

Semantisk annotering av læringsmateriale

Hvordan gjenbruk av ressurser i diskusjonsforumer
hjelper studenter i deres kollaborative læringsprosess

Masteroppgave for Richard Persen
Institutt for informasjons- og medievitenskap
Universitetet i Bergen
Vår 2008



richard.persen@student.uib.no

richard.persen@gmail.com

Nøkkelord: semantisk web, datastøttet læring, semantikk, annotering, ontologi, konseptuell domenemodell, gjenbruk av læringsressurser, diskusjonsforumer

Forord

Denne oppgaven hadde ikke vært mulig å gjennomføre uten hjelp fra en rekke personer. Jeg vil først og fremst få takke min veileder, Weiqin Chen, som har vært til stor hjelp under hele arbeidet, og som har kommet med verdifulle innspill når jeg har stått fast. Jeg vil også få takke respondentene mine som stilte opp til evalueringen, uten dere hadde ikke oppgaven vært mulig å gjennomføre.

Jeg vil også få takke venner, bekjente og medstudenter som har stilt opp underveis. Særlig må jeg få takket Kari Gjerdevik som leste språklig korrektur, og Helge Holm som leste faglig korrektur. Uten deres hjelp ville nok kvaliteten på oppgaven blitt en helt annen. Til slutt vil jeg få takke alle som har kommet med oppmuntrende ord underveis, som har vært tålmodige når det har skrantet med motivasjonen min, og som har kommet med tips til hvordan oppgaven kunne bli bedre. Dere vet hvem dere er!

-Richard Persen

Bergen, april 2008.

Abstrakt

For å kunne gjenbruke ressurser fra tidligere semesters diskusjoner ble det utviklet et system for semantisk annotering av innlegg i diskusjonsforumer. Dette systemet, kalt AnnForum (Annotating Discussion Forums), benytter seg av en konseptuell domenemodell og en nøkkelordgjenkjenner til å klassifiserer innlegg fra tidligere semestres diskusjonsforumer. Disse innleggene kan da gjenbrukes som nye læringsressurser i det nåværende semesteret, noe som fører til mindre duplisering av informasjon, ved at man får muligheten til å benytte seg av det som andre allerede har lært og erfart.

Tidligere systemer har vist at brukerne som regel ikke er motiverte nok til å manuelt tilordne innlegg i henhold til konsepter i domenemodellen. AnnForum tilbyr derfor en dynamisk klassifisering av nye innlegg i henhold til konseptene i domenemodellen. Dette har den fordelen at systemet ikke forstyrrer læringsprosessen til brukerne.

For å teste ut systemet, ble det foretatt seks eksperimentelle evalueringer opp mot studenter som hadde tatt kurs relatert til kunstig intelligens. Disse evalueringene fokuserte på brukervennlighet og brukbarhet sett i en læringssammenheng. Datainnsamlingen ble gjort i form av observasjoner og intervju, og resultatet viste at respondentene likte funksjonaliteten som AnnForum tilbydde, hovedsaklig fordi at de fikk en oversiktlig og enkel tilgang til ressurser som kunne være relevante for den nåværende diskusjonen.

Innholdsfortegnelse

1. Innledning.....	1
1.1 Problemet med tradisjonelle diskusjonsforum.....	2
1.1.1 Eksisterende systemer – hva andre tilbyr.....	4
1.1.2 Løsningsforslag.....	5
1.2 Forskningsspørsmål.....	7
1.2.1 Problemstilling og hypotese.....	7
1.3 Organisering av teksten.....	8
2. Teoretisk rammeverk.....	10
2.1 Læring, samhandling og teknologi.....	10
2.1.1 Datastøttet Læring.....	11
2.1.2 Diskusjonsforumer i læring.....	12
2.2 Teorier og begreper.....	14
2.2.1 Semantikk.....	15
2.2.2 Semantisk Web.....	16
2.2.2.1 Knowledge Representation.....	17
2.2.2.2 Ontologier.....	18
2.2.2.3 Fokusområder.....	19
2.2.2.4 Teknologier.....	20
2.2.2.5 Implikasjoner og eksempler på praktisk bruk.....	21
2.2.3 Annoteringer.....	22
2.2.4 Taksonomier og folksonomier.....	23
2.2.5 Domenemodell.....	23
2.3 Gjenbruk av læringsressurser.....	24
2.3.1 Annotering og klassifisering av innlegg i diskusjonsforumer.....	24
2.3.2 Knowledge Management Systems.....	26
3. Design og utvikling.....	28
3.1 Systemkrav.....	29
3.1.1 Use cases.....	30
3.1.2 Funksjonelle krav – systemet skal.....	30
3.1.3 Ikke-funksjonelle krav – systemet bør.....	32
3.1.4 Scenarioer.....	32
3.2 Arkitektur og hovedkomponenter.....	33
3.2.1 Systemarkitektur.....	34
3.2.2 Domenemodellen.....	35
3.2.3 Databasen.....	36
3.2.4 Klassifiseringsalgoritmen.....	37
3.2.5 Relevante innlegg og presentasjon.....	39
3.3 Integrasjon med FLE3.....	40
3.4 Gjennomgang av systemet.....	41
3.4.1 Brukergrensesnittet for relevante innlegg.....	41
3.4.2 Lærerens annoteringsverktøy.....	45
4. Metode og evaluering.....	48
4.1 Metode.....	48
4.1.1 Eksperimentell evaluering.....	48
4.1.2 Elementer i et eksperiment.....	49
4.2 Gjennomføringen av selve eksperimentet.....	50
4.2.1 Målet med evalueringen.....	50
4.2.2 Utvalg av respondenter.....	51
4.2.3 Datainnsamling – brukeroppgaver og intervju.....	52

4.2.3.1 Brukeroppgaver.....	53
4.2.3.2 Datainnsamling – intervju.....	53
4.3 Analyse og oppsummering av data.....	55
4.3.1 Første evaluering.....	55
4.3.2 Andre evaluering.....	56
4.3.3 Tredje evaluering.....	57
4.3.4 Fjerde evaluering.....	58
4.3.5 Femte evaluering.....	59
4.3.6 Sjette evaluering.....	60
4.3.7 Oppsummering av resultatet.....	61
4.3.7.1 Førsteintrykket av de foreslåtte innleggene.....	61
4.3.7.2 Relevansen for de innleggene som faktisk ble lest.....	61
4.3.7.3 Positive og negative aspekter – forslag til endringer.....	62
4.3.7.4 Generell nytteverdi av systemet.....	63
4.3 Diskusjon av resultatet.....	64
5. Konklusjon.....	66
5.1 Tilbakeblikk på evalueringen.....	66
5.2 Tilbakeblikk på utviklingen.....	67
5.3 Forbedringer og fremtidig arbeid.....	67
6. Litteraturoversikt.....	70
Appendiks.....	74
Appendiks A: Klassediagram.....	74
Appendiks B: Domenemodellen.....	79
Appendiks C: Brukeroppgaver.....	83
Appendiks D: Intervjutranskriberinger.....	85
Respondent 1.....	85
Respondent 2.....	87
Respondent 3.....	89
Respondent 4.....	91
Respondent 5.....	93
Respondent 6.....	95

Figurer

Figur 1.1: Et typisk diskusjonsforum med lite beskrivende titler.....	3
Figur 1.2: FLE3 sin kunnskapsbyggingsmodul.....	6
Figur 2.2: Studentportalen ved UiB.....	14
Figur 3.1: De forskjellige delene i AnnForum.....	28
Figur 3.2: Use case scenarios.....	30
Figur 3.3: Konseptuell oversikt over AnnForum.....	34
Figur 3.4: Diskusjonsforumet til FLE3.....	42
Figur 3.5: Vis relevante innlegg.....	42
Figur 3.6: Brukergrensesnittet for relevante innlegg.....	43
Figur 3.7: Et utvalgt relevant innlegg.....	44
Figur 3.8: Forum View.....	45
Figur 3.9: Topic View.....	46
Figur 3.10: Et utvalgt innlegg i Topic View.....	46

1. Innledning

1. Innledning

Begrepet annotering er sentralt innen semantisk web, og en annotering defineres i dette tilfellet som en kommentar eller et notat som beskriver et objekt eller et dokument (Euzenat, 2002). Innen dagens moderne utdanning blir læringsmateriale stort sett gjort tilgjengelig på Internett, og i en læringskontekst er det ikke bare læringsmateriale og ressurser lagt ut av lærere som kan bli annotert, men også materiale lagt ut av studenter som ledd i deres kollaborative læringsprosess. Blant annet diskusjonsforumer har blitt mye brukt i forskjellige læringskontekster, for eksempel til å stille faglige spørsmål, diskutere prosjektarbeid, samarbeide i forskjellige situasjoner, presentere hypoteser, lenke opp til relevante ressurser og så videre.

Disse diskusjonsforumene kan derfor ved riktig bruk inneholde mye viktig informasjon relatert til læring, og de har derfor et stort potensiale til å kunne øke kunnskapsbyggingen hos studenter. Azouaou, Chen og Desmoulins (2005), samt Helic (2003, 2005) har skrevet en del om dette temaet, og de sier at det blant annet ville vært meget nyttig om man kunne brukt informasjonen i innleggene fra tidligere semesters diskusjoner som nye læringsressurser, slik at nye studenter i inneværende semester kunne dratt fordel av det som andre studenter har lært, erfart og diskutert. På denne måten ville man fått en situasjon der kunnskapsbyggingen bygger videre på allerede oppsamlet kunnskap. Slik det er nå, må man starte hele prosessen på nytt hvert eneste semester, noe som er både tidkrevende og ineffektivt.

Selv om det ligger mye læringsinformasjon tilgjengelig på Internett og i diskusjonsforumer, er disse ressursene vanskelige å finne frem til, samt vanskelige å gjenbruke. Dette henger til dels sammen med at de fleste diskusjonsforumer som eksisterer i dag, er rotete og uoversiktlige, og lærings-diskusjonsforumer er intet unntak (Helic, 2003). Selv om det er den enkle, trådbaserte strukturen til diskusjonsforumer som har gjort de så populære, så er det også den samme enkle strukturen som kan gjøre det til en tidkrevende og vanskelig prosess å finne informasjon og få oversikt over relevante innlegg. Et av de vanligste problemene er at innlegg dyttes lenger og lenger bak i hierarkiet etter hvert som nye innlegg legges til,

1. Innledning

noe som medfører at eldre, ofte relevante innlegg, havner bakerst i listen. Siden brukerne stort sett konsentrerer seg om å lese de innleggene som er lettest tilgjengelige, altså de nyeste innleggene, så vil disse eldre innleggene da ofte bli ignorert. Et annet vanlig problem er at det kan være veldig vanskelig å se om et innlegg er relevant eller ikke ut fra en lite beskrivende tittel, noe som har sammenheng med at de fleste brukerne ofte ikke er motiverte nok til å bruke tid på å formulere en beskrivende tittel. Som oftest ender de bare opp med å trykke på «Reply» og bruker dermed den autogenerated tittelen som kommer opp, for eksempel «RE: Hva er en annotering» eller «RE: Hjelp!!!».

Alle de ovennevnte problemene, kombinert med søkefunksjoner som ofte ikke er gode nok, har den konsekvensen at det vil gå fortere for brukerne å legge inn et nytt innlegg, fremfor å måtte bruke mye tid på å lete etter tidligere innlegg som omhandler det ønskede temaet. Dette er en lite effektiv prosess da det fører til mye duplisering av allerede eksisterende informasjon og spørsmål som tar opp samme type tema/spørsmål. Eldre og ofte relevante innlegg blir av denne grunnen ofte oversett, noe som fører til at mye verdifull informasjon som kunne vært brukt i læringsprosessen går tapt. Problemet består derfor både i informasjon fremfinning (Information Retrieval), kunnskapshåndtering (Knowledge Management) og i å gjenbruke denne informasjonen på en fornuftig måte ved hjelp av semantisk web.

1.1 Problemet med tradisjonelle diskusjonsforum

Som nevnt så er det den enkle strukturen til diskusjonsforumer som har gjort disse så populære, men det er også den samme strukturen som har gjort det så problematisk å finne frem til ønsket informasjon på en enkel og intuitiv måte. Mye av dette henger sammen med det som plager Internett som helhet, nemlig en grunnleggende mangel på en konseptuell struktur. Informasjon er ikke klassifisert på en måte som gjør at maskiner kan tolke den (Antoniou, 2004: 1-2). Noen av de mest sentrale problemene ved tradisjonelle diskusjonsforum er listet opp nedenfor:

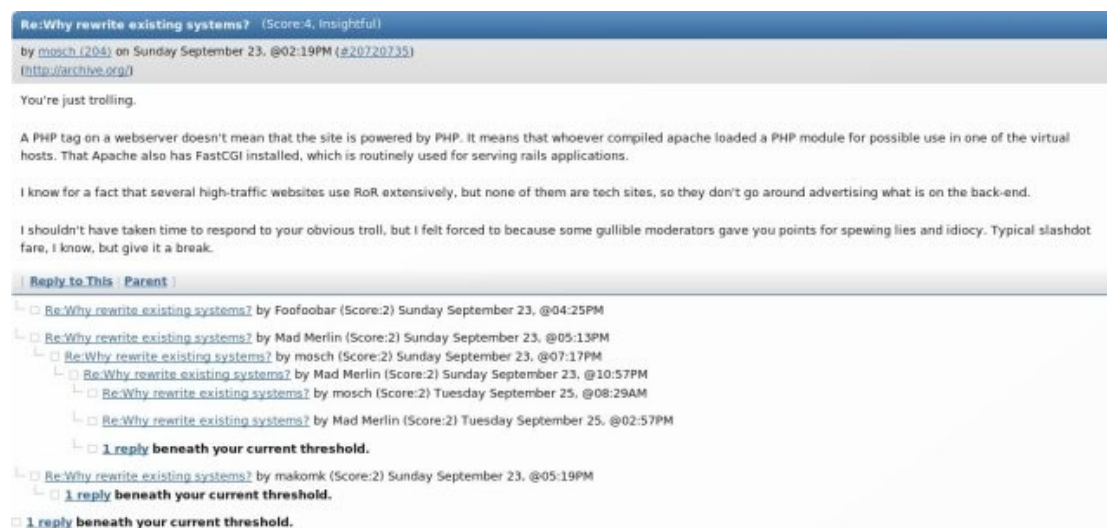
1. De har for dårlig støtte for informasjon fremfinning (kun rene tekstsøk), noe som gjør det vanskelig å finne relevante innlegg og ressurser tilknyttet et

1. Innledning

spesifikt tema.

2. De blir fort uoversiktlige etter hvert som antall innlegg i forumet vokser.
3. Eldre, og ofte relevante, innlegg blir dyttet lenger og lenger bak i køen, og har derfor en tendens til å bli oversett og glemt.
4. Titlene for de enkelte innleggene er ofte lite beskrivende og sier lite om innleggets innhold. Det vanligste er at tittelen er av formen «RE: <tittel på hovedinnlegg>» (se figur 1.1).

Det må likevel nevnes at i noen tilfeller, for eksempel der forumet er forholdsvis lite, eller der diskusjonen handler om et flyktig tema, så er ikke disse problemene grunn til en altfor stor bekymring; for eksempel i en diskusjon som omhandler en nyhetsartikkel. Men i situasjoner der diskusjonsforumene kan inneholde mye nyttig/generell informasjon, og som mange kan ha interesse av å finne frem til (for eksempel innen læring eller forskning), så kan det vært interessant å ta vare på denne informasjonen. De skisserte problemene blir derfor i disse tilfellene litt mer alvorlige fordi de kompliserer prosessen med å lett finne frem til den aktuelle informasjonen.



Figur 1.1: Et typisk diskusjonsforum med lite beskrivende titler

Et system som bruker teori og teknologi relatert til Semantisk Web for å kategorisere innleggene, hovedsaklig gjennom semantisk annotering, og som i tillegg hjelper

1. Innledning

brukeren til å finne frem til relevante og ønskede innlegg på en enkel måte, kan derfor være et steg i riktig retning. Dette kan drastisk forenkle informasjonsfremfinningen, og potensialet med dette er at brukerne slipper å bruke så mye tid på å lete etter ønsket informasjon. I tillegg kan kvaliteten på informasjonen i diskusjonsforumene bli bedre, hovedsaklig som følge av at det blir mindre «støy» i form av duplikate innlegg.

1.1.1 Eksisterende systemer – hva andre tilbyr

Innen informasjonsfremfinning har det blitt gjort mye arbeid når det kommer til å finne (lærings)informasjon på Internett, blant annet også til å lage systemer som finner frem til og gjenbraker informasjon i diskusjonsforumer. Men det er fortsatt flere problemer relatert til dette. Blant annet har Helic (2003, 2005) utviklet et verktøy som lar brukerne manuelt tilordne et innlegg til et konsept i en predefinert domenemodell, der den konseptuelle domenemodellen er uavhengig av diskusjonsforumet. Men denne fremgangsmåten har vist seg å ikke være optimalt av to årsaker: For det første er studentene ofte ikke motiverte nok til å manuelt tilordne et innlegg til et spesifikt konsept, noe som har sammenheng med at det medfører mer arbeid for studenten. For det andre kan det oppstå situasjoner der et innlegg kan tilordnes mer enn ett konsept.

