609670/no.ntnu%3Ainspera%3A2338765.pdf?sequence=1&isAllowed=y> nedlastet 30.01.2020

lenke til UNIT sine ressurser for arbeid med GDPR:

<https://www.unit.no/ressurser-arbeidet-med-gdpr> nedlastet 09.03.2020

lenke til tomme godaddy forum:

<https://no.godaddy.com/community> nedlastet 01.06.2020

lenke til stackexchange-tråd som forklarer hvordan registrere forms i WP:

<https://wordpress.stackexchange.com/questions/177076/post-form-request-with-admin-post> nedlastet

12.05.2020

lenke til artikkel som beskriver hvordan eksportere sql til fil via php:

<https://www.itechempires.com/2016/05/export-data-from-mysql-to-csv-using-php/> nedlastet

28.04.2020

lenke til github repository som beskriver hvordan eksportere \$wpdb-spørring til .csv:

<https://gist.github.com/ddbs/9545223> nedlastet 13.05.2020

Lenke til en artikkel som beskriver hvordan man looper seg gjennom JavaScript objekter:

<https://zellwk.com/blog/looping-through-js-objects/> nedlastet 15.05.2020

lenke til artikkel som beskriver hvordan man setter opp et AJAX-request i WP:

<https://premium.wpmudev.org/blog/using-ajax-with-wordpress/> nedlastet 14.05.2020

lenke til stackoverflow-tråd som beskriver hvordan dynamisk oppretter et HTML-table med jQuery:

<https://stackoverflow.com/questions/171027/add-table-row-in-jquery> nedlastet 20.05.2020

lenke til hvordan data nonces fungerer i WP:

<https://developer.wordpress.org/plugins/security/nonces/> nedlastet 14.05.2020

Vedlegg

Oppgavetekst

Arbeidstittel: Digitalt samhandlingsforum med registrering av data

Hensikten med oppgaven:

Bygge videre på en web-portal for et nettbasert samhandlingsforum (Community of Practice) for ca 2000 brukere innen helse og sosialsektoren (brukerrepresentanter). Et sted der brukerne kan samle relevant informasjon og utveksle erfaringer. Portalen finnes i dag på URL: <https://brukerforum-mh.org/>

En ønsker å bygge opp en database / kunnskapsbase med data og dokumenter knyttet til bruken av forumet. I tillegg til evt. forbedringer / utvidelser av funksjonaliteten.

Kort beskrivelse av oppgaveforslag:

Våren 2019 ble det gjennomført en BA-oppgave som utviklet en web-portal i et samarbeid mellom Mental Helse, TISIP og NTNU. Resultatet var en kravspesifikasjon og en web-side (<https://brukerforum-mh.org/>). Det er ønskelig å bygge videre på denne oppgaven med fokus på lagring / registrering av data knyttet til bruken av forumet. Aktuelle data er mulige metadata, trafikkdata og dokumenter (som pdf). Evt. lagring av spørreskjema. Mulige formelle behov knyttet til arkivering. Noe av denne datalagringen kan senere knyttes opp mot forskning / dataanalyse. Det er aktuelt å vurdere mulige databaseverktøy som MySQL og NoSQL (lagring av semistrukturelle data), samt bruk av filsystem (lagring av dokumenter). En vurdering og implementasjon av utvidelser av funksjonaliteten til forumet kan også være aktuelt.

Opgaven passer for (kryss av de(t) som passer og skriv evt. en kommentar til oss):

- Bacheloroppgave
- passer for Informasjonsbehandling eller ingeniørstudenter

Kan oppgavestiller stille arbeidsplass med ikke arbeidsplass, men programvare ved behov nødvendig utstyr og programvare:

Hvis ikke, hva kreves av maskin og programvare: PC med aktuelle plug-ins

Begrensninger i tilgjengelighet av opplysninger o.l.: Testdata inneholder personopplysninger som må behandles i henhold til Personopplysningsloven

Opgaven passer best for, antall studenter:

- 1
- 2

Opplysninger om oppgavestiller

Navn på bedrift eller organisasjon: AIT (NTNU) i samarbeid med TISIP og Mental Helse

Adresse TISIP, Postboks 4419

Postnummer 7418

Poststed	Trondheim
Navn på kontaktperson/veileder:	Tore Mallaug
Telefon:	99238232
Epost:	tore.mallaug@ntnu.no
Navn på kontaktperson 2/veileder 2:	Thorleif Hjeltnes
Telefon kontaktperson 2/veileder 2:	95293091
Epost kontaktperson 2/veileder 2:	thorleif@tisip.no

Utfyllende kommentarer til hva oppgaven gjelder:

Oppgaven er en del av prosjektet «Forum for brukerrepresentanter» som er et samarbeid mellom Mental Helse (prosjekteier), Stiftelsen TISIP og NTNU, IDI - AIT. Prosjektet går over 2 år og data-løsningene som utvikles / velges skal i første omgang benyttes av en gruppe frivillige bruker-representanter som skal evaluere løsningen sammen med en

masterstudent fra AIT.

Tilrettelegging for forskning på data ved nettbasert forum – Forstudierapport

1 Introduksjon – hensikten med dokumentet

Dette dokumentet skal være en forstudierapport for utviklingsprosjektet som utføres for Mental Helse på vegne av studenten i samarbeid med NTNU og TISIP. Hensikten er å få oversikt over hvilke oppgaver som gjenstår, drøfte risikomomenter, klargjøre mål, presentere interessenter, drøfte suksessfaktorer, og fremlegge retningslinjer og standarder som gjelder prosjektet. Sist vil det skrives en anbefaling om arbeidet bør videreføres.

2 Bakgrunn for prosjektet

2.1 Beskrivelse av problemer og behov

Grunnen til at prosjektet settes i gang er at Mental Helse lyst å forbedre plattformen deres for deling av ideer, erfaringer og beste praksis blant brukerrepresentanter, og fordi de potensielt vil drive med datanalyse av forumbrukernes aktivitet.

Forumet brukes svært lite, og vi må undersøke og finne ut hvorfor, og deretter prøve å forbedre forumet, slik at flere finner nytteverdien høy nok til å ta det i bruk.

Den dag i dag finnes det ikke noe godt og lett tilgjengelig forum for akkurat dette. Jeg antar her at oppdragsgiver og tidligere arbeidende har undersøkt markedet for et slikt, ettersom prosjektet startet et år siden.

De vil koble sammen brukerrepresentanter over hele landet.

2.2 Kort om dagens systemer og rutiner

Forumbrukerne kan delta på webinarer og logge inn og sende hverandre meldinger i dagens system.

De skal kunne diskutere på forumene, men forumene fungerer ikke slik de skal.

En må registreres og logge inn for å ta i bruk forumfunksjonaliteten.

I dag kan administratorer opprette og slette innlegg og dokumenter som ligger ute.

Oppdragsgiver har frem til nå hatt ansvar for forvaltning og drift av nettstedet. Samme person er med Mental Helse eier og finansierer.

Brukerne er alle som er interesserte i å være, eller er brukerrepresentanter.

- Hvem har bruk for å vite noe om produktet?
Dem som administrerer, andre brukere, gjerne også NTNU og TISIP

- Hvem har bruk for å vite noe om prosjektforløpet?
NTNU og spesielt Mental Helse m. samarbeidspartnere.

- Hvem utfører arbeidet?
Felix Grimsrud, student ved NTNU

- Hvem bidrar eller leverer?
Karl-Johan Johansen, Tore Mallaug, Thorleif Hjeltnes og andre administrerende brukere som brukertester

Interessent	Suksesskriterier	Bidrag til prosjektet
Eksterne (kunde) <ul style="list-style-type: none"> - Oppdragsgiver - Brukerrepresentanter - TISIP 	Et velfungerende nettforum for brukerrepresentanter Få kontakt med andre brukerrepresentanter, og få en arena for konstruktiv diskusjon og meningsdeling. Et nettsted som følger standarder for sikkerhet og brukeropplevelse	Beslutninger, brukerbehov Brukertesting Veiledning, kunnskap
Interne (leverandør) <ul style="list-style-type: none"> - Prosjektleder - Veileder 	Vellykket produkt og prosjektrapport Et godt gjennomført prosjekt	Utføre de beslutninger en kommer frem til Kunnskap, veiledning

4.2 Rammebetingelser

Absolutte krav til ferdigdato

Prosjektet skal være ferdig til den 20.05.2020

Absolutte krav til kostnadsramme

500 arbeidstimer

Drifts- og utviklingsmiljø

Skal helst kunne teste endringer uten å påvirke det aktive nettstedet.