En annen fremgangsmåte som har vært forsøkt av Craven (1998) er å gi diskusjonsforumet en predefinert struktur basert på domenets konseptuelle skjema, men dette er heller ikke optimalt da det begrenser brukerens frihet når innlegg skal postes. For eksempel kan det oppstå en situasjon der innlegget ikke passer inn i domenemodellen. I tillegg har en slik fremgangsmåte de samme ulempene som Helic sitt system: Brukerne er ofte ikke motiverte nok til å utføre det ekstra arbeidet med å manuelt tilordne informasjon til konsepter, samt at innlegget kan passe inn under mer enn ett konsept.

Vi ser med andre ord at et system som er for komplisert gjennom å sette for mange begrensninger, eller som gjør at brukeren får en ekstra arbeidsbyrde, vil være mer til hinder enn til nytte og vil føre til at brukerne fort mister motivasjonen og gir opp. Det er derfor nødvendig at et slikt system gir brukeren tilstrekkelig frihet når det kommer

1. Innledning

til det å poste nye innlegg, og som i tillegg hjelper til med å automatisk tilordne et innlegg til eventuelle relevante konsepter.

Om disse kravene oppfylles, kan dette medføre at:

1. Det blir lettere å finne frem til relevante innlegg.
2. Søkefunksjonene blir bedre ved at mengden relevant informasjon som hentes frem blir mer presis: *Recall* og *precision* øker (Yates *et al.*, s. 75, 1999).
3. Det blir mindre «overhead», altså at samme type spørsmål gjentas gang på gang.
4. Man får en oversikt over relevante ressurser tilknyttet de forskjellige innleggene.

Men det aller viktigste er den økte muligheten for *gjenbruk av informasjon fra tidligere diskusjoner*. Se kapittel 2 for flere detaljer om eksisterende systemer, samt en diskusjon og en kort analyse av systemene til Helic og Craven.

1.1.2 Løsningsforslag

AnnForum (*Annotating discussion Forum*) er en prototype som først ble utviklet ved Universitetet i Bergen (Chen, 2004), og er et eksempel på et annoteringsverktøy som tilbyr en semi-automatisk, semantisk annotering av innlegg i diskusjonsforumer. Systemet er integrert i gruppevaresystemet FLE3, og har tidligere blitt brukt i kurset INFO281: Kunstig Intelligens. AnnForum ble imidlertid ikke testet med tanke på brukbarhet, noe som har sammenheng med at prototypen var uferdig.

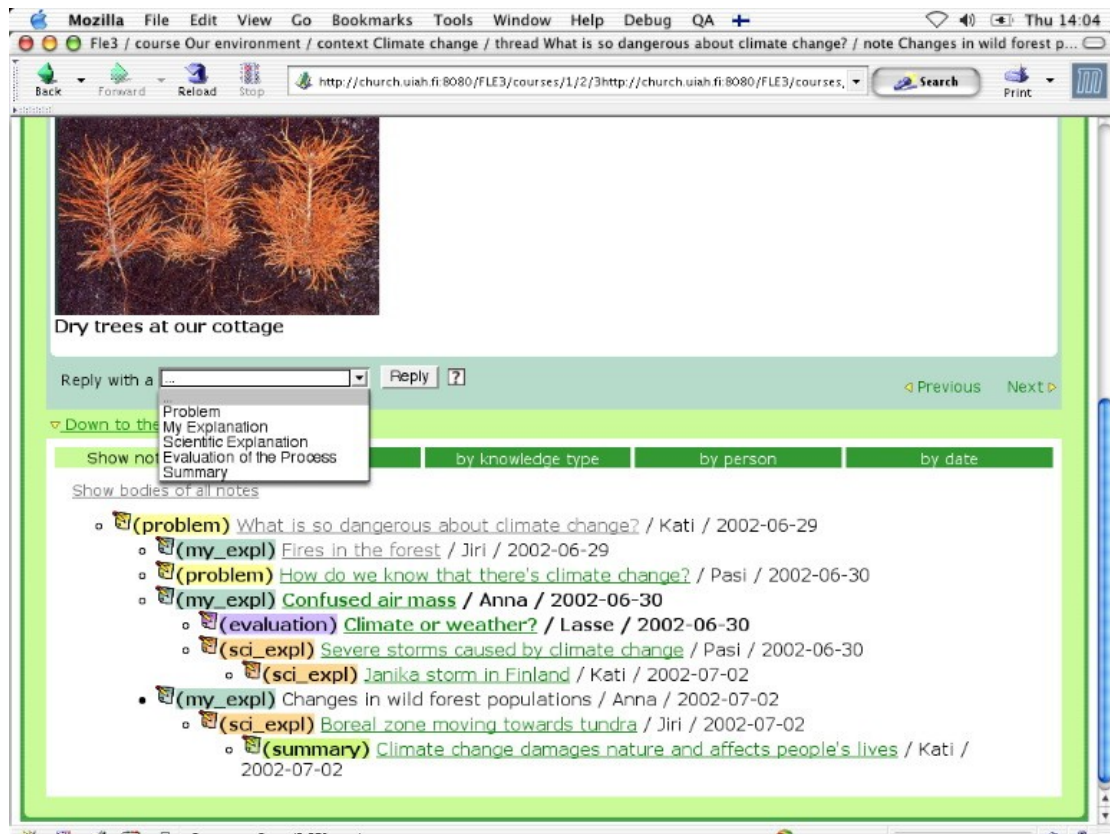
FLE3, eller *Future Learning Environment*¹, er utviklet ved Universitetet i Helsinki, og er et gruppevaresystem som benyttes hovedsaklig i forbindelse med datastøttet samarbeid. Systemet er utviklet for å støtte opp om kunnskapsbygging sett i en gruppesammenheng, og har flere moduler som støtter opp om dette. Målet er å sette søkelys på kunnskapsbygging, problemløsning og forskningsmessige metoder i en

¹ <http://fle3.uiah.fi/>

1. Innledning

læringsammenheng. *WebTop* er et personlig arbeidsområde hvor lærere og studenter kan lagre ressurser relatert til studiene, som for eksempel dokumenter og bilder.

Knowledge Building er et diskusjonsforum som gjør det mulig å organisere innleggene i henhold til fem forskjellige kunnskapstyper: *Problem*, *My Explanation*, *Scientific Explanation*, *Evaluation of the Process* og *Summary*. *Jamming* er en modul som brukes som et felles arbeidsområde for digitale ressurser, som for eksempel bilder, video etc. Kunnskapsbyggingsmodulen har blitt mye brukt i INFO281, og inneholder derfor flere hundre innlegg som har blitt brukt som nye læringsressurser. Figur 1.2 viser et skjermbilde av kunnskapsbyggingsmodulen til FLE3.



Figur 1.2: FLE3 sin kunnskapsbyggingsmodul

For å gjenbruke innleggene som ligger i diskusjonsforumet til FLE3 har det manuelt blitt konstruert en konseptuell domenemodell, eller en ontologi, i form av et XML Topic Map (XTM) som beskriver konsepter innenfor kunstig intelligens-domenet. Det har også blitt utviklet en klassifiseringsalgoritme som kategoriserer innleggene i diskusjonsforumet i henhold til disse konseptene. Den ovennevnte

1. Innledning

klassifiseringsalgoritmen og domenemodellen, sammen med et brukergrensesnitt hvor studentene kan se relevante innlegg samt hvilke emner innlegget er relatert til, gjør det dermed mulig for systemet å presentere relevant informasjon fra tidligere semesters diskusjoner. Studentene får dermed en mulighet til å få en oversikt over innlegg som andre studenter har skrevet tidligere, og som kan være relevante for den nåværende diskusjonen.

Den automatiske klassifiseringen av nye innlegg løser blant annet også problemet med at studentene manuelt må foreta tilordningen, i tillegg til at et innlegg nå kan kategoriseres under mer enn et relevant konsept. Etter at studentene har lest et relevant innlegg har de dessuten mulighet for å avgi en stemme for om de synes et foreslått innlegg var relevant eller ikke. Denne stemmen vil påvirke rangeringen av de relevante innleggene for neste bruker da urelevante innlegg havner langt nede på listen, mens relevante innlegg vil havne langt oppe. Man kan altså si at systemet «lærer» ved å ta imot tilbakemelding fra brukerne.

Under utformingen av systemet har det vært vektlagt at selve annoteringsprosessen skal gi studenten tilstrekkelig frihet når han poster innlegg, i tillegg til at prosessen heller ikke skal forstyrre læringsprosessen. Ved siden av grensesnittet som viser relevante innlegg til studentene har det også blitt utviklet et administrativt verktøy for læreren. Dette verktøyet gir læreren mulighet til å opprette nye domenemodeller, få en oversikt over og redigere eksisterende modeller, åpne tidligere diskusjonsforumer samt redigere annoteringene for de klassifiserte innleggene. Se kapittel 3 for mer informasjon om lærerens administreringsverktøy.

1.2 Forskningsspørsmål

Denne delen tar for seg forskningsspørsmålet, problemstillingen og hypotesen min.

1.2.1 Problemstilling og hypotese

Forskingsspørsmålet og problemstillingen min har gått ut på å se om semantisk annotering av innleggene i diskusjonsforumer kan hjelpe studenter i deres kollaborative læringsprosess, ved at læringsressurser fra tidligere semesters

1. Innledning

diskusjoner gjenbrukes som nye læringsressurser i den nåværende diskusjonen.

Fremgangsmåten for å få svar på dette spørsmålet har vært todelt:

1. Jeg har videreutviklet et system for semi-automatisk, semantisk annotering av innleggene i diskusjonsforum.
2. Jeg har evaluert dette systemet ved å se på brukbarheten og brukervennligheten når det kommer til annoteringen av *ikke-prefabrikkert læringsmateriale*.

Med andre ord har oppgaven min vært å implementere systemet AnnForum (*Annotating discussion Forum*) som blant annet inkluderer et administrativt verktøy, en klassifiseringsalgoritme, i tillegg til et brukergrensesnitt for relevante innlegg. Deretter har jeg gjennomført en funksjonell brukertest med tanke på brukbarhet og brukervennlighet av systemet. Et viktig moment har vært at annoteringsprosessen ikke skal forstyrre læringsprosessen for studentene; de skal selv kunne velge om de vil bruke funksjonen eller ikke.

1.3 Organisering av teksten

Ved siden av innledningen, konklusjonen, innholdsfortegnelsen og appendikset så er denne oppgaven delt opp i fire andre kapitler.

Kapittel 2 tar for seg aktuelle teorier og relevant litteratur som har vært viktige i forbindelse med utformingen, utviklingen og evalueringen av AnnForum. Det som blir tatt opp her, fungerer som et utgangspunkt for argumentasjonen i resten av oppgaven i tillegg til å begrunne de teoretiske og designmessige valgene som har blitt gjort underveis. Viktige bidrag gjort av andre aktører i relasjon til problemstillingen, som for eksempel eksisterende systemer for annotering av læringsressurser, vil også bli diskutert.

Kapittel 3 beskriver oppbyggingen av AnnForum ut fra et teknisk perspektiv. Dette inkluderer blant annet de formelle kravene til systemet, systemarkitekturen, sentrale komponenter, eksempelkode samt hvordan systemet har blitt integrert med FLE3. I

1. Innledning

tillegg gis det en gjennomgang av de mest sentrale brukergrensesnittene til AnnForum.

Kapittel 4 evaluerer AnnForum i forhold til problemstillingen og er delt opp i tre deler. Den første delen tar for seg selve evalueringsmetoden, den andre delen beskriver hvordan evalueringen ble gjennomført i praksis med tanke på utvalg av respondentene, intervjuprosessen og lignende, mens den tredje delen analyserer innsamlet data og oppsummerer resultatet av evalueringen.

Kapittelet fem oppsummerer teksten i form av en konklusjon, samt peker ut fremtidig arbeid i forbindelse med AnnForum. I tillegg diskuteres resultatet fra den eksperimentelle evalueringen i korte trekk, med den hensikten å se på om resultatet var som forventet i forhold til problemstillingen og hypotesen. Var det spesielle momenter som skilte seg ut, eventuelt var som forventet, og hvorfor respondentene svarte som de gjorde.

2. Teoretisk rammeverk

Dette kapitlet presenterer sentrale teorier og relevant litteratur som har vært aktuelle i forbindelse med problemstillingen skissert i forrige kapittel. Noen av de viktigste teknologiene og konseptene beskrives også, i tillegg til at deler av kapitlet tar for seg relaterte systemer som har likheter med AnnForum. Kapitlet er delt opp i tre deler.

Første del omhandler bruk av teknologi som et hjelpemiddel sett i en læringskontekst, og da særlig med tanke på diskusjonsforumer. Dette innebærer blant annet en introduksjon til datastøttet læring og relaterte områder. Hvordan kan datamaskiner benyttes i en klasseroms-/studiesammenheng når det kommer til samhandling og kommunikasjon mellom elever, studenter og lærere?

Andre del tar for seg de teoriene og begrepene som har vært relevante for designet av AnnForum, samt hva som er de potensielle fordelene ved å annotere og gjenbruke læringsressurser. Noen av de sentrale teoriene og begrepene som vil bli tatt opp underveis, er blant annet: semantikk, informasjonsfremfinning, semantisk web, annoteringer, ontologier og datastøttet samarbeid. En kort introduksjon til aktuelle teknologier som blant annet XTM Topic Maps vil også bli gitt, i tillegg til en beskrivelse av systemene til Helic og Craven.

Tredje del beskriver prosessen med å gjenbruke læringsressurser, og da særlig med søkelys på annotering og klassifisering av informasjon i diskusjonsforumer. Denne delen gir også en liten introduksjon til kunnskapshåndtering, noen eksempler på slike systemer, og hvordan domenekart og ontologier kan forbedre disse systemene.

2.1 Læring, samhandling og teknologi

De siste femten til tjue årene har datateknologien utviklet seg med rekordfart, noe som har medført en stor interesse for å bruke datamaskiner i forbindelse med utdanning (Koschmann, 2006). Helt fra barnehage til universitetsnivå har man sett på de potensielle fordelene ved å innføre bruk av datamaskiner i forskjellige læringssammenhenger. Dette innebærer ikke bare bruk av datamaskiner i vanlig

2. Teoretisk rammeverk

klasseromsundervisning, men kan også inkludere undervisning gjort over Internett ved hjelp av web-baserte læringssystemer (Web-Based Training, WBT). Men utdanning og læring er en komplisert prosess som involverer mange forskjellige aspekter, herunder psykologi, sosiologi, språk, problemløsning, sosial samhandling, forskning, metodologi og data-teoretiske fag. Det er derfor ingen triviell sak å gjøre endringer i denne prosessen, og selv om gevinstene ved riktig bruk av teknologi i en undervisningssammenheng er mange, så er det også mange fallgruver. For web-basert læring er for eksempel én av fordelene at informasjon er lett tilgjengelig, mens en ulempe er en mangel på ansikt-til-ansikt-interaksjon mellom elever og lærere. Det er derfor viktig at man tenker nøye gjennom fordelene og ulempene ved bruk av teknologi i en læringssammenheng. De ovennevnte momentene har dermed gitt en økt interesse for å forske på dette feltet når det kommer til å bruke ny teknologi relatert til utdanning og læring.

2.1.1 Datastøttet Læring

Datastøttet læring, eller Computer Supported Collaborative Learning (CSCL), er et teoretisk felt relatert til forskning, utdanning, læring og metode og ble først karakterisert av Timothy Koschmann på 1990-tallet (Koschmann, 1996). Koschmann gir en kort forklaring av begrepet og sier at man innenfor datastøttet samarbeid er opptatt av å studere «[...] hvordan mennesker kan lære sammen ved hjelp av datamaskiner [...]» (Koschmann, 2006:1-2). Lipponen bygger videre på dette og sier blant annet at man innen datastøttet læring er interessert i å se på hvordan teknologi og datamaskiner kan brukes til å støtte opp om samhandling mellom brukere og grupper av brukere i ulike, kollaborative læringssituasjoner (Lipponen, 2002). Videre sier Lipponen at feltet bygger på teori fra områder som sosiologi, språk, pedagogikk, informasjonsvitenskap og psykologi, og at det av denne grunnen har vært problematisk å komme til en felles forståelse av hva datastøttet læring egentlig betyr, og da særlig begrepet om samhandling («collaboration»). Dette henger blant annet sammen med at feltet har sitt utspring fra datastøttet samarbeid, og at de to områdene har mange likhetstrekk når det kommer til aspekter som sosial interaksjon ved hjelp av datamaskiner.

2. Teoretisk rammeverk

Samhandling kan ifølge Lipponen i den bredeste forstand defineres som «en aktivitet eller handling der to eller flere personer utfører noe sammen» (Lipponen, 2002:2).

Dillenbourg (1999) forsøker i sin artikkel å forklare hva som menes med samhandling relatert til læring (kollaborativ læring). Kollaborativ læring, i sin bredeste form, er en situasjon der to eller flere personer forsøker å lære noe sammen. Denne vage definisjonen gir tre elementer som man kan bygge videre på: Hvor mange som er involvert, selve prosessen med å lære noe, og til slutt det å utføre læringen sammen, altså i samhandling. Men det å gi en nøyaktig definisjon av begrepet er vanskelig da ordet «kollaborativ» brukes i veldig mange forskjellige sammenhenger. Men grovt sett kan man si at man fokuserer på å løse oppgaver i fellesskap. Med andre ord har deltakerne en felles forståelse av et problem relatert til læring, og de forsøker å løse dette problemet ved at alle deltakerne er aktive. Sosial interaksjon aktørene imellom er en sentral del av kollaborativ læring, og man er ikke fullt så interessert i selve teknikkene for kommunikasjon, som hva det er som blir kommunisert.

Datastøttet læring ser ikke bare på hvordan aktører kan samhandle ansikt til ansikt, men også hvordan læring foregår over avstander, som for eksempel ved bruk av Internett (Koschmann, 2006). Men til forskjell fra e-læring, som i mange tilfeller kun tar for seg digitaliseringen og publisering av læringsinformasjon på Internett, så er man innen datastøttet samarbeid veldig opptatt av at både lærere og studenter skal være aktive i selve læringsprosessen. Et eksempel er et kurs som blir tatt over Internett, og som bruker et system for datastøttet læring som lar studentene diskutere og finne frem til informasjon i en kollaborativ kontekst. Brukerne kan da bruke dette systemet til lettere å løse problemer, utveksle ideer samt dele informasjon og læringsressurser uten å måtte oppholde seg i samme rom. Målet er med andre ord at elevene aktivt skal være med på å konstruere og oppsøke kunnskap, så vel som ta del i denne prosessen i samhandling med sine medstudenter. Det krever altså betraktelig med planlegging og arbeid for å tilrettelegge undervisning til datastøttet læring.

2.1.2 Diskusjonsforumer i læring

Et diskusjonsforum, også kalt et Internett-forum, er et populært verktøy som brukes i

2. Teoretisk rammeverk

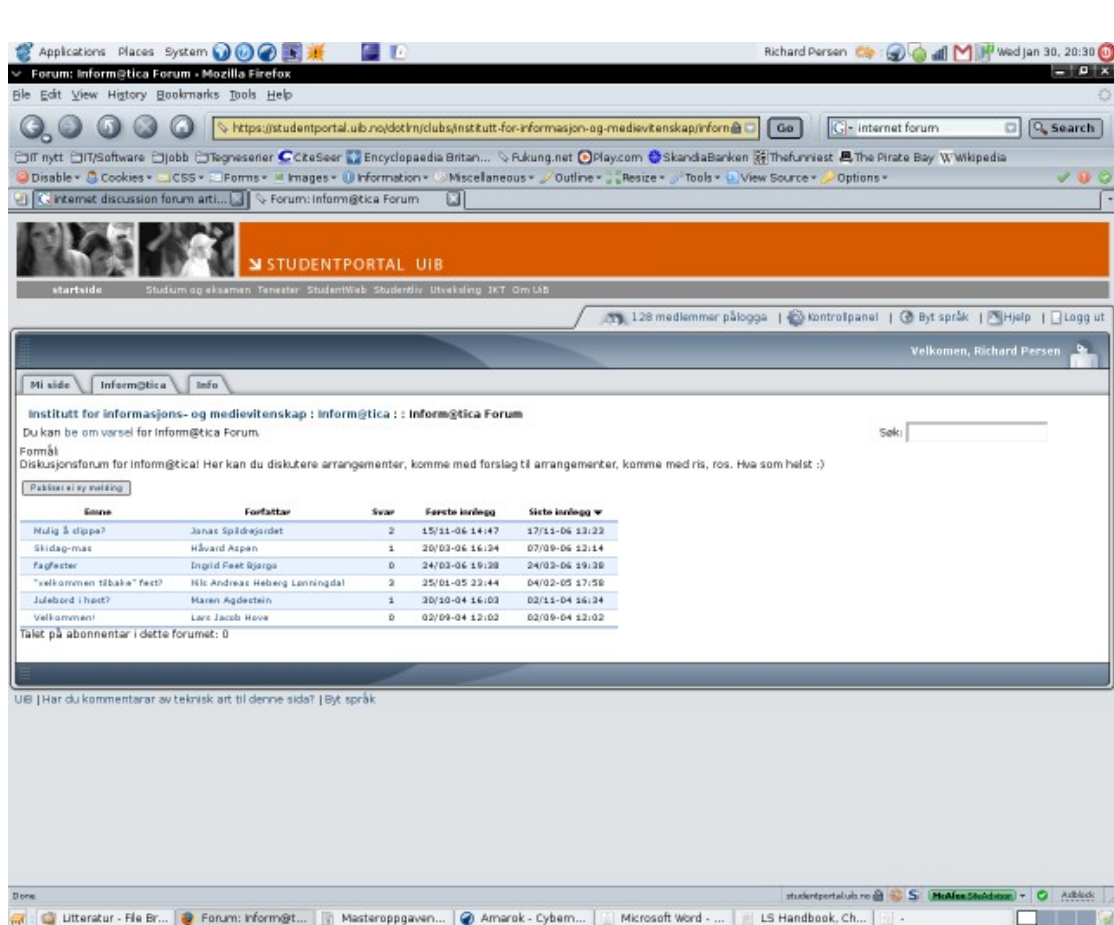
situasjoner der man ønsker å kunne holde en diskusjon og debattere omkring et spesifikt tema (Helic, 2005). Dette kan for eksempel være for å diskutere vitenskaplige artikler, komme med forslag til forskjellige problemstillinger eller rett og slett relatert til en hobby. Diskusjonsforumer blir derfor hovedsakelig brukt til å kommunisere med andre. Som oftest har man tilgang til dette forumet via Internett, og temaene som diskuteres, kan være innenfor forskning, hobby, utdanning, nyheter og så videre. Hovedgrunnen til at diskusjonsforumene har blitt så populære, er at de er enkle og intuitive av natur. Man blar gjennom innlegg via en trådbasert struktur og kan enkelt svare på innlegg og legge til nye innlegg ved å fylle ut skjema.

Et diskusjonsforum er et asynkront system, noe som innebærer at brukerne ikke trenger å være pålogget samtidig for å holde en diskusjon gående, men at de heller logger seg på når det måtte passe dem best. Diskusjonsforum kan ta mange forskjellige former, men felles for de fleste er at de lar brukerne legge inn brukergenerert innhold i form av egne poste/innlegg. Noen forumer lar også brukerne legge ved bilder eller dokumenter. Disse postene er som oftest organisert i en kronologisk, trådbasert eller hierarkisk struktur og ligger tilgjengelig slik at andre brukere kan gå inn og lese hva som har blitt diskutert, samt legge til et svar eller motargument om de skulle ønske det.

De fleste forumene har en administrator samt én eller flere moderatører i tillegg til vanlige brukere. Administratoren har det overordnede ansvaret for hele diskusjonsforumet og har som regel full kontroll. Dette kan for eksempel være sletting og redigering av poster, blokkering eller oppretting av brukere, åpning av nye tråder og så videre. En moderatør er en vanlig bruker som har fått noen av administratorens privilegier, slik som for eksempel sletting og redigering av innlegg. De aller fleste diskusjonsforumene har også en del uformelle normer og regler som det forventes at deltakerne skal overholde, og moderatorens hovedoppgave er i denne sammenhengen å påse at poster som legges inn, er i tråd med forumets retningslinjer, samt å advare eller blokkere brukere som bryter disse. Eksempler på diskusjonsforumer er studenportalen til Universitetet i Bergen (figur 2.1), FLE3 sin kunnskapsbyggingsdel (figur 1.2) og diskusjonsforumet for artikler på Slashdot² sine sider (figur 1.1).

² <http://slashdot.org/>

2. Teoretisk rammeverk



Figur 2.1: Studenportalen ved UiB

Studentportalen kan også muligens defineres som et gruppevaresystem relatert til datastøttet læring, hovedsaklig fordi studentene kan bruke funksjonaliteten i portalen til å ha kontakten med forelesere og andre studenter, så vel som til å lagre læringsmateriale i form av personlige dokumenter og innleveringer.

2.2 Teorier og begreper

Det har lenge vært en kjent sak at det å søke etter informasjon på Internett til tider kan være en møysommelig og tidkrevende oppgave, og en av hovedårsakene som oftest blir dratt frem i denne sammenhengen, er at Internett mangler en konseptuell struktur (Antoniou, 2004:1–2). Med en konseptuell struktur menes det her at informasjonen i hovedsak ikke kategoriseres eller sorteres på noen måte, men at den kun eksisterer i sin reneste form, for eksempel en side eller en artikkel på et nettsted med til dels store mengder ustrukturert tekst. Selv ved riktig bruk av søkemotorer som for eksempel

2. Teoretisk rammeverk

Yahoo! eller Google vil mye tid likevel bli brukt til å lete gjennom urelevante sider før man finner frem til ønsket informasjon. Tim Berners-Lee, mannen som oppfant World Wide Web, har sagt at hovedgrunnen til dette er at data som per i dag ligger på nettet, hovedsaklig er designet, strukturert og organisert med tanke på menneskelig prosessering (Berners-Lee, 2004).

Denne organiseringen gjør det vanskelig for søkemotorer og andre software-baserte agenter å vite hvilken type informasjon det er brukeren vil ha tak i. Av denne grunnen kan man si at det oppstår et skille mellom brukerens informasjonsbehov og måten software-agenten tolker dette informasjonsbehovet på. Dette skillet blir ofte referert til som *The Semantic Gap* (Smoulders, 2000:5). Der et menneske kan se på et dokument/bilde og fort få en oversikt over innholdet, vil en maskin i samme situasjonen, for eksempel en søkemotor, kun se en sekvens med tegn. Software-baserte agenter tolker med andre ord ikke innholdet (semantikken) i web-dokumenter på noen måte når de søker etter informasjon, de prosesserer kun tekstens tegn ved å se på antall forekomster av ord og hvor i dokumentet ordene forekommer. Rangeringen foretas deretter på bakgrunn av den tekstbaserte analysen. Det at innholdet på nettet ikke er gjort tilgjengelig på et maskin-tilgjengelig språk, gjør det dermed vanskelig, om ikke umulig, for maskiner å bruke automatiske metoder for å finne og tolke data basert på semantikk. Dette er som allerede nevnt også et stort problem i diskusjonsforumer.

2.2.1 Semantikk

For å kunne snakke om semantisk annotering av læringsmateriale er det først nødvendig å gi en forklaring på hva semantikk er. Semantikk kan i den originale, filosofiske sammenhengen sies å være studiet av mening, og da mer spesifikt hvordan ulike ord og termer brukes for å representere informasjon i forskjellige kontekster (Merriam-Webster Online Dictionary³). Særlig innen lingvistikk har semantikk blitt hyppig brukt når det kommer til å se på mening og syntaks i språk. Semantikk brukes også i forbindelse med semiotikk der man ser hvordan tegn og symboler blir brukt i forskjellige sammenhenger til å symbolisere og tolke mening. En annen variasjon er logisk semantikk, også kalt modell-teori, eller formell semantikk (Stanford

3 <http://www.merriam-webster.com/dictionary/semantics>

2. Teoretisk rammeverk

Encyclopedia of Philosophy, 2005). Innen modell-teori er man opptatt av å se på den semantiske relasjonen mellom forskjellige uttrykk i formelle språk og hvordan disse formelle språkene kan brukes i forbindelse med blant annet klassifiseringer.

Semantiske teorier, som for eksempel modell-teori, blir brukt innenfor en rekke forskjellige felt. I tillegg til lingvistikk er semantikk mye brukt innenfor matematikk og informasjonsvitenskap, og da særlig i relasjon til semantisk web.

2.2.2 Semantisk Web

Begrepet *semantisk web* ble opprinnelig utformet av Word Wide Web (www) sin grunnlegger, Tim Berners-Lee, og var hans opprinnelige visjon av hvordan nettet burde vært bygd opp (Berners-Lee, 2000). Målet med et slikt nettverk var å forbedre og effektivisere prosessen med å søke etter informasjon, blant annet ved å utvikle nye teknologier og teknikker, slik at maskiner lettere kunne tolke data og innhold på Word Wide Web. *Semantisk web* er nå et prosjekt drevet av The World Wide Web Consortium⁴, under ledelse av Tim Berners-Lee, hvor hovedmålet er å endre på noen av problemene nevnt innledningsvis. Dette innebærer blant annet å kunne representere informasjon i en maskin-prosesserbar form, slik at maskiner lettere kan tolke den informasjonen som ligger der (Berners-Lee, 1998). Dette vil man få ved å utvikle nye teknologier og konsepter for representasjon av data, og to av de mest sentrale teknologiene i denne sammenhengen er Resource Descriptive Framework (RDF) og Ontology Web Language (OWL).

I et intervju fra 2007 snakker Berners-Lee om et *nettverk av data*, som i motsetning til et *nettverk av dokumenter* (designet for mennesker), kan ses på som en slags global database. Denne globale databasen vil i mye større grad legge vekt på å uttrykke meningen, altså semantikken, til informasjon enn det som per i dag er tilfellet (Berners-Lee, 2007). Maskiner kan da bruke avanserte teknikker og åpne standarder til å hente ut og tolke semantikken til informasjonen som ligger på nettet. Disse åpne standardene vil i en mye større grad gjøre det mulig å skille data og semantikk fra applikasjonslogikk og dokumentinnhold, noe som igjen vil gjøre det enklere å beskrive og lenke sammen data på en måte som gjør at maskiner bedre kan «integrere,

⁴ <http://www.w3.org/>

2. Teoretisk rammeverk

aksessere, gjenbruke og dele informasjon på tvers av applikasjoner, forretningsområder og landegrenser»⁵.

Det å gjenbruke og dele informasjon handler om å kunne finne alternative måter å representere kunnskap på, og da mer spesifikt hvordan man kan få datamaskiner og programmer til å forstå og behandle informasjon som ellers bare ville vært tilgjengelig for mennesker. Denne problemstillingen har lenge vært sentral innenfor et område av kunstig intelligens som heter Knowledge Representation.

2.2.2.1 Knowledge Representation

Innenfor Knowledge Representation er man opptatt av å se på hvordan man kan representere kunnskap om den virkelige verden, på en slik måte som gjør at maskiner kan tolke denne informasjonen (Nilsson, 1998:301–316). Dette krever i de fleste tilfeller at representasjonene man bruker, er formaliserte og standardiserte, slik at aktørene kan ha et felles utgangspunkt. En slik formalisering innebærer blant annet at man må være enige om de overordnede regler, konsepter og terminologier for det aktuelle området, noe som igjen innebærer at man har en sentralisert struktur. Når man har etablerte standarder, kan aktørene bruke automatiske teknikker til å behandle, hente ut og til og med trekke konklusjoner ut fra data på samme måte som et menneske ville ha gjort det. Dessverre har nytteverdien av Knowledge Representation så langt vært begrenset, og en av hovedårsakene som oftest blir trukket frem, er den ovennevnte sentraliserte strukturen.

Den enorme mengden av tilgjengelig informasjon gjør at det vil være en ekstremt tidkrevende prosess å sette opp en formell representasjon av selv et ganske enkelt område. Et annet problem er hvordan man skal beskrive ulike konsepter fordi et konsept kan ha forskjellige meninger i forskjellige situasjoner. Så selv om en sentral autoritet har vært sett på som nødvendig for at systemet i det hele tatt skal kunne fungere, så fører også denne samme strukturen til at det fort blir veldig vanskelig å administrere ting etter hvert som systemet vokser (Berners-Lee, 2001). Man har rett og slett ikke funnet en enkel løsning på hvordan man kan representere menneskelig

⁵ <http://www.w3.org/2001/sw/>

2. Teoretisk rammeverk

kunnskap innenfor kunstig intelligens.

Ved utformingen av semantisk web var det derfor nødvendig med en litt annen fremgangsmåte (Berners-Lee, 2001). Den enorme mengden informasjon som finnes tilgjengelig på Internett, gjør det bortimot umulig å finne en entydig beskrivelse av ulike konsepter. Mennesker definerer konsepter og tolker informasjon på forskjellig måte ut fra situasjonen de befinner seg i, og dette fører til at man ikke kan ha en ovenfra-og-ned-struktur hvor alt er hundre prosent konsist. Der hvor man innen kunstig intelligens og tradisjonell Knowledge Representation har vært nødt til å ha faste, veldefinerte og rigide regler i tillegg til en sentralisert struktur, har man innen semantisk web sagt at usikkerhet og tvetydighet er noe som er nødt til å kunne forekomme. Med så store mengder informasjon, og med så mange brukere som det er på Internett, så må man kunne gi rom for at forskjellige tolkninger kan forekomme.