Tilrettelegge for forskning og analyse.

5 Kritiske suksessfaktorer

5.2 Suksessfaktorer

- Det er kritisk at vi finner raskt ut om det er mulig å sette opp en egen database som kan hente og lagre informasjon fra WordPress sine tjenester.
- Det er kritisk at forumet virker slik det skal slik at brukere får tilstrekkelig brukernytte og faktisk tar nettstedet i bruk.
- Det er kritisk at alle parter bidrar til at møter blir så effektive som mulig. Dersom en bes om å ha noe klart til møtetid bør det forventes at alle har gjort det de skal. Hvis ikke blir møter uten mål og mening og mister sin spissende hensikt. I slike tilfeller kan det være mer aktuelt å utsette møter eller bare få info gjennom andre kommunikasjonskanaler.
- Det er kritisk at studenten får satt seg inn i redskapene som er brukt tidligere for å kunne arbeide videre med dem.

5.2 Informasjonsbehov

- Studenten vil ha behov for å vite brukerbehov og regler/standarder for behandling av informasjon.
- Oppgavestiller vil ha behov for å vite hva som er mulig å gjøre med teknologien anvendt, og hva som kan gjøres når systemet er ferdigstilt. (oppgave er stilt slik at en akkurat nå er litt usikker på hva som er bestilt, hva som skal gjøres, og hva som kan gjøres).
- TISIP vil gjerne ha behov for å vite hvordan div. informasjon benyttes og kommuniseres av systemene, samt hvor brukervennlig systemet er.
- Veileder vil ha behov for å vite hvordan fremdriften er i prosjektet.

6 Risikoanalyse

Den teknologien som i dag er i bruk, er for prosjektlederen ny og uprøvd og kan derfor være utfordrende å ta hensiktsmessig i bruk. Det kan oppstå problemer med å skaffe ressurser til prosjektlederen. Vi kjenner ikke datakvaliteten i eksisterende database, det kan bety problemer når vi skal laste over data. Programmer og utvidelser tatt i bruk slutter å fungere eller være kompatible. Hvem kan så rette opp i dem?

7 Retningslinjer og standarder

Det er kun en som lager dokumentene, dermed faller alt ansvar for godkjenning og revisjon på denne. Det er nok likevel lurt å forhøre seg med veileder for å sikre at de blir gjennomført på forsvarlig vis.

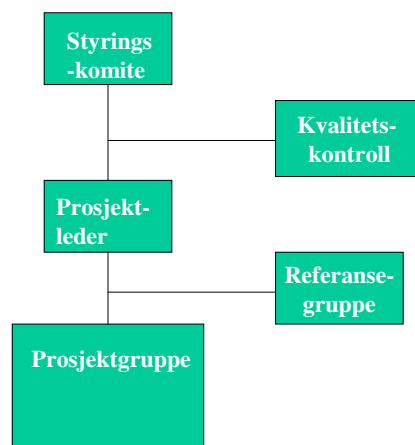
Dokumenter som skal produseres er:

- Forstudie-rapport
- Kravspesifikasjon
- Sluttrapport

8 Prosjektorganisering

I dette kapitlet skal vi vise hvem som er med i prosjektet og hvordan de har fordelt arbeidet mellom seg.