2.2.2.2 Ontologier

Ontologi som begrep har sin opprinnelse i filosofien og blir definert som «teorien eller studiet av eksistens.» (Gruber, 1993:1). Ontologier henger sammen med den delen av metafysikk hvor man er opptatt med å identifisere og beskrive, med så abstrakte termer som mulig, objekter som eksisterer i den virkelige verden. Man snakker altså om et filosofisk system lagd for å kunne representere ulike tolkninger av faktiske objekter (Guarino, 1998). Gruber har skrevet en del om ontologier i forbindelse med informasjonsvitenskap og gir blant annet den følgende definisjonen: «En ontologi er en eksplisitt spesifisering av en konseptualisering» (Gruber, 1993:1). En konseptualisering blir i denne sammenhengen beskrevet som en forenklet og abstrakt representasjon av den virkelige verden og kan bestå av hierarki av klasser og objekter, samt relasjonene mellom disse objektene. En ontologi består som regel derfor av et avgrenset antall konsepter og relasjoner.

Ontologier har vist seg å være nyttige i situasjoner der man forsøker å kategorisere og klassifisere forskjellige typer entiter og objekter, samt relasjonene mellom disse. I denne sammenhengen snakker man mer om *én ontologi*, hvor det menes at man formelt vil beskrive et utvalgt domene. Ontologier har også vist seg å være nyttige når

2. Teoretisk rammeverk

det kommer til å organisere navigasjonen på nettsider, og en ontologi blir da som oftest i form av en datamodell som beskriver konseptene for det utvalgte domenet, samt relasjonene mellom disse konseptene (Antoniou, 2004). Alle konseptene innenfor et spesifikt domene kalles for et vokabular og blir brukt som et felles utgangspunkt for aktørene. Eksempler på praktisk bruk kan være en oversikt over dyreriket (pattedyr, fugler, fisker osv.) eller en oversikt over alle personer på et universitet (lærere, studenter osv.) samt relasjonene mellom disse. Også innen legevitenenskapen har man gjort bruk av ontologier til å klassifisere sykdommer, lage genkart osv.

Et av de største problemene med å lage formelle ontologier har vært at brukerne av en felles, delt ontologi kan ha forskjellige måter å representere kunnskap på. Den ideelle situasjonen ville vært å kunne spesifisere ontologiene uavhengig av et representasjonsspråk. Dette er i hovedsak et teknisk problem: Hvordan skal man kunne lage ontologier som kan representeres på forskjellige måter, uavhengig av plattformer (Gruber, s. 2, 1998)? Innenfor semantisk web har det blitt utviklet et språk for å representere ontologier, kalt *Ontology Web Language* (se punkt 2.2.2.4).

2.2.2.3 Fokusområder

Med utgangspunkt i det som allerede har blitt nevnt i del 2.2.2, kan Semantisk Web sies å fokusere på tre hovedområder:

1. Å lage et universelt medium / felles rammeverk for informasjonsutveksling, altså et felles format for å beskrive og representere data slik at det skal kunne hentes fra flere forskjellige kilder på tvers av applikasjoner.
2. Å utvikle språk og teknologier som gjør det mulig å si noe om hvordan data kan relateres til objekter i den virkelige verden, og som i tillegg kan tolkes av både mennesker og maskiner gjennom å tolke semantisk innhold.
3. Å gjøre det mulig å utvikle mer avanserte Knowledge Management-systemer ved å bedre organisere informasjon (Antoniou, 2004:4).

2. Teoretisk rammeverk

Hensikten er altså ikke at semantisk web skal erstatte det nåværende Internett, men heller utvide det gjennom en gradvis evolusjon. Man ser for seg at dette vil føre til en rekke endringer når det gjelder hvordan applikasjoner utvikles og derfor også hvordan mennesker bruker Internett til å finne informasjon på. Noen av implikasjonene tar Tim Berners-Lee opp i en av sine første artikler om semantisk web (Berners-Lee, 2001). Der gir han blant annet et eksempelscenario hvor en såkalt «semantisk web-agent» brukes til alt fra å skru ned volum på musikken når telefonen ringer, til å bestille time hos legen for din syke mor. Et annet eksempel er at en slik agent kan brukes når man skal planlegge ferien sin, for eksempel ved at agenten finner det beste hotellet og den beste flyruten basert på predefinerte kriterier som pris, sted, bostandard osv. Igjen blir det tatt opp at semantisk web handler om å *uttrykke mening og representere informasjon*, slik at maskiner skal kunne utføre oppgaver som tidligere måtte gjøres manuelt av mennesker.

2.2.2.4 Teknologier

To av de mest sentrale teknologiene som vil bli brukt for å gi semantikk til data, er på nåværende tidspunkt W3C-anbefalingene Resource Descriptive Framework og Ontology Web Language.

Resource Descriptive Framework⁶ (RDF) er en datamodell som tar utgangspunkt i å beskrive ressurser i form av egenskaper og verdier (objekt-attributt-verdi). RDF er særs egnet til å representere metadata om dokumenter på Internett og blir hovedsaklig brukt for å representere ulike typer informasjon på nettet (www). Hensikten er å representere informasjon på en måte som gjør at denne informasjonen kan bli delt på tvers av forskjellige typer applikasjoner.

Ontology Web Language⁷ (OWL) er en standard for å spesifisere ontologier og blir blant annet brukt av applikasjoner som har behov for å kunne tolke innholdet av informasjon. OWL har en kraftigere syntaks enn RDF, som gjør det mulig å beskrive informasjon i en maskin-prosesserbar form. Blant annet kan OWL brukes til å representere meningen av forskjellige konsepter samt relasjonene mellom disse

6 <http://www.w3.org/TR/REC-rdf-syntax/#intro>

7 <http://www.w3.org/TR/2004/REC-owl-features-20040210/#s1>

2. Teoretisk rammeverk

konseptene.

En tredje teknologi som også er relatert til semantisk web, er XML Topic Maps⁸ (XTM). Topic Maps (domenekart) gjør at man kan lage en indeks over informasjon, hvor informasjonen man indekserer til, ligger utenfor selve domenekartet (Garshol, 2002). Det er tre elementer som er sentrale for XTM: konsepter, relasjoner og forekomster (se del 3.2.2.1). Topic Maps brukes for det meste til å organisere ulike typer informasjon på nettsteder, men kan også brukes i flere andre sammenhenger, blant andre AnnForum.

2.2.2.5 Implikasjoner og eksempler på praktisk bruk

Selv om semantisk web har kommet et godt stykke på vei, er det nok ennå et stykke igjen til man vil se noen av de applikasjonene nevnt ovenfor. Swoogle⁹ er et eksempel på en semantisk web-applikasjon, mer spesifikt en semantisk søkemotor som lar bruker søke etter blant annet ontologier, dokumenter og termer. Et annet eksempel er konseptet om tagging og folksonomier (Shadbolt, 2005). Disse teknikkene har vært brukt en god stund nå og har vist seg å være veldig kraftige i forbindelse med organisering av informasjon. Blant annet blir de ofte brukt av forskjellige nyhetssider, som for eksempel i forbindelse med artiklene Dagbladet publiserer på nettsidene sine.

En *tag* kan i dette tilfellet sies å være et nøkkelord, en terminologi eller et konsept som brukes for å beskrive informasjon om et spesifikt tema. Det kan på mange måter minne om en annotering. Teknikken går ut på at ressurser klassifiseres i henhold til kategorier/konsepter/nøkkelord, for eksempel «boligmarkedet» eller «nokassaken», noe som gjør det mulig å lettere finne informasjon som omhandler samme tema. Et annet konkret eksempel på en applikasjon som bruker tagging, er Flickr¹⁰. Flickr lar brukerne dele personlige bilder med hverandre ved at bildene som blir lastet opp, kategoriserer nøkkelord (tags). YouTube¹¹ gjør også utstrakt bruk av tagging, slik at når man har valgt en film, så har man også muligheten til å se relaterte filmer.

8 <http://www.topicmaps.org/xtm/>

9 <http://swoogle.umbc.edu/>

10 <http://www.flickr.com/>

11 <http://www.youtube.com/>

2. Teoretisk rammeverk

2.2.3 Annoteringer

Det finnes flere forskjellige definisjoner av hva en annotering er, avhengig av hvilken situasjon man bruker de i, noe som gjør det vanskelig å gi en konsis, entydig definisjon av begrepet. Merriam-Webster Online Dictionary definerer en annotering som en to-steps prosess: et notat eller en kommentar lagt til et dokument, så vel som selve prosessen med å annotere¹². Annoteringer er med andre ord ekstra informasjon som legges til et dokument, enten i form av et notat eller en tilleggskommentar, med det målet å utdype innholdet i dokumentet. En annen definisjon er gitt av W3C¹³, som definerer annoteringer som eksterne kommentarer, notater eller forklaringer som kan legges til web-dokumenter, eventuelt utvalgte deler av et web-dokument.

Annoteringer har en lang rekke bruksområder, og et typisk eksempel er metadata (data om data) som ofte blir brukt i forbindelse med web-dokumenter. Et annet eksempel er innenfor semantisk web, hvor et av hovedmålene er å beskrive informasjon på Internett ved hjelp av en formell struktur, slik at maskiner kan tolke denne informasjonen på bakgrunn av innhold. Dette innebærer da å bruke formaliserte representasjoner i form av teknologier som RDF og OWL for å beskrive og lenke sammen informasjon. Et tredje eksempel kan være et kort sammendrag eller abstrakt av en bok, slik at man kan lese annoteringen før man bestemmer seg for å lese hele boken¹⁴.

Euzenat (2002) har også skrevet en del om annoteringer i semantisk web, og han sier i denne sammenhengen at hovedmålet med en annotering er å gjøre det enklere å finne frem til informasjon på bakgrunn av en formell beskrivelse av innholdet. En formalisering av begrepet kan derfor beskrives som en relasjon mellom to objekter: dokumenter og formelle representasjoner. Ut fra disse to objektene har vi så to forskjellige funksjoner, nemlig indekseringer og annoteringer. En *indeksering* blir definert som en funksjon fra en formell representasjon til et dokument og gjør det blant annet mulig å hente frem et dokument på bakgrunn av en formell representasjon. En annotering på den andre siden er en funksjon fra et dokumentets innhold til en formell representasjon. Denne representasjonen blir siden lagt til dokumentet og gjør

¹² <http://www.merriam-webster.com/dictionary/annotations>

¹³ http://www.w3.org/Amaya/User/attaching_annotations/what_is_an_annotation.html

¹⁴ http://myrin.ursinus.edu/help/resrch_guides/annotate.htm

2. Teoretisk rammeverk

det mulig å tolke dokumentet på bakgrunn av innholdet.

2.2.4 Taksonomier og folksonomier

Som nevnt i del 2.2.2.5 henger begrepet tagging sammen med klassifisering av ressurser, noe som er veldig relevant for semantisk web. Flickr og YouTube har allerede blitt nevnt som eksempler på nettsteder som benytter seg av tagging, og Gruber (2007) nevner i sin artikkel at tagging først dukket opp som en måte for folk å assosiere ressurser på Internett med personlige nøkkelord. Dette fenomenet har senere fått navnet *folksonomi* av Vander Wahl (2007) og brukes hovedsaklig i sosiale kontekster. Folksonomier har flere likhetstrekk med ontologier, men med den forskjellen at en folksonomi ikke benytter seg av et like formelt språk som en ontologi. Dette gjør det lettere for «vanlige» folk å benytte seg av folksonomier fremfor ontologier da de kan bruke sitt eget vokabular til å beskrive ressurser og informasjon (Vander Wahl, 2007). En folksonomi kan ses på litt i motsetning til en taksonomi, som også tar for seg klassifisering av informasjon, men innenfor et mer rigid regelsett. En taksonomi er en form for semantisk modellering som definerer et antall konsepter, samt relasjonene mellom disse konseptene som et hierarki av konsepter og underkonsepter, og har tradisjonelt sett blitt brukt innenfor vitenskapen til å for eksempel klassifisere dyr og planter (Merriam-Webster¹⁵).

2.2.5 Domenemodell

Et domenekart, også kalt en domeneontologi, er en type ontologi som gir en detaljert beskrivelse av konseptene innenfor et avgrenset domene, og er nært knyttet opp til blant annet kunstig intelligens og semantisk web (Antoniou, 2004:10). Domenekart blir vanligvis konstruert manuelt av eksperter, for eksempel forskere, da det er disse som oftest innehar den nødvendige kunnskapen om det aktuelle området.

Det finnes flere forskjellige teknologier for å lage domeneontologier i forbindelse med semantisk web, hvor Ontology Web Language (OWL) nok er den mest kjente. OWL er en W3C-spesifikasjon som gjør det mulig å formelt beskrive et konseptuelt domene i relasjon til semantisk web. Andre standarder som også kan brukes for å lage

¹⁵ <http://www.merriam-webster.com/dictionary/taxonomy>

2. Teoretisk rammeverk

ontologier, er XML Topic Maps og Resource Descriptive Framework. Topic Maps gjør det blant annet mulig å lage en oversikt over alle konseptene innenfor et utvalgt domene samt å trekke konklusjoner på bakgrunn av konseptene og relasjonene mellom dem.

2.3 Gjenbruk av læringsressurser

Det finnes flere forskjellige typer semantisk annotering og klassifisering av web-baserte ressurser, både når det gjelder diskusjonsforumer og i andre sammenhenger. Som det ble nevnt i del 1.1.3, har blant annet Chen (2004), Helic (2003) og Craven (1998) lagd systemer for semantisk annotering, men med litt forskjellige fremgangsmåter. Denne delen vil gi en forklaring til de forskjellige systemene, i tillegg til å diskutere fordelene og ulempene mellom dem. De fleste fremgangsmåtene går ut på å klassifisere ressurser på bakgrunn av semantikk, og da er teori fra semantisk web-feltet en sentral del.

2.3.1 Annotering og klassifisering av innlegg i diskusjonsforumer

Helic, Maurer og Scerbakov (2003, 2005) har skrevet en del annotering av ressurser i diskusjonsforumer, og de sier blant annet at diskusjonsforumer ikke bare kan være nyttige som et kommunikasjonsverktøy, men at de også kan inneholde mye informasjon relatert til læring. Denne informasjonen kan være i form av spørsmål, dialoger, referering til relevant litteratur lagt til av tidligere studenter, og det ville derfor vært meget nyttig om man kunne tatt vare på disse ressursene og gjort de tilgjengelige for nye studenter som nytt læringsmateriale.

Men det å gjenbruke innlegg i diskusjonsforumer er ingen enkel sak av flere årsaker. Det største problemet er at diskusjonsforumer har veldig dårlige funksjoner for informasjonsfremfinning. Bortsett fra fulltekst-søk er det så godt som ingen funksjoner for å finne frem til relevante innlegg på en enkel måte (Helic, 2005). Det er heller ikke er mulig å foreta kontekstbaserte søk eller navigere etter forskjellige konsepter. I tillegg er titlene for de forskjellige innleggene ofte lite beskrivende, noe som har sammenheng med at brukeren sjelden er motivert nok til å bruke tid på å finne en beskrivende tittel. Kombinert fører disse problemene til at det blir ekstremt

2. Teoretisk rammeverk

vanskelig å bla seg frem til relevante innlegg på en enkel måte. Hovedgrunnen for de ovennevnte svakhetene er ifølge Helic at diskusjonsforumer har en total mangel på konseptuell struktur, og de har derfor i denne sammenhengen utviklet et system kalt WBT-master.

WBT-master et system beregnet på web-baserte diskusjonsforumer og baserer seg på at man gir en konseptuell, semantisk struktur til diskusjonsforumet (Helic, 2003). Dette gjøres med utgangspunkt i et konseptuelt skjema som inneholder konseptene for et utvalgt domene samt relasjonene mellom konseptene. Innleggene i diskusjonsforumet blir deretter tilordnet/klassifisert innleggene i henhold til disse konseptene. Selve tildelingen gjøres manuelt av brukerne. Det er to hovedfremgangsmåter for å klassifisere innlegg i diskusjonsforumer i henhold til den konseptuelle modellen. Den første måten er å gi selve forumet en predefinert struktur basert på den konseptuelle modellen. Den andre måten er benytte seg av en automatisk klassifisering av innleggene i forumet basert på å sammenlikne nøkkelord i innleggene mot konseptnavnene.

Det kan i følge Helic være problematisk å gi selve forumet en predefinert struktur hovedsaklig fordi at det kan forekomme tilfeller der et innlegg kan tilordnes flere enn et konsept. Når det gjelder den automatiske klassifiseringen så sier Helic (2005) at dette kan fungere tilfredsstillende i mange situasjoner, men at det kan føre til problemer i de tilfellene der man bruker flere språk i forumet. Man vil da bli nødt til å oversette konseptene og å ha flere forskjellige applikasjoner for de forskjellige språkene, noe som fører at det kan bli for komplisert å både implementere og administrere. I stedet tok de en litt annen fremgangsmåte gjennom å adskille den konseptuelle modellen fra selve diskusjonsforumet. Selve tilordningen av innleggene til de forskjellige konseptene i WBT-master foretas manuelt før innlegget legges til, noe som krever en aktiv deltakelse fra studenten sin side. Helic nevner her at for at denne fremgangsmåten skal fungere må brukerne være motiverte nok, hvis ikke kan det hende at de ikke gidder å bruke tid på den manuelle operasjonen. Men i dette tilfellet er motivasjonen at brukerne faktisk deltar på kurset av egen fri vilje. Det ble også foretatt en evaluering i forbindelse med WBT-master, og resultatene var stort sett positive. Studentene så fordelene med et slikt system, og at de tilegnet seg ekstra

2. Teoretisk rammeverk

kunnskap ved hjelp av systemet.

2.3.2 Knowledge Management Systems

Større bedrifter og organisasjoner har ofte store mengder data, informasjon og kunnskap som må holdes styr på. Denne kunnskapen består som regel av tekstdokumenter, videoer og arkivfiler og er som oftest i en dårlig strukturert form (Antoniou, 2004:3–7). Knowledge Management ser på forskjellige måter for kunne å håndtere organisatorisk informasjon og kunnskap, slik at det skal bli enklere å organisere, identifisere, vedlikeholde, representere og dele til dels store mengder kunnskap.

Sett i denne sammenhengen er *Knowledge Management Systems* (KMS) datasystemer som kan hjelpe bedrifter må å få et tilrettelagt miljø hvor det blir enklere å lagre, analysere, gjenbruke og finne frem til organisatorisk informasjon (Maier, 2007). Brukerne av et slikt system vil da få lettere tilgang til sentral kunnskap, noe som har mange fordeler når det kommer til for eksempel utveksling av ideer, samarbeid og så videre. Andre potensielle fordeler ved å ha en slik delt kunnskapsdatabase er at det fører til mindre ekstraarbeid da man har en slags sentral database hvor man kan søke etter informasjon. I tillegg kan slike systemer gjøre det lettere å trene opp nye ansatte, noe som minsker usikkerheten ved å være avhengig av såkalt taus kunnskap.

Taus kunnskap er kunnskap som kan være særdeles viktig for en organisasjon, men som ikke er formelt skrevet ned noen plass. I stedet er denne typen kunnskap personlig og kommer ofte på bakgrunn av for eksempel erfaring (Gourlay, 2003). Det kunne derfor være ekstremt verdifullt om man kunne tatt vare på denne informasjonen når en sentral person slutter, hvis ikke vil den gå tapt. Eksempler på Knowledge Management-systemer inkluderer blant annet Lotus Notes¹⁶, SAP Knowledge Warehouse¹⁷, MoinMoin¹⁸, KnowledgeTree¹⁹, ulike typer diskusjonsforumer og ikke minst Internett.

16 <http://www-306.ibm.com/software/lotus/products/notes/>

17 http://help.sap.com/saphelp_nw04/helpdata/en/8f/d82d37ba466b5be10000009b38f936/frameset.htm

18 <http://moinmo.in/>

19 <http://www.knowledgetree.com/>

2. Teoretisk rammeverk

Selv om Knowledge Management systemer kan bidra til å gjøre organiseringen av informasjon enklere så har de også flere svakheter (Antoniou, s. 3-7, 2004). Noen av disse problemene har allerede blitt nevnt i del 2.2.2 i forbindelse med Internett, og da særlig dårlige funksjoner for å søke etter informasjon, samt begrenset eller ingen konseptuell struktur. Knowledge Management som felt er derfor særdeles relevant for Semantisk Web, hovedsaklig fordi at ny teknologi vil gjøre det mulig å lage mer avanserte KMS som ikke har de ovennevnte svakhetene. Konseptuelle domenemodeller og ontologier kan i denne sammenhengen brukes til å lage en helhetlig oversikt over et spesifikt område, for eksempel innenfor en organisasjonsstruktur, slik at konsepter og informasjon lettere kan klassifisere.

Grunnen til at ontologier er så godt egnet til slike formål er at de gjør det lettere å beskrive forskjellige typer entiteter og relasjonene mellom disse entitetene på bakgrunn av innhold (se del 2.2.2.2). Når man har organisert informasjon ved hjelp av en ontologi eller endomenemodell i så blir det lettere å bruke automatiske teknikker til å hente frem ønskede data. Blant annet vil man kunne benytte seg spørringer i stedet for vanlige tekstbaserte søk; noe som blir mulig å få til nettopp fordi at man har tilgang til informasjon i en strukturert form. For eksempel vil søkemotorer kunne finne nettsider etter både konsept og innhold, fremfor å bare returnere alle sider som inneholder søkeordene. Flere eksempler på hvordan ontologier kan brukes i forbindelse med Semantisk Web og informasjonsfremfinning finnes i del 2.2.5.

Ontologier og domenemodeller er som nevnt veldig godt egnet til å klassifisere informasjon om utvalgte området, og dette gjelder også i en organisasjonssammenheng der man ofte har mye kompleks informasjon. For eksempel kan man bruke en domenemodell som et organisasjonskart, eller man kan gruppere informasjon i henhold til blant annet avdelinger. Men andre ord så kan ontologier/domenekart brukes til å forbedre funksjonaliteten i Knowledge Management systemer ved hjelp av momentene nevnt i forrige avsnitt.

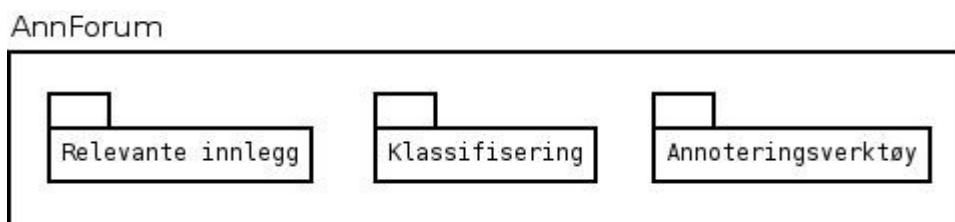
3. Design og utvikling

3. Design og utvikling

Dette kapittelet vil gi en oversikt av AnnForum og hvordan det er bygd opp, fra selve utviklingsprosessen til designmessige valg som har blitt gjort underveis. Blant annet så inkluderer dette formelle krav til systemet, sentrale funksjoner, skjermbilder av viktige brukergrensesnitt, klassesdiagram, systemarkitektur, utviklingsverktøy, samt eksempelkode av de mest sentrale klassene og metodene.

Kort forklart så er AnnForum et system for semi-automatisk, semantisk annotering av innlegg i et diskusjonsforum hvor hovedmålet er å kunne gjenbruke innlegg fra tidligere semesters diskusjoner. Disse tidligere innleggene blir gjort tilgjengelige i diskusjonen for det nåværende semesteret som relevante/relaterte ressurser.

Diskusjonsforumet som har blitt benyttet er FLE3 sin kunnskapsbyggingsmodul, og måten AnnForum fungerer på er at en konseptuell domenemodell blir brukt til å beskrive konseptene innenfor Kunstig Intelligens domenet, samt relasjonene mellom disse konseptene. Etter at domenemodellen har blitt opprettet så kan man lage annoteringer ved at læreren lenker stillingsressursene, altså innlegg fra tidligere semestres diskusjoner, opp til de relevante konseptene i domenemodellen. Med andre ord er det snakk om en klassifisering av innleggene. Denne klassifiseringen gjøres automatisk av systemet, men læreren har også mulighet til å gå inn i etterkant og redigere annoteringene, i tillegg til å endre selve domenemodellen ved å legge til og fjerne konsepter. Figur 3.1 illustrerer den konseptuelle oppbyggingen av systemet.



Figur 3.1: De forskjellige delene i AnnForum

Grensesnittet for relevante innlegg er integrert i kunnskapsbyggingsmodulen til FLE3 og brukes kun av studenten. Klassifiseringsregelen er mer uavhengig og blir blant annet benyttet av FLE3 når et nytt innlegg legges til, eller når et helt diskusjonsforum fra et tidligere semester skal importeres og annoteres. Annoteringsverktøyet er helt

3. Design og utvikling

adskilt fra FLE3 og brukes kun av læreren for å redigere domenemodellen og annoteringene for de klassifiserte innleggene. Det som skiller AnnForum fra systemene til Helic (2003) og Craven (1998), er at klassifiseringen av innleggene skjer i sanntid. Det vil si at i det øyeblikket studenten legger til innleggene, vil klassifiseringsalgoritmen i AnnForum automatisk klassifisere innlegget i henhold til de riktige konseptene. Dette har den fordelen at studenten slipper å bruke tid på å gjøre denne tilordningen manuelt.

3.1 Systemkrav

De formelle kravene er delt opp i «use cases», funksjonelle krav og ikke-funksjonelle krav og gir en oversikt over den overordnede funksjonaliteten for systemet. «Use cases» brukes som utgangspunkt for å finne funksjonelle krav og tar for seg interaksjonen mellom aktøren og systemet, i dette tilfellet studenten/læreren og AnnForum. Funksjonelle krav er de funksjonene som systemet må støtte, og er på mange måter systemets planlagte oppførsel. Disse blir som regel listet opp i formatet «systemet skal» eller «systemet må». Ikke-funksjonelle krav er sekundære funksjoner som støtter opp om de funksjonelle kravene, og som ikke er direkte kritiske for applikasjonen, men heller beskriver hvordan systemet fungerer, for eksempel med tanke på skalerbarhet, ytelse, kostnader og så videre.

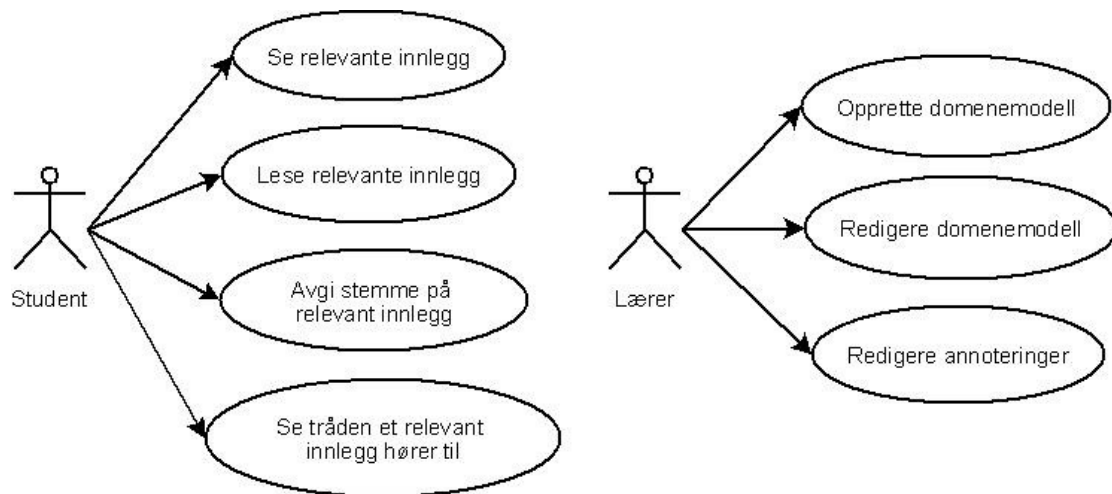
Læringsystemer har ofte spesielle krav, og når man skal lage semantiske annoteringsverktøy for læringsmateriale, er det en del praktiske hensyn man må ta. Chen har identifisert og oppsummert noen av disse kravene på bakgrunn av eksisterende annoteringsverktøy (Chen, 2005). Disse kravene knytter seg til blant annet brukbarhet, delbarhet og brukervennlighet, mer spesifikt at man skal fokusere på selve læringskonteksten ved utviklingen av systemet, at de forskjellige aktørene skal kunne kommunisere gjennom å gjenbruke de semantiske annoteringene, samt at systemet skal være lettfattelig og enkelt å bruke. Den beste fremgangsmåten ifølge Chen er om «[...] verktøyet er automatisk, kognitivt og semantisk [...]», med andre ord at datamaskinen håndterer klassifiseringen av innleggene på vegne av brukeren. Fordelen med dette er at det fører til en minimal forstyrrelse av brukerens arbeidsflyt, og at det reduserer arbeidsmengden. Denne problemstillingen er veldig relevant i

3. Design og utvikling

forbindelse med diskusjonsforumer, hvor brukerne som regel ikke er motiverte nok til å manuelt tilordne innleggene til konsepter i domenemodellen.

3.1.1 Use cases

De funksjonelle systemkravene kan kort oppsummeres ved hjelp av use case diagrammet nedenfor (figur 3.2: Scenarioer), som igjen er utformet på bakgrunn av problemstillingen. Diagrammet viser hvilke handlinger som må kunne utføres av studenten og hvilke handlinger som må kunne utføres av læreren. Tilsvarende diagrammer og use-cases for FLE3 finnes i dokumentasjonen²⁰.



Figur 3.2: Use case scenarios

3.1.2 Funksjonelle krav – systemet skal

Funksjonene som systemet må støtte, er listet opp nedenfor og tar utgangspunkt i use-case-diagrammene i forrige del. Systemet må/skal oppfylle følgende krav:

- Det skal være en domeneuavhengig tilleggsmodul utviklet i blant annet Java, Python og PHP. I første omgang vil det kun være beregnet på læringssystemet FLE3 sitt diskusjonsforum.
- Det skal kunne annotere/klassifisere innlegg basert på en predefinert domenemodell. Denne domenemodellen inneholder de viktigste og mest

²⁰ <http://fle3.uiah.fi/> - UML diagrams of FLE3

3. Design og utvikling

brukte konseptene innenfor kunstig intelligens.

- Det skal kunne klassifisere innlegg fra tidligere semesters diskusjoner. Dette innebærer blant annet en funksjon for å importere og annotere et helt diskusjonsforum. Klassifiseringsalgoritmen gjenkjenner konsepter i domenemodellen som går igjen i innleggene, og finner også relevansverdien for de enkelte innleggene.
- Det skal kunne klassifisere nye innlegg for det nåværende semesteret. Dette innebærer blant annet at klassifiseringen må skje i det øyeblikket innlegg legges til. Selve klassifiseringen foregår på samme måte som i forrige punkt.
- Det skal ikke forstyrre læringsprosessen for studentene. Dette innebærer at klassifiseringen må gjøres automatisk av systemet for ikke å ødelegge arbeidsflyten til aktørene.
- Det skal kunne hente frem innlegg fra tidligere semester som er relevante for den nåværende diskusjonen, og vise frem disse til studentene. De relevante innleggene skal kunne hentes frem for både nye så vel som eksisterende innlegg. Innleggene skal rangeres etter relevansverdi når de listes opp.
- Det skal la studentene kunne avgi stemme for om de synes et tidligere innlegg er relevant for den nåværende diskusjonen eller ikke. Når et innlegg har fått flere negative stemmer enn positive, vil det fjernes fra listen.
- Det skal tilby et brukergrensesnitt som lar læreren opprette og redigere en domenemodell for kunstig intelligens. Dette inkluderer å kunne legge til og fjerne relevante innlegg for de enkelte konseptene innen domenemodellen.
- Det skal tilby et brukergrensesnitt som lar læreren redigere annoteringene for de klassifiserte innleggene, både for tidligere og nåværende innlegg. Dette innebærer at man skal kunne legge til og fjerne hvilke konsepter som er relatert til de enkelte innleggene.

3. Design og utvikling

- Det skal tilby et brukergrensesnitt som lar studenten se hele tråden som et relevant innlegg hører inn under. Studenten vil da kunne få en bedre oversikt over problemområdet.
- Det skal tilby et brukergrensesnitt som lar studenten bla gjennom alle innlegg tilknyttet et spesifikt konsept. Dette vil gi studenten muligheten til å få en oversikt over ressurser gruppert etter tema.

3.1.3 Ikke-funksjonelle krav – systemet bør

Funksjoner som støtter opp om de funksjonelle kravene, er listet opp nedenfor.

Systemet bør:

- være så intuitivt og lettfattelig som mulig for ikke å forstyrre læringsprosessen til studenten. Dette innebærer blant annet et enkelt og oversiktlig brukergrensesnitt.

3.1.4 Scenarioer

De to mest typiske scenarioene for gjenbruk og klassifisering av innlegg i FLE3 er skissert opp nedenfor:

SCENARIO 1: STUDENTEN SKRIVER ET NYTT INNLEGG:

1. Studenten er ferdig å skrive innlegget og trykker på «Submit».
2. Innlegget legges til FLE3 sitt diskusjonsforum.
3. Klassifiseringsmodulen som er integrert i FLE3, finner de relevante konseptene for innlegget.
4. For hvert relevante konsept regnes det ut en relevansverdi. Denne relevansverdien regnes ut ved å se på antall forekomster av konseptet samt om konseptet forekommer i innleggets tittel eller i selve innholdet.
5. Klassifiseringsresultatet, altså de relevante konseptene sammen med

3. Design og utvikling

relevansverdiene, legges til i MySQL-databasen.