I en "klassisk" prosjektorganisasjon har vi følgende roller:



I studentprosjekter er det vanligvis naturlig med en enkel prosjektorganisasjon. Det som er viktig er å klarlegge

følgende:

1. Hvem har oppdragsgivers rolle? Vedkommende navngis og fyller rollen som styringskomite for prosjektetsammen med veileder. Eventuelt kan oppdragsgiver være representert med flere enn en person.
2. Hva gjør vi med rollene: kvalitetskontroll og referansegruppe? Veileder vil ha rollen som kvalitetskontroll, sett fra skolen side. I tillegg må gruppen vurdere om en eller flere representanter fra brukerne kan fylle rollen som kvalitetskontroll evt. referansegruppe. Oppgaven for disse blir å vurdere kvaliteten på resultatet, sett fra oppdragsgivers side.
3. Hvordan skal rollen som prosjektleder ivaretas? Skal en person i prosjektgruppa ha denne rollen permanent eller skal den gå på omgang, og hvordan skal skifte skje?
4. Hvordan skal prosjektmedlemmene dele arbeidet mellom seg? Skal vi allerede nå ha forskjellige roller eller skal arbeidsfordelingen vurderes etterhvert som oppgavene klargjøres?

Det gir ikke mening at oppdragsgiveren og veilederen alene er styringskomite ettersom de har sagt til meg (prosjektleder) at det er jeg som har ansvar for fremdriften og videreføring av prosjektet. Slik jeg oppfattet det på oppstartsmøtet gav de meg styringen, og sa at det var opp til meg om prosjektet ble noe. Jeg må vel (minst?) delta i styringen på lik linje med veileder og oppdragsgiver.

Veileder ved NTNU skal være ansvarlig for kvalitetssjekke fremgangen av arbeidet. Vi skal ha flere brukere som kan være en del av en referansegruppe. De koordineres av oppdragsgiver. TISIP skal gjerne også hentes inn for å kvalitetskontrollere sikkerheten til personopplysninger o.l., samt kanskje brukervennligheten.

Det er vel et statlig prosjekt, og det skaper høyere krav til universell utforming, som kanskje ikke er hovedfokus i denne omgang, men som kan være fint å ha i bakhode. Det vil være mulig å skrive noe om i en eventuell konklusjon og videreføringsrapport.

Felix Grimsrud blir prosjektleder ettersom han er den som har ansvaret for at prosjektet ferdigstilles.

9 Anbefaling om videre arbeid

Arbeidet anbefales videreført ettersom det ser ut til at forumet kan bli en god ressurs for brukerne.

Kravspesifikasjon for: Tilrettelegging for forskning på data ved et nettbasert forum

Innholdsfortegnelse

- 1 introduksjon
- 2 use case og funksjonelle krav
- 3 wireframes
- 4 andre krav
- 5 referanser
- 6 vedlegg/bilde av wireframes

1 Introduksjon

Dette dokumentet presenterer krav og behov som er satt til et forum for brukerrepresentanter, igangsatt av Mental Helse og TISIP. Kravene er utformet av studenten basert på møter, ønsker og annen informasjon som er kommet til lys i løpet av samarbeidsprosessen. Kravene presenteres i detalj senere i dokumentet, og man får på denne siden kun en kjapp introduksjon av dem.

Krav:

- Administrator må få tak i data som de kan bruke til analyse og helst forskning.
- Bør følge DIFI sine retningslinjer. Ettersom det er en offentlig nettside, er det et automatisk krav at det følger DIFI sine retningslinjer for universell utforming.
- Ettersom det er snakk om personopplysninger i en profesjonell sammenheng, med taushetsplikt, vil krav til sikkerhet være ganske høye. Disse kravene ser ut til å ivaretas gjennom den nåværende løsningen i all hovedsak.
- Møter og webinar bør være krypterte. Dette ordner Zoom. De er også EU-sertifisert gjennom Secure Shield Sertifisering. Oppdragsutlyser vil også kunne ta logg av chatten og opptak av webinar. Alt dette skal være mulig å få til i Zoom.
- Må følge EUs krav til GDPR
- Administrerende behov:
 - De administrerende av nettstedet ønsker å ha en nyhetsfeed der de kan legge ut nyheter der de nyeste innleggene kommer øverst.
 - De administrerende ønsker å ha et nyhetsbrev som er frivillig å abonnere på, der de kan sende ut info om møter/webinarer etc.
 - Det er beskrevet behov for et ordentlig skjema for tilbakemeldinger for nettstedet.
 - Karl Johan ønsker å analysere innhold i innlegg, kommentarer og meldinger. Dette ligger til grunn i forskning og saklighet. Informasjonen han vil ha tak i er bl.a. Innhold, bruker, tidspunkt, tittel, Tema, bakgrunnsdata (er skribent brukerrepresentant, region, hvor etc.)