6. Resultatet for det nye innlegget sendes tilbake til brukergrensesnittet for relevante innlegg. Dette inkluderer altså de relevante konseptene sammen med relevansverdiene.
7. Forekomstene for de relevante konseptene, altså innleggene fra tidligere semester, hentes frem fra domenemodellen (Topic Map-filen) og vises til brukeren dersom ønsket. Innleggene er rangert etter relevansverdi i prosent.

SCENARIO 2: STUDENTEN LESER ET EKSISTERENDE INNLEGG I FLE3:

1. Studenten leser et innlegg i FLE3.
2. Studenten trykker på knappen som viser relevante innlegg.
3. Brukergrensesnittet for relevante innlegg henter frem de relevante konseptene fra MySQL databasen for nåværende semester.
4. Forekomstene for de relevante konseptene, altså innleggene fra tidligere semester, hentes frem fra domenemodellen (Topic Map-filen) og vises til brukeren. Innleggene er rangert etter relevansverdi i prosent.

Fremgangsmåten for å importere et tidligere diskusjonsforum følger de samme prinsippene når det kommer til klassifisering og lagring, men er litt mer komplisert. Det første som må gjøres, er å hente ut alle innleggene fra FLE3 sin database, og i dette tilfellet ble dette gjort ved å eksportere databasen til en XML-fil ved hjelp av grensesnittet i Zope. Deretter ble denne XML-filen lest og innleggene satt inn i MySQL-databasen ved hjelp av et Python-script.

3.2 Arkitektur og hovedkomponenter

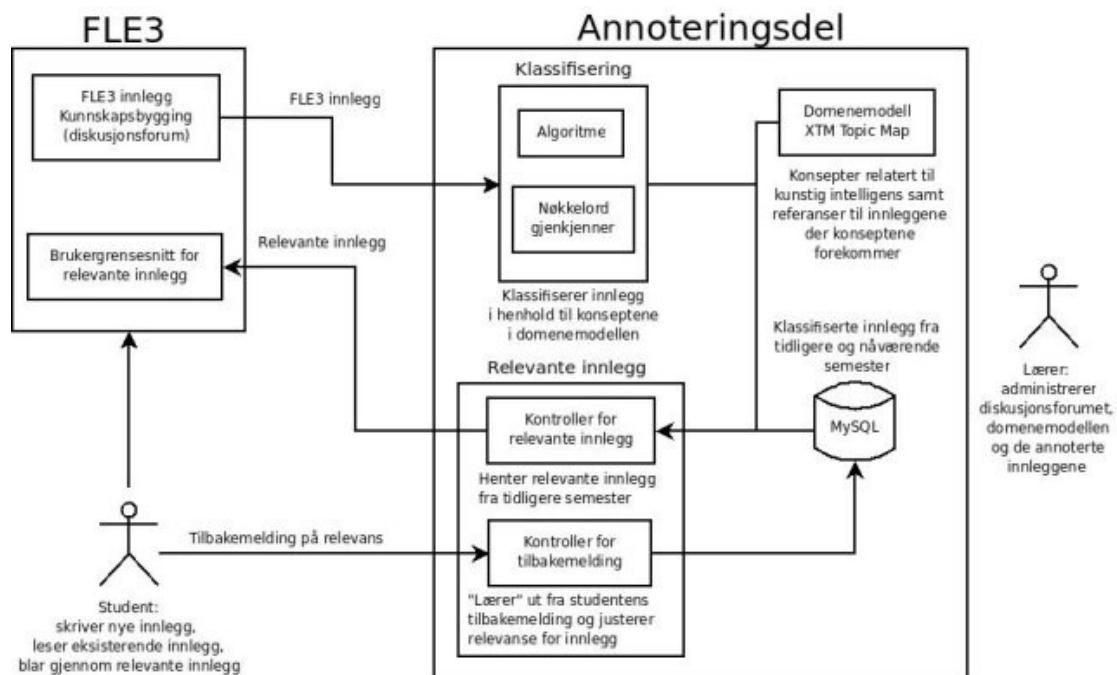
Her beskrives utformingen og oppbyggingen av AnnForum i sin helhet, noe som blant annet inkluderer en oversikt over generelt design og systemarkitektur. Dette innebærer blant annet en oversikt over de forskjellige modulene, komponentene,

3. Design og utvikling

brukergrensesnitt og data som er nødvendige for at systemkravene skal tilfredsstilles. Systemarkitekturen med tanke på den konseptuelle domenemodellen, MySQL-databasen, klassifiseringsalgoritmen og relasjonene mellom disse samt presentasjonen av de relevante innleggene går også gjennom i detalj. Klassediagrammet i appendiks A gir en oversikt over de viktigste klassene.

3.2.1 Systemarkitektur

Systemet er ment å være en domeneuavhengig modul bestående av to undermoduler: en klassifiseringsmodul og en modul for relevante innlegg. Klassifiseringsmodulen har ansvar for å klassifisere innleggene i diskusjonsforumet etter konseptene i domenemodellen, noe som gjøres ved hjelp av en nøkkelordgjenkjenner som finner forekomstene av konseptene, samt en algoritme som bestemmer relevansverdien for et innlegg sammenlignet med disse konseptene.



Figur 3.3: konseptuell oversikt over AnnForum

Modulen for relevante innlegg har ansvar for å hente frem relevante innlegg fra tidligere semesters diskusjoner som er relevante for den nåværende diskusjonen, vise disse til studenten ved hjelp av brukergrensesnittet for relevante innlegg samt ta imot

3. Design og utvikling

tilbakemeldinger fra studenten angående de relevante innleggene.

Det har også blitt utviklet en applikasjon som lar læreren få en oversikt over diskusjonsforumene og administrere den konseptuelle domenemodellen, i tillegg til la å ham redigere annoteringene for de enkelte innleggene.

Hele systemet består av to hoveddeler som illustrert i figur 3.3. Disse to delene er et brukergrensesnitt for relevante innlegg som er integrert i FLE3, og annoteringsdelen som blant annet inneholder klassifiseringsalgoritmen som knytter de forskjellige innleggene opp mot riktige konsepter.

Inne i annoteringsdelen er MySQL-databasen som lagrer innleggene og de relevante konseptene for hvert innlegg. Den konseptuelle domenemodellen som inneholder alle konseptene for AI-domenet, og referanser til innlegg der konseptene forekommer, hører også til under annoteringsdelen.

3.2.2 Domenemodellen

Den konseptuelle domenemodellen som er brukt i forbindelse med AnnForum, er i form av et XML Topic Map (XTM)²¹. Et Topic Map er, som allerede nevnt i kapittel 2, en spesifisering som gjør det mulig definere ulike konsepter innenfor et utvalgt domene. Man kan også definere relasjonene mellom disse konseptene, i tillegg til å lenke hvert konsept opp til ulike ressurser. Et Topic Map er med andre ord basert på følgende tre konsepter:

1. Topic – et emne i forbindelse med det utvalgte domenet. For kunstig intelligens kan for eksempel topic (baseName) være «Machine Learning». Emnet kan også ha ett eller flere synonymer (variantNames).
2. Association – relasjonene mellom de forskjellige konseptene.
3. Occurrences – forekomster av ressurser som er relevant i forhold til *topic*.

Syntaksen for domenemodellen til AnnForum ble noe forenklet i forhold til den

²¹ <http://www.topicmaps.org/xtm/>

3. Design og utvikling

offisielle XTM-spesifikasjonen. Grunnen til dette var at det forenklet utviklingen, samt at den fullstendige spesifikasjonen ikke var nødvendig for at systemet skulle fungere. Men for en fremtidig versjon av AnnForum, når standardisering blir mer viktig, bør man nok etterstrebe å følge XTM-spesifikasjonen. Nedenfor er et lite utdrag fra domenemodellen for kunstig intelligens med et par klassifiserte innlegg. Se forøvrig appendiks B for en oversikt over hele modellen uten occurrences.

```
<topicMap name="Artificial Intelligence">
<topic id="147">
  <baseName>Turing Test</baseName>
  <association>
    <member name="Objections"/>
  </association>
  <occurrence id="15">
    <title>Think differently</title>
    <resourceRef href=""/>
  </occurrence>
</topic>
</topicMap>
```

3.2.3 Databasen

De klassifiserte FLE3-innleggene for nåværende og tidligere semester blir lagret i en MySQL-database med en tabell for hvert semester. Innleggene blir lagret i sin helhet med tittel, innhold, kategori og så videre samt at det finnes et felt som sier hvilke konsepter innlegget er relatert til. Hvert konsept svarer til et konsept i den konseptuelle domenemodellen og har også en relevansverdi knyttet til seg. Skjemaet for tabellene ser da slik ut:

(*id*, *title*, *content*, *category*, *keywords*, *path*, *user*, *reply_to_id*, *related_concepts*)

Foreløpig er det kun to tabeller i databasen: en tabell for innlegg fra forrige semester og en tabell for innlegg fra det nåværende semesteret. Men ideen er at det skal være mulig å ha en tabell for hvert semester som kunstig intelligens har blitt undervist i. På

3. Design og utvikling

den måten vil man da etter hvert få et stort antall relevante innlegg som kan brukes som relevante læringsressurser for det nåværende semesteret.

Et klassifisert FLE3-innlegg fra et tidligere semester kan da eksempelvis se slik ut:

id	15
title	Think differently
content	In my opinion, the Turing test is not a good criteria to determine artificial intelligence ...
category	my_expl
keywords	<i>foreløpig ikke i bruk</i>
path	<i>foreløpig ikke i bruk</i>
user	Arild
reply_to_id	14
related_concepts	Reasoning:0,26;Turing Test:0,37;

Feltet «related_concepts» er en semikolon-separert liste med alle konseptene som innlegget er relatert til, i dette tilfellet «Reasoning» med relevansverdi 0,26 og «Turing Test» med relevansverdi 0,37. Feltene «keywords» og «path» er foreløpig ikke i bruk, men er tatt med da de finnes i objekt databasen til FLE3. Relevansverdien brukes når de relevante innleggene skal rangeres, og blir i hovedsak benyttet av klassifiseringsalgoritmen.

3.2.4 Klassifiseringsalgoritmen

Som allerede nevnt er det klassifiseringsalgoritmen som har ansvaret for å gruppere nye innlegg etter de aktuelle, relevante konseptene i den konseptuelle domenemodellen. Algoritmen har også ansvaret for å regne ut relevansverdien for hvert konsept tilknyttet et innlegg. Det har også blitt konstruert en nøkkelordgjenkjenner som tar «baseName» (hovednavnet) og «variantName» (synonym) for hvert konsept i domenemodellen, for deretter å søke etter forekomstene av disse ordene i FLE3-innlegget. Et treff som forekommer i tittelen på innlegget, vil gi en høyere rangering enn et treff i selve innholdet fordi dette gir en bedre indikasjon på hva innlegget handler om. Merk at denne beslutningen ble tatt på bakgrunn av at

3. Design og utvikling

man i FLE3 er nødt til å skrive en beskrivende tittel for hvert nytt innlegg.

Pseudokoden nedenfor skisserer gangen i klassifiseringen av et nytt innlegg:

Gitt et nytt innlegg M (for nåværende semester)

For hvert konsept K i den konseptuelle domenemodellen

Sett relevansverdi $R1 = 0$

Sett relevanseverdi $R2 = 0$

Hvis K forekommer i tittelen til M :

$R1 = \text{hentRelevansVerdiFor}(M)$

Hvis K forekommer i innholdet til M :

$R2 = \text{hentRelevansVerdiFor}(M)$

$R = \text{MAX}(R1, R2)$

Hvis R er større enn grenseverdi $(0,19)$

Sett konsept K til å ha relevansverdi R

Legg til konsept K til innlegg M

Slutt

Dersom innlegget som klassifiseres er fra et tidligere semester, vil innlegget også bli lagt til som en forekomst (occurrence) i den konseptuelle domenemodellen. Se del 3.2.2.1 for eksempel av en forekomst. Måten klassifiseringen skjer på, kan også beskrives ved hjelp av scenarioene i del 3.1.4.

Selve klassifiseringsalgoritmen ble utviklet i Java, og for å finne relevansverdiene ble det brukt et Java-bibliotek som heter Classifier4J²². Dette biblioteket inneholder en «vector space search»-algoritme, som blant annet passer veldig bra i forbindelse med klassifisering av tekst. Java-koden nedenfor viser hvordan Classifier4J bruker «vector space search»-algoritmen til å finne relevansverdien for tittelen til et aktuelt innlegg:

```
double result = 0; //relevanseverdien  
vectorStorage = new HashMapTermVectorStorage();  
vectorClassifier = new VectorClassifier(vectorStorage);  
vectorClassifier.teachMatch("title", title);
```

²² <http://classifier4j.sourceforge.net/>

3. Design og utvikling

```
result = vectorClassifier.classify("title", t.getBaseName());  
// tilsvarende gjelder for variantNames - t.getVariantName()
```

Pseudokoden for koden over er som følger:

```
Gitt et innlegg I og et konsept K  
Sett resultat = 0  
Opprett et vektor-objekt V  
resultat = k.klassifiser(I.tittel, K.baseName)  
Gjenta for innleggets synonymer (variantNames) og innhold (content)
```

Samme fremgangsmåte gjelder også for å finne relevansverdien for selve innholdet i et innlegg og for å sjekke for forekomstene av synonymer («variantNames»).

3.2.5 Relevante innlegg og presentasjon

De relevante innleggene fra tidligere semester, tilknyttet det innlegget som studenten leser, presenteres i en rangert liste som vist i figur 3.6. Hvert av innleggene i listen kan ha ett eller flere relevante konsepter i tillegg til en relevansverdi tilknyttet hvert konsept. Det samme gjelder for konseptene til FLE3-innlegget for det nåværende semesteret. Relevansverdien for hvert av innleggene i figur 3.6 blir regnet ut som en vektlagt term (verdi) av forholdet mellom konseptene sine relevansverdier, og relevansverdiene for det nåværende innlegget sine konsepter.

For eksempel: Sett at det ene innlegget (som studenten leser) har to relevante konsepter, konsept A og konsept B, og hvor konsept A og konsept B har hver sin relevansverdi på for eksempel 0,7 og 0,3. I den konseptuelle domenemodellen, altså XML-filen, er det to innlegg for konsept A og fire innlegg for konsept B, for en total av seks relevante innlegg. Disse seks innleggene er som nevnt lagret i sin helhet i MySQL-databasen, sammen med relevante konsepter og relevansverdier. De seks innleggene er også relatert til konsept A og konsept B fordi de allerede ligger i domenemodellen.

For å utdype enda litt mer kan vi si at det i databasen er to relevante innlegg fra

3. Design og utvikling

tidligere semester: innlegg M1 og innlegg M2. Relevansverdien vil da bli regnet ut på følgende måte, gitt kriteriene nedenfor:

- M1 er relevant til konsept A med en verdi ($V_a, m1$) på 0,7.
- M1 er relevant til konsept B med en verdi ($V_b, m1$) på 0,3.
- M2 er relevant til konsept A med en verdi ($V_a, m2$) på 0,3.
- M2 er relevant til konsept B med en verdi ($V_b, m2$) på 0,2.

Det nåværende innlegget som studenten leser, innlegg M3, har to relevante konsepter: konsept A og konsept B. Konsept A har to innlegg i databasen, M1 og M2, og relevansverdien blir da som følger:

- Mellom M3 og M1: $V_a, m1 * 0,5 + V_b, m1 * 0,5 = 0,7 * 0,5 + 0,3 * 0,5 = 0,55$.
- Mellom M3 og M2: $V_a, m2 * 0,5 + V_b, m2 * 0,5 = 0,3 * 0,5 + 0,2 * 0,5 = 0,25$.

Når brukeren har valgt et spesifikt relevant innlegg, som vist i figur 3.7, vil han også her ha muligheten til å avgi en stemme for om innlegget var nyttig eller ikke. For hver stemme vil relevansverdien for innlegget økes eller reduseres med femti prosent avhengig av om man stemmer positivt eller negativt. Brukergrensesnittet ble lagd ved hjelp av henholdsvis HTML og PHP som kalte på MySQL-databasen og den konseptuelle domenemodellen.

3.3 Integrasjon med FLE3

Proessen med å integrere AnnForum med FLE3 var todelt. Det første som ble gjort, var å legge til en knapp i FLE3 sin kunnskapsbyggingsmodul (figur 3.5), slik at denne knappen åpnet et nytt popup-vindu til brukergrensesnittet for relevante innlegg. Dette ble gjort ved å legge til én enkel linje med DHTML-kode i */ui/Note/index_dhtml-filen*.