2 Use case og funksjonelle krav

Navn	Uthenting av data fra databasen
Mål	Hente ut data fra databasen i et format som er hensiktsmessig til å studere/analysere
Aktør	Administrerende for nettstedet, samt forskere for MH
Trigger	Aktøren vil først og fremst drive innholdsmessig forskning på data om innlegg som ligger ute på nettstedet. Vedkomne ønsker senere å sette sammen en søknad til utbygging eller økt støtte for prosjektet. Derfor anses all informasjon om bruk og nyttig innhold som aktuelt.
Pre-betingelse	Vi antar at administrator er pålogget og at nettstedet fungerer slik det skal.
Post-betingelse	Aktøren får (laste ned) dataene i et format som kan tas i bruk i vedkomnes analyse-programvare
Hovedløp	<p>Normal hendelsesflyt (hva skjer når alt går bra).</p> <p>Steg 1: brukeren går i menyen/nav-bar og trykker seg frem til siden for uthenting av informasjon fra databasen.</p> <p>Steg 2: brukeren velger alternativ «kjør spørring etter posts» i en nedtrekksmeny, og lar søkefeltet stå tomt.</p> <p>Steg 3: brukeren krysser av på hvilket filformat de ønsker at spørringen skal lagres i.</p> <p>Steg 4: brukeren trykker på en knapp «hent info».</p> <p>Steg 5: Hele eller deler av tabellen som er et resultat av spørringen kommer frem på skjermen.</p> <p>Steg 6: brukeren trykker på «last ned» knapp for å laste ned filen.</p> <p>Steg 7: fil med full tabell lastes ned på lokal maskin.</p>
Sideløp	<p>Alternativ hendelsesflyt fra hovedløp. F.eks:</p> <p>Steg 2a: brukeren velger alternativ «kjør spørring etter brukerdata» i en nedtrekksmeny, og lar søkefeltet stå tomt.</p> <p>Steg 3a: brukeren krysser av på hvilket filformat de ønsker at spørringen skal lagres i.</p> <p>Steg 4a: brukeren trykker på knapp «hent info».</p> <p>Steg 5a: Hele eller deler av tabellen som er et resultat av spørringen kommer frem på skjermen.</p>

	<p>Steg 6a: brukeren merker på tabellen at dette ikke er data om posts, men heller data om brukere, og innser at de har valgt feil alternativ i nedtrekksmenyen.</p> <p>Steg 7a: bruker går tilbake til steg 2 og sjekker denne gang at alt er på plass, slik at prosedyren går etter normalt løp.</p> <p>Steg 5b: Skjermdumpen viser en feilmelding som tilsier at webløsningen ikke fikk kontakt med databasen og at dataauthenticeringen var mislykket.</p> <p>Steg 6b: brukeren får ikke hentet ut dataene den ønsker.</p>
Spesielle krav	<p>Brukeren vil få noe mer begrenset data, men slipper da også å sette seg like godt inn i databasens struktur. Spørringene vil være programmert på forhånd, som gjør det litt mindre fleksibelt.</p>

Jeg har fått høre fra en ressursperson at DIFI sine retningslinjer til universell utforming ikke oppfylles med dagens tema. Etersom det er en offentlig nettside, er det et automatisk krav at det følger DIFI sine retningslinjer for universell utforming. Jeg vet egentlig ikke hva DIFI sine retningslinjer nøyaktig er per dags dato, men har hatt et fag som har vært innoen det og vet derfor generelt hva det går i. Derfor må jeg undersøke retningslinjene nok en gang før jeg retter opp i eventuelle mangler.

Møter og webinar skal være krypterte. Dette ordnes gjennom Zoom. De er også EU-sertifisert gjennom Secure Shield Sertifisering. Oppdragsutlyser vil også kunne ta logg av chatten og opptak av webinar, slik det er blitt meldt et behov for. Alt dette skal være mulig å få til i Zoom, og det trengs nok bare mer opplæring/eksperimentering for å få til en mer effektiv bruk av verktøyet.

Det vil være et krav å følge det nye EU-reglementet for GDPR. Dette går på personvern, og er omtrent likt som det vi i Norge hadde fra før, med noen unntak, spørsmålstegn og mye større bøter. Et eksempel på større bøter er når Bergen kommune fikk millionbot for å ha lekket student-brukernavn og passord uten vilje. Et eksempel på spørsmålstegn er om norske leger fortsatt kan sende e-post med personopplysninger til pasienter dersom e-posten er kryptert.

Med dagens redskaper skal GDPR være mulig å ivareta på webløsningen slik vi nå har den. Både Zoom, WordPress og GoDaddy skal være GDPR-godkjent så lenge en bruker det rett.

Etersom det er snakk om personopplysninger i en profesjonell sammenheng, med taushetsplikt, vil krav til sikkerhet være ganske høye. Disse kravene ser ut til å ivaretas gjennom den nåværende løsningen i all hovedsak.

De administrerende for nettstedet har kommet med en rekke behov til funksjonalitet som de ønsker dekket av nettstedet:

De administrerende for nettstedet har meldt om et behov for en nyhetsfeed, der de kan legge ut innlegg om nyheter som sammenkomster og webinarer.

De administrerende ønsker å ha et nyhetsbrev som er frivillig å abonnere på, der de kan sende ut info om møter/webinarer etc.

Det er beskrevet behov for et ordentlig skjema for tilbakemeldinger for nettstedet. Dette kommer av et behov for å vite hva brukerne syntes og for å skape et godt utgangspunkt for forbedring.

Administrator ønsker å analysere innhold i innlegg, kommentarer og meldinger. Dette ligger til grunn i forskning og saklighet. Informasjonen han vil ha tak i fra databasen er bl.a:

Innhold, bruker, tidspunkt, tittel, Tema, bakgrunnsdata (er skribent brukerrepresentant, region, hvor etc.)

Spøringer til databasen:

Det er 5 tabeller i databasen som er spesielt interessante i databasen. Det er tabell for profil, data om profil, post, data om post og tabellen om aktivitet. I data om profil-tabellen, kan man finne informasjon om alle de punktene en må fylle når en registrerer seg. Her ligger altså alle svaralternativene i registreringer.

Profil-tabellen er lenket til data om profil-tabellen, og bruker den til å vise de forskjellige svarene som er avgitt hos den enkelte brukeren i ulike rader. Her ligger også rader som «navn» og «telefonnr», etc.

I data om post-tabellen finner vi informasjon om poster i post-tabellen. Disse to er veldig rotete og krungete å navigere. I post-tabellen lagres nemlig samtlige poster, og dersom en går og redigerer en post vil denne lagres i samme post-tabell som et nytt innslag. Det gjør det litt stress å finne den mest oppdaterte posten i en rekke revideringer, spesielt når de alle har omtrent de samme verdiene, med unntak av ID og dato. Data om post-tabellen sier noe om innholdet i hver post i post-tabellen.

I aktivitets-tabellen lagres informasjon om når en bruker sist var aktiv, invitasjoner til nye brukere/medlemmer, invitasjoner til grupper, og noen andre hendelser.

3 Wireframes

Det er ganske store behov for å hente informasjon ut fra databasen til å bruke i diverse analyseprogrammer. Disse dataene har varierende kvalitet og nytte for de som vil drive analysen. Dette gjør det vanskeligere å hente ut denne informasjonen. Det må lages et grensesnitt som en kan hente den nyttige informasjonen ut fra. Utvikleren/studenten skal da helst utvikle dette, men jobber akkurat nå heller med å finne ut av hvilken informasjon som er relevant å hente ut, slik at en har en mer håndfast ide over kravene til utseende.

Det er laget to wireframes for utseende av siden for uthenting av data fra databasen. Disse ligger vedlagt på slutten av dokumentet.