3. Design og utvikling

Neste steg var å få til den dynamiske klassifiseringen, slik at nye FLE3-innlegg ble lagt til både i databasen og den konseptuelle domenemodellen etter at de var lagt til i FLE3 sin objektdatabase. Dette ble gjort ved å modifisere Python-koden som behandlet et nytt innlegg (*Note.py*), slik at innlegget ble sendt inn som parameter til AnnForum sin klassifiseringsmodul for videre behandling der.

Som det ble skissert, var prosessen med å integrere AnnForum i FLE3 ikke altfor omstendelig, og det å bruke samme fremgangsmåte for andre diskusjonsforumer bør ikke være stort vanskeligere. Så lenge man kan få utført følgende tre steg for det aktuelle diskusjonsforumet, så bør det ikke være vanskelig å gjenbruke innlegg:

1. Man må få importert eksisterende innlegg som kan brukes som relevante ressurser. For FLE3 ble dette gjort ved å lese et XML-dokument med alle innleggene.
2. Man må få vist brukergrensesnittet for relevante innlegg. Dette vil hovedsaklig gjøres ved å legge til en knapp der man ønsker å vise de relevante ressursene.
3. Man kan få kalt på klassifiseringsmodulen til AnnForum, slik at modulen kan klassifisere nye innlegg dynamisk etter hvert som de legges til.

3.4 Gjennomgang av systemet

Denne delen gir en beskrivelse og gjennomgang av brukergrensesnittet for relevante innlegg. Lærerens annoteringsverktøy vil også bli gått gjennom i del 3.4.2.

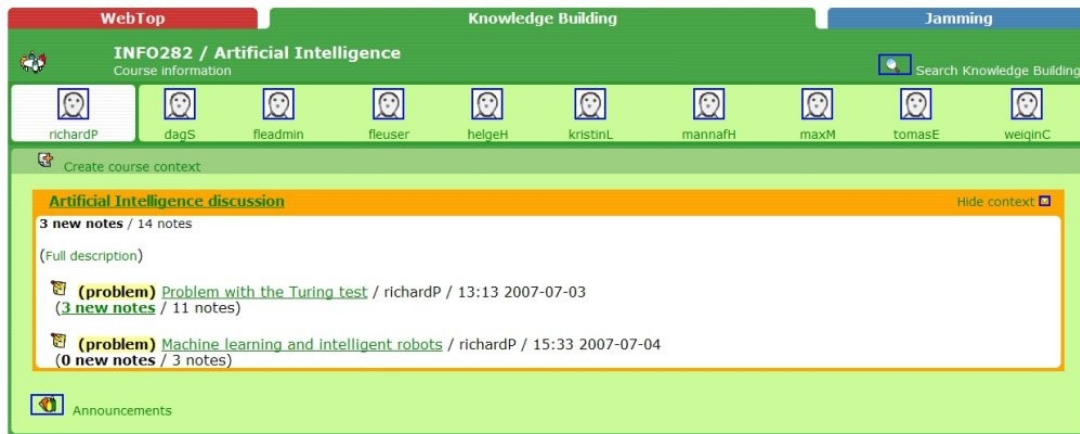
3.4.1 Brukergrensesnittet for relevante innlegg

Brukergrensesnittet for relevante innlegg er den mest sentrale delen av AnnForum fordi det er den funksjonen som vil bli hyppigst brukt av studentene i løpet av et semester. Utgangspunktet for grensesnittet er gruppevaresystemet FLE3 sitt diskusjonsforum: «Knowledge Building». Dette er et helt vanlig trådbasert diskusjonsforum, som vist i figur 3.4.

Forumet er her satt opp for at studentene skal kunne diskutere temaer relatert til

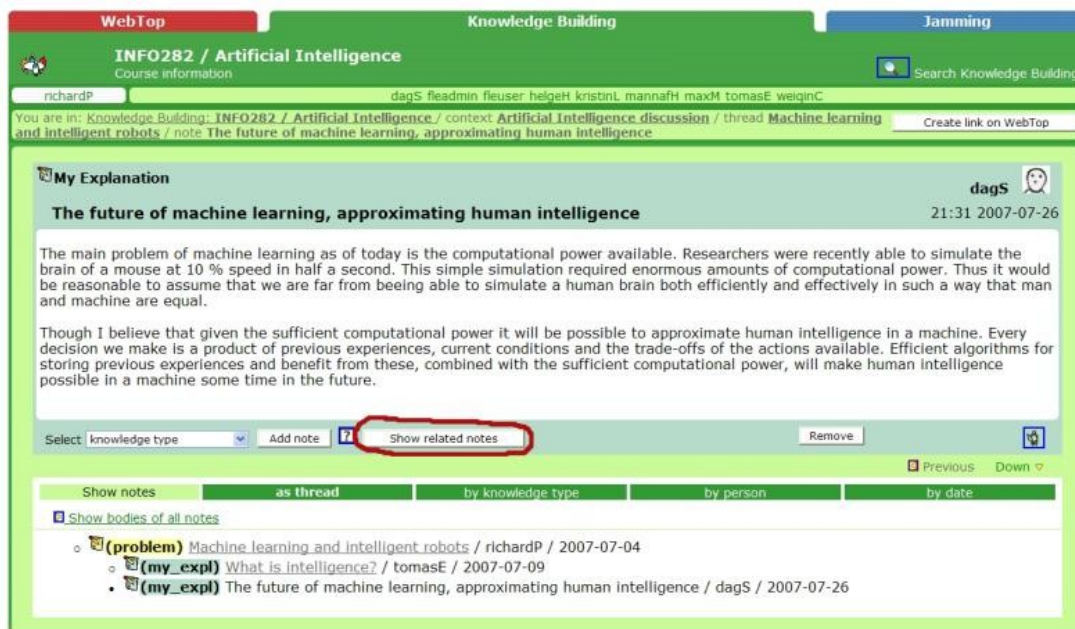
3. Design og utvikling

kunstig intelligens. Det eneste som skiller FLE3 sitt diskusjonsforum litt fra andre diskusjonsforumer, er at brukeren må tilordne et innlegg til en kunnskaps-kategori.



Figur 3.4: Diskusjonsforumet til FLE3

Disse kategoriene er «Problem», «My Explanation», «Scientific Explanation» og «Evaluation of the process», og meningen med disse kategoriene er at det skal være enklere å se hva som er problemområdet for et innlegg.

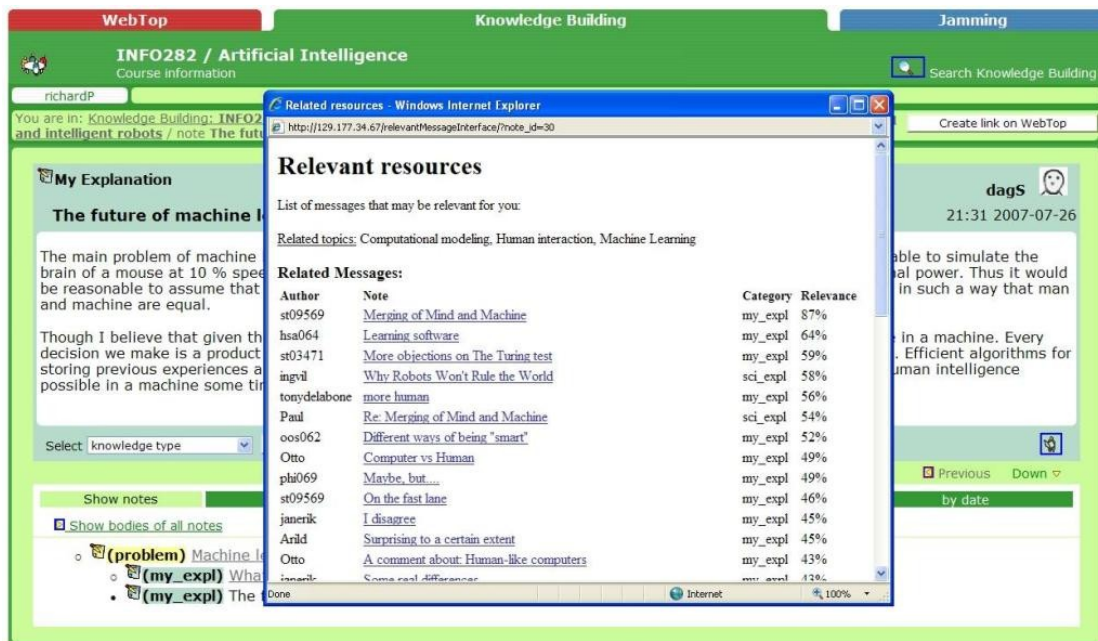


Figur 3.5: Vis relevante innlegg

Grensesnittet er integrert i FLE3, og måten det fungerer på, er at brukeren vil få

3. Design og utvikling

muligheten til å se relevante ressurser når har valgt et innlegg (figur 3.5). Hensikten er at brukeren selv skal kunne velge om de relevante ressursene skal vises. Når brukeren trykker på «Show related notes»-knappen, vil innlegg fra tidligere semesters diskusjoner, som er relevante for den nåværende FLE3-diskusjonen, hentes frem og vises i et nytt vindu, som vist i figur 3.6. Innleggene rangeres etter relevansverdi i prosent og gir informasjon om blant annet hvem som er forfatter av innlegget, en tittel som er en lenke til hele innlegget, samt hvilken FLE3-kategori det hører til under.



Figur 3.6: Brukergrensesnittet for relevante innlegg

I tillegg kan brukeren se hvilke konsepter/emner som innleggene er hentet ifra, i dette tilfellet Computational Modeling, Human Interaction og Machine Learning.

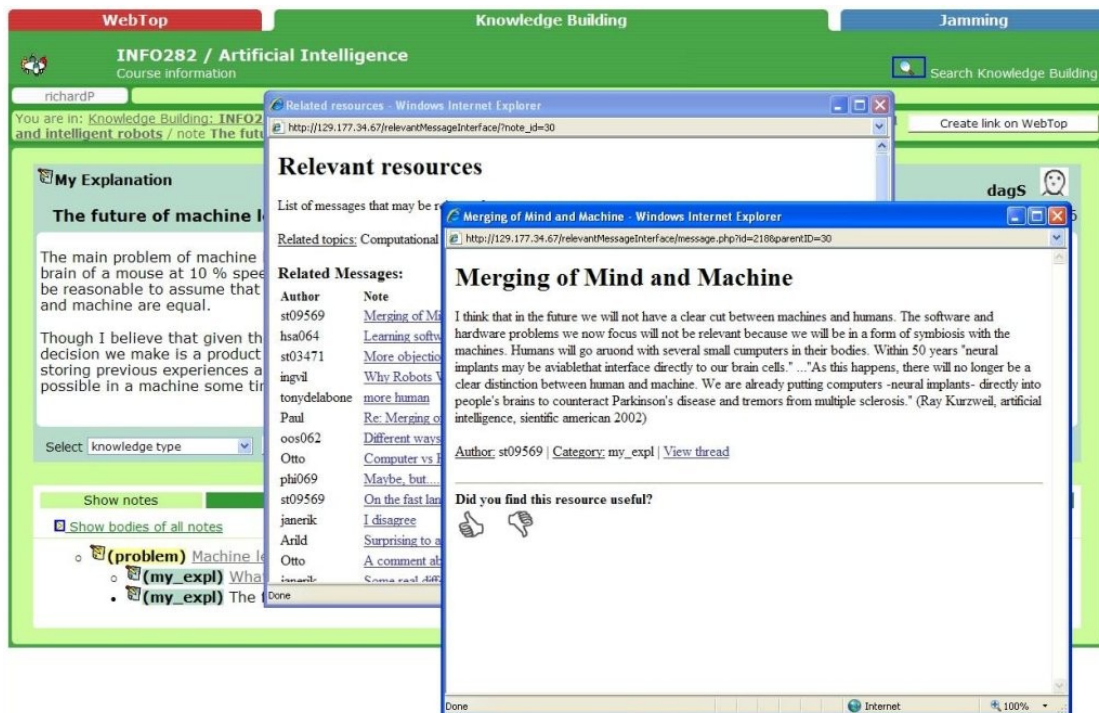
Grensesnittet er skrevet i HTML og PHP og kaller på MySQL-databasen og den konseptuelle domenemodellen for å hente ut relevante innlegg, konsepter og relevansverdier.

Brukergrensesnittet for relevante innlegg er tilgjengelig ikke bare for eksisterende innlegg i FLE3, men også etter at brukeren har skrevet og lagt til et nytt innlegg. Rangeringen av innleggene er basert på en vektlagt term mellom to innlegg som beskrevet i del 3.3.2.4, hvor det ene innlegget er fra nåværende semester, mens det

3. Design og utvikling

andre innlegget er fra et tidligere semester.

Figur 3.7 viser et utvalgt, relevant innlegg. Innlegget vises i sin helhet med tittel, innhold, forfatter og FLE3-kategori. Brukeren har også muligheten til å avgi en stemme for om innlegget er relevant eller ikke ved å trykke på tommel-opp- eller tommel-ned-ikonene. Når en stemme avgis, vil relevansverdien for innlegget økes/reduseres med 50 %. I tillegg var det meningen at brukerne skulle kunne se hele tråden som innlegget hørte til under («View thread»). Men på grunn av tidsmangel ble ikke denne funksjonen implementert. En slik funksjon, når den blir ferdigstilt, vil gi en oversikt over konteksten som det relevante innlegget befinner seg i.

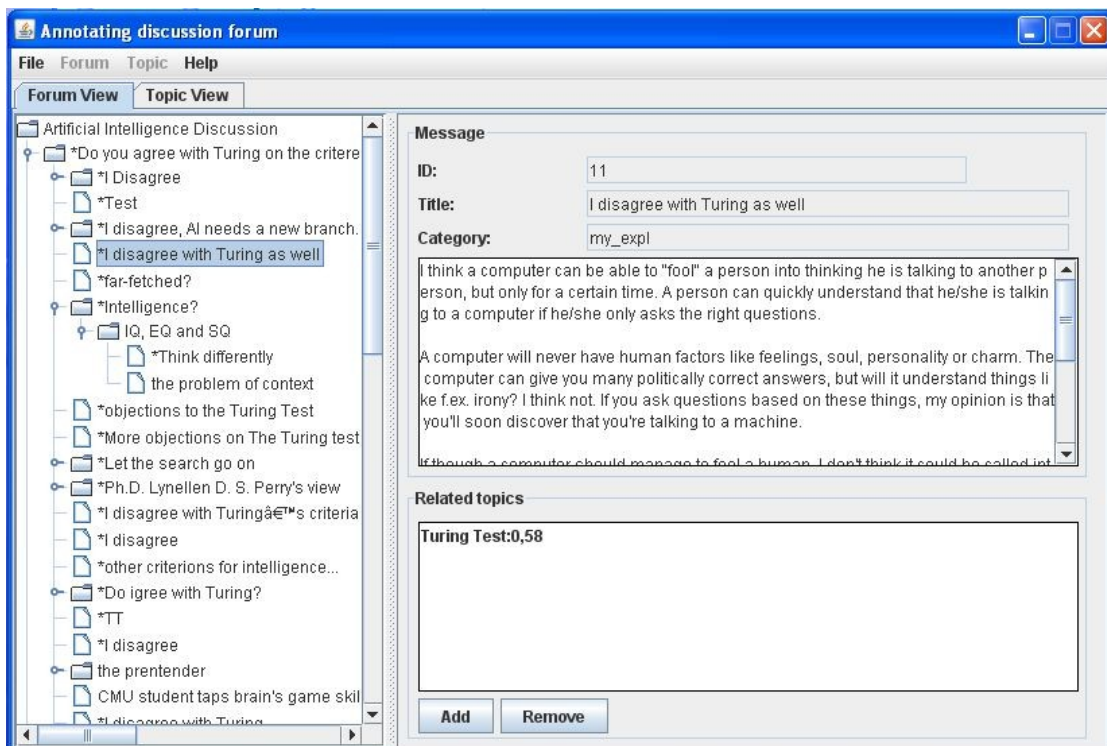


Figur 3.7: Et utvalgt relevant innlegg

3. Design og utvikling

3.4.2 Lærerens annoteringsverktøy

Det har også blitt utviklet et verktøy som gir læreren muligheten til å redigere den konseptuelle domenemodellen i tillegg til å endre på annoteringene for de enkelte innleggene. Nedenfor beskrives de mest sentrale brukergrensesnittene, mer spesifikt Forum View og Topic View.