Digital samhandlingsplattform



Plattformen skal bidra til erfaringsoverføring mellom brukerrepresentanter og styrke rekruttering.

Man må være registrert og logget inn på nettsiden for å delta i samhandlingen og få tilgang til alle funksjoner.

Kom i gang

Populære innlegg

I dag ser forsiden slik ut, men jeg har hørt at oppdragsutlyser gjerne vil ha bildet mindre, ettersom det går an å scrolle ned, og det blir for langt å scrolle før en finner noe nyttig. Jeg tenker også å gjøre navbaren noe tynnere, ettersom jeg syntes den nå er en smule stor. Bildene i populære innlegg kan kanskje også være mindre. Jeg håpte på å få frem dette i den vedlagte wireframen, men vet ikke om det ble like tydelig som jeg ville.

4 Andre krav

Hele hensikten med nettstedet:

Ansatte brukerrepresentanter trenger en arena for diskusjon av arbeidet de gjør.

«De føler seg ensomme» - interessant fra TISIP

5 Referanser

<https://www.unit.no/ressurser-arbeidet-med-gdpr>

<https://zoom.us/docs/doc/Zoom-Security-White-Paper.pdf>

<https://www.dn.no/teknologi/personvern/gdpr/bergen-kommune/datatilsynet-varsler-millionbot-til-bergen-kommune/2-1-503797>

Wireframes ligger vedlagt under

HENT DATA TIL ANALYSE

Velg tabell/datagrunnlag

Søk i tabell/datagrunnlag

LAST NED DATA/TABELL

(valgt)
tabell : Tabellnavn

Søkeord : "Søkeord"

Filttype

(eksempel på)

Utseende Tabell : Tabellnavn

Id	K1	K2	K3	K4
1	mm	mm	mm	19.02.2020
2	m	a	m	22.02.2020
3	...			

-
- ¹ <https://www.dam.no/prosjekter/forum-for-brukerrepresentanter/>
 - ² <https://mentalhelse.no/>
 - ³ http://tisip.no/faces/public/om_tisip/index.jsp
 - ⁴ <https://www.regjeringen.no/no/tema/kommuner-og-regioner/regionreform/regionreform/nye-fylker/id2548426/>
 - ⁵ <https://www.dn.no/teknologi/personvern/gdpr/bergen-kommune/datatilsynet-varslar-millionbot-til-bergen-kommune/2-1-503797>

 - ⁶ <https://dzone.com/articles/what-are-the-pros-and-cons-of-using-a-graph-databa>
 - ⁷ <https://oracle-patches.com/en/databases/3978-relational-model-versus-document-model>
 - ⁸ <https://www.aerospike.com/what-is-a-key-value-store/>
 - ⁹ https://en.wikipedia.org/wiki/Key-value_database
 - ¹⁰ <https://bi-insider.com/posts/wide-column-column-family-nosql-database/>
 - ¹¹ <https://riak.com/products/riak-kv/index.html>
 - ¹² <https://stackoverflow.com/questions/14371819/can-i-use-nosql-database-in-wordpress>
 - ¹³ <https://stackoverflow.com/questions/3801438/store-and-retrieve-word-documents-with-mysql>
 - ¹⁴ <https://plugins.trac.wordpress.org/browser/hello-dolly/trunk/hello.php>
 - ¹⁵ <https://developer.wordpress.org/plugins/plugin-basics/best-practices/#prefix-everything>
 - ¹⁶ <https://stackoverflow.com/help/badges>
 - ¹⁷ <https://stackoverflow.com/questions/10646445/read-word-document-in-php>
 - ¹⁸ <https://www.sitepoint.com/community/t/php-get-text-from-word-document-doc/6922>

 - ¹⁹ <https://ntnuopen.ntnu.no/ntnu-xmlui/bitstream/handle/11250/2609670/no.ntnu%3Ainspera%3A2338765.pdf?sequence=1&isAllowed=y>
 - ²⁰ <https://www.unit.no/ressurser-arbeidet-med-gdpr>
 - ²¹ <https://no.godaddy.com/community>