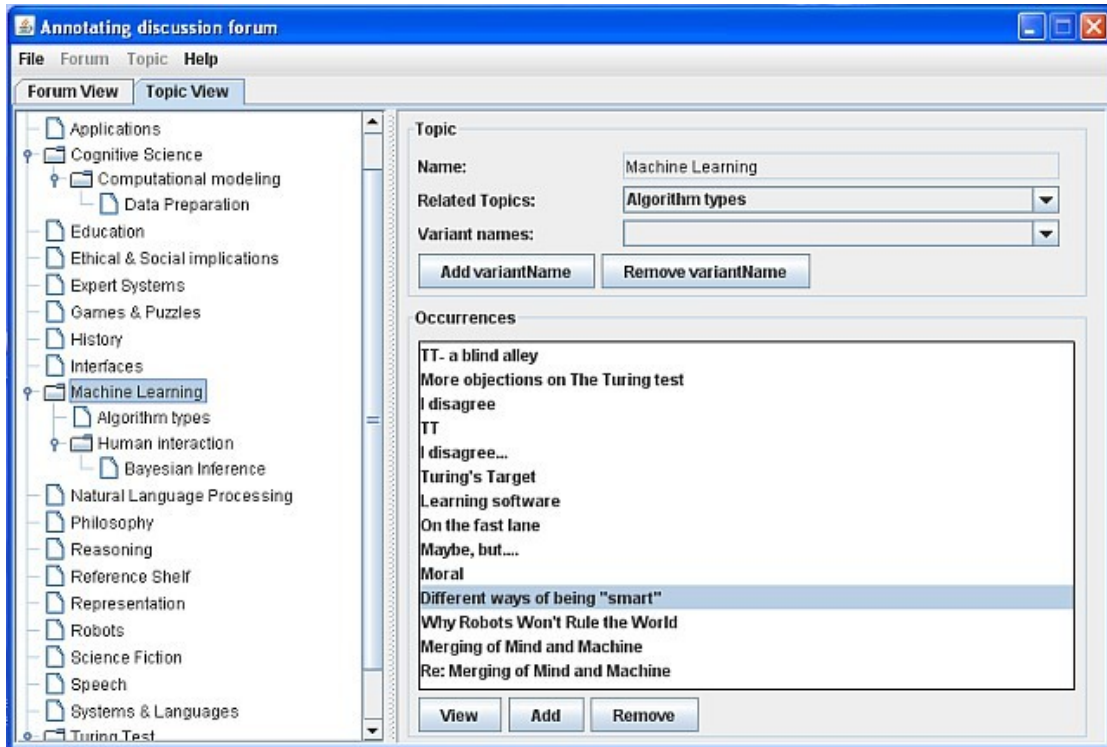


Figur 3.8: Forum View

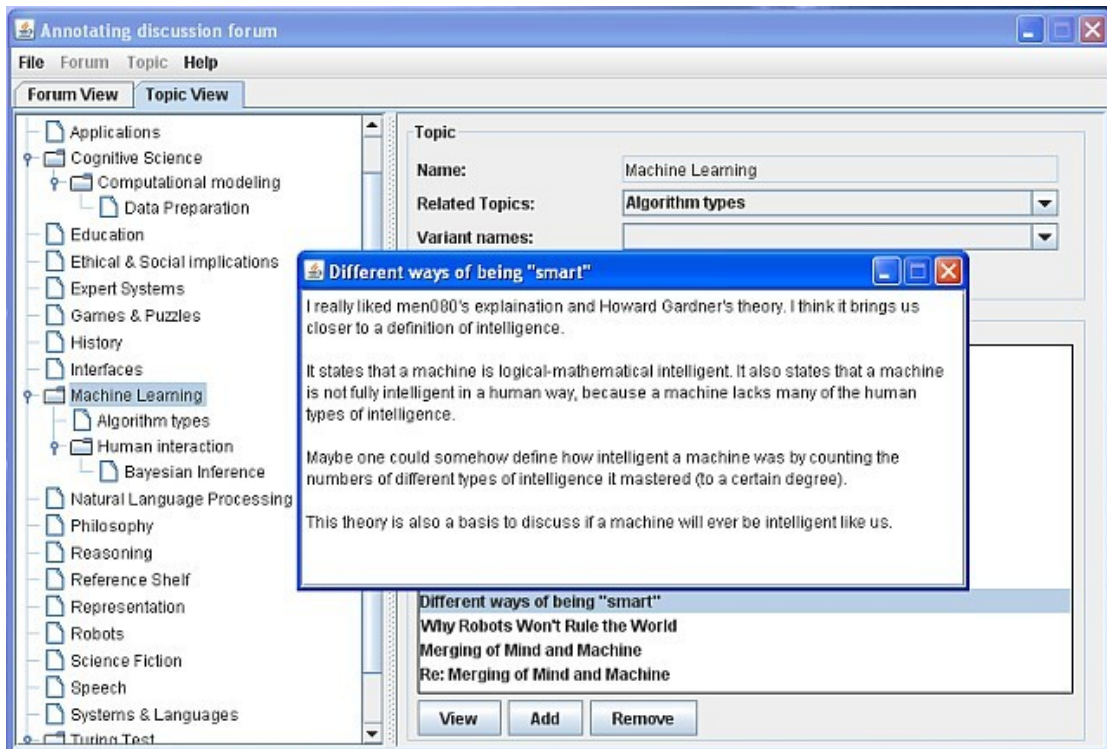
Figur 3.8 viser «Forum View», som er grensesnittet der læreren kan få en oversikt over alle innleggene i et valgt diskusjonsforum. Det valgte diskusjonsforumet kan enten være for det nåværende semesteret eller et tidligere semester. Læreren har her muligheten til å bla gjennom alle innleggene ved hjelp av en trebasert struktur, i tillegg til å se innlegget i sin helhet. For hvert innlegg er det også mulig å se de relevante konseptene med relevansverdi, samt at man kan legge til og fjerne konsepter.

Figur 3.9 viser «Topic View», som er det grensesnittet hvor læreren kan redigere den konseptuelle domenemodellen.

3. Design og utvikling



Figur 3.9: Topic View



Figur 3.10: Et valgt innlegg i Topic View

3. Design og utvikling

Som for Forum View er det også her mulig å bla gjennom domenemodellen ved hjelp av en trestruktur. Trestrukturen representerer konsepter og eventuelle relaterte konsepter, som for eksempel «Algorithm Types» som er relatert til «Machine Learning». Til høyre i grensesnittet ser man detaljene for konseptet: navn, relaterte konsepter og eventuelle alias («variantNames»). Andre funksjoner for Topic View inkluderer muligheten til å legge til og fjerne et synonym for et valgt konsept, samt å legge til og fjerne forekomster/innlegg for et konsept. Det er også mulig å se et innlegg som forekommer under et valgt konsept (figur 3.10), med andre ord et innlegg som har blitt klassifisert under konseptet som et «occurrence».

4. Metode og evaluering

Dette kapittelet beskriver hvordan systemet har blitt evaluert og testet opp mot problemstillingen og er delt opp i tre deler. Den første delen tar for seg selve evalueringsmetoden på et generelt grunnlag, den andre delen beskriver hvordan eksperimentet ble gjennomført i praksis, mens den tredje delen analyserer de data som har blitt samlet inn og oppsummerer resultatet av evalueringen.

4.1 Metode

Systemet ble testet med tanke på brukbarhet og brukervennlighet gjennom å foreta en eksperimentell brukerevaluering med et sett med ekspertbrukere ved Universitetet i Bergen. Denne delen gir en beskrivelse av hva som kjennetegner en eksperimentell evaluering, samt hvordan man skal kan gå frem for å sette opp og gjennomføre slike evalueringer.

4.1.1 Eksperimentell evaluering

En *eksperimentell evaluering* er en empirisk forskningsmetode som blir mye brukt når man ønsker å teste brukergrensesnitt og design med tanke på aspekter som brukbarhet, brukervennlighet og nytteverdi (Dix, 2004:327–351). Til forskjell fra en systemanalyse, som er en metode der designeren selv evaluerer systemet, så bruker man innen en eksperimentell evaluering faktiske brukere som respondenter for å teste designet. Dette er en veldig effektiv teknikk som gir den fordelen at man får sett hvordan systemet fungerer i en virkelig situasjon, samt at man også får muligheten til å spørre brukerne direkte om deres meninger angående konkrete aspekter ved systemet. Særlig innenfor menneske–maskin-interaksjon (Human-Computer Interaction) blir eksperimentelle evalueringer mye brukt, og resultatet fra slike eksperiment gir empiriske data som kan brukes til å støtte opp om en hypotese eller en påstand. For at en eksperimentell evaluering skal gi tilstrekkelig med relevante data, må man tenke nøye gjennom følgende tre faktorer: utvalget av respondenter, hvilke variabler det er som skal testes, samt at man må ha utformet en god hypotese som spesifiserer nøyaktig hva det er man vil undersøke.

4. Metode og evaluering

4.1.2 Elementer i et eksperiment

Det første steget i oppsettet av en eksperimentell evaluering er som nevnt utformingen av en hypotese. Hypotesen er i denne sammenhengen hva det er vi forventer å få ut av eksperimentet, og den spesifiserer i klartekst fokusområdet vårt. Målet med en eksperimentell evaluering er å få samlet inn tilstrekkelig med empiriske, kvalitative data som kan brukes til å bekrefte, eventuelt falsifisere, påstanden vår, og derfor også problemstillingen (Dix, 2004:329).

Neste steg i evalueringen er å velge ut respondentene til eksperimentet. Det å ha et godt utvalg er på mange måter det viktigste punktet da et dårlig utvalg ikke vil gi oss de nødvendige data. Den beste situasjonen er om man har mulighet til å foreta eksperimentet med faktiske brukere av systemet, men om disse ikke er tilgjengelige, bør vi fokusere på å få et utvalg som er så representativt som mulig for den faktiske brukergruppen når det gjelder alder, utdanning, dataferdigheter og så videre. Hvor mange respondenter man skal bruke, avhenger av problemstilling og hypotese, men som regel bør man ha så mange at utvalget blir representativt for den faktiske brukergruppen. Hvis ikke vil man ikke klare å avdekke potensielle problemer på en god nok måte.

Siste steg er å velge ut et antall variabler som kan testes under kontrollerte forhold, slik at man kan få testet hypotesen. Når man har utformet en god hypotese, vet man også ofte konkret hvilke variabler det er som skal testes, og om dette ikke er tilfelle, bør man revurdere hypotesen. Disse variablene må kunne manipuleres slik at man senere kan teste resultatet ved å gjenskape eksperimentet under liknende omstendigheter.

Et eksperimentet kan gjennomføres i form av en strukturert observasjon. Ifølge Grønmo (2004:126) er en slik observasjon en undersøkelsesmetode der evaluator observerer respondentene under kontrollerte forhold. Men det viktigste er at man får testet ut designet ved hjelp av brukere (Dix, 2004:331). Innsamling av data utføres som regel ved hjelp av utspørringer i etterkant av eksperimentet. Disse utspørringene kan for eksempel være strukturerte intervju der forskeren bruker et spørreskjema med et sett predefinerte spørsmål. Intervjuer og respondent kan møtes personlig for å

4. Metode og evaluering

gjennomføre intervjuet, men respondenten kan også selv kan fylle ut spørreskjemaet. Dette kan være fordelaktig dersom man ikke har muligheten å møtes ansikt til ansikt. For å få kvalitative data ut fra eksperimentet kan man bruke en kvalitativ innholdsanalyse til å analysere svarene. Dette er en analyseform der man forsøker å behandle og analysere data sett i forhold til den aktuelle problemstillingen. Men hovedmålet er altså å få samlet inn kvalitative data ved hjelp av observasjon og interjvu. De data som bli samlet inn kan man igjen bruke for å analysere problemstillingen/hypotesen.

Men andre ord kan oppsettet av en eksperimentell evaluering oppsummeres slik:

1. Velg en hypotese som klart spesifiserer hva det er vi ønsker å teste.
2. Velg ut respondenter som er så representative som mulig for den faktiske brukergruppen.
3. Velg ut variabler som kan testes under kontrollerte forhold.

4.2 Gjennomføringen av selve eksperimentet

Denne delen gir en utfyllende beskrivelse av hvordan eksperimentet ble utformet og gjennomført. Først beskrives målet med evalueringen, deretter utvalget av respondenter, og til slutt prosessen med å samle inn data.

4.2.1 Målet med evalueringen

Hovedmålet med evalueringen har vært å teste ut AnnForum med tanke på aspekter som brukervennlighet, nytteverdi og brukbarheten relatert til annoteringen av ikke prefabrikkerte læringsressurser i FLE3 sitt diskusjonsforum. Dette har blant annet innebåret å bruke faktiske studenter som respondenter, for på den måten å se om systemet kunne bidra til å øke graden av læring hos studenter sett i en læringssammenheng. Hypotesen skissert i innledende kapittel har hele tiden vært den faktoren som mest har påvirket oppsettet av evalueringen:

4. Metode og evaluering

Kan semantisk annotering av innlegg i diskusjonsforum hjelpe studenter i deres kollaborative læringsprosess ved at læringsressurser fra tidligere semesters diskusjoner gjenbrukes?

Med andre ord var hensikten med å få relevante tilbakemeldinger fra respondentene å kunne få en idé om nytteverdien av AnnForum og mer spesifikt brukergrensesnittet for relevante innlegg. Et delmål var også å se på hvor godt systemet klarte å annotere innlegg etter de riktige konseptene i den konseptuelle domenemodellen, det vil si hvor godt klassifiseringsalgoritmen fungerte.

4.2.2 Utvalg av respondenter

Seks respondenter, studenter ved Universitetet i Bergen, ble valgt ut til evalueringen av AnnForum. Årsakene til at disse seks ble valgt ut, var at først og fremst at alle hadde bakgrunn fra kunstig intelligens, hovedsaklig ved at de hadde tatt fag ved UiB. Alle respondentene hadde også erfaring med bruk av diskusjonsforumer. Disse to punktene gjorde at det ikke var nødvendig å gi spesielt mye opplæring av verken verktøy eller fagområdet, noe som sparte mye tid. I tillegg var respondentene veldig like faktiske brukere.

De ovennevnte kriteriene for utvalg av respondenter ble satt opp for å få et så representativt utvalg som mulig sammenlignet med en virkelig brukergruppe. Respondentene var alle ekspertbrukere med kunnskap om det aktuelle domenet, noe som også ville vært tilfellet ved en evaluering gjennom et helt semester med faktiske brukere.

Den opprinnelige planen var å få utført evalueringen gjennom et helt semester ved å teste studenter som var aktive deltakere på et kurs i kunstig intelligens, men på grunn av tidsmangel, samt at de aktuelle kursene ikke gikk det semesteret denne evalueringen ble gjennomført, var dette dessverre ikke mulig. Ved å evaluere systemet gjennom et helt semester kunne man ha observert bruken av systemet i en virkelig situasjon, man kunne ha spurt studentene om momenter ved systemet, i tillegg til å ha satt opp loggføringsmekanismer for å sjekke bruken av annoteringsmodulen.

4. Metode og evaluering

4.2.3 Datainnsamling – brukeroppgaver og intervju

Den eksperimentelle evalueringen ble gjennomført i form av seks enkeltstående evalueringer. Hver evaluering ble utført i et kontrollert miljø med kun evaluator og en respondent til stede. Dette ble gjort for at respondenten skulle kunne observeres under selve evalueringen. To av de seks studentene var ikke tilgjengelige i løpet av disse to ukene, men disse ble i stedet tildelt et alternativt evalueringssopplegg som kunne utføres over Internett. Respondentene fikk ikke noe informasjon i forkant av evalueringene, annet enn at varigheten for hver evaluering ville bli på én til én og en halv time, at det ville innebære kunstig intelligens og diskusjonsforumer, og at målet med evalueringen var å teste ut en ny funksjon i FLE3 sitt diskusjonsforum.

De fire evalueringene ble gjennomført i et kontrollert miljø, det vil si et stille rom med kun evaluator og respondent til stede. Hver evalueringssesjon ble satt opp på den måten at respondenten først kort ble fortalt hva som var hensikten med evalueringen: å teste ut en ny funksjon i FLE3 sitt diskusjonsforum for å se om denne kunne forenkle prosessen med å finne informasjon. Respondenten fikk også en kort opplæring/introduksjon til FLE3 sin kunnskapsbyggingsmodul, samt at de ble vist hvor de kunne finne brukergrensesnittet for relevante innlegg (figur 3.5). Når introduksjonen var ferdig, fikk de utdelt et sett med brukeroppgaver som de skulle gjennomføre på egenhånd. De ble også bedt om å si i fra om det var noe de fikk problemer med mens de jobbet med oppgavene. For å få en konkret tilbakemelding på brukbarheten av AnnForum, ble respondentene intervjuet rett i etterkant etter at brukeroppgavene hadde blitt utført. Hensikten med intervjuet var todelt:

1. å få bekreftet resultatet fra observasjonen ved å stille spørsmål om observerte elementer som måtte være uklart.
2. å få en mer nøyaktig tilbakemelding på brukerens holdninger i relasjon til systemet, og da særlig med tanke på aspekter som brukervennlighet, brukbarhet og nytteverdi.

De to respondentene som ikke hadde mulighet til å møte opp personlig, fikk tilsendt evalueringssopplegget på e-post. I tillegg til at brukeroppgavene og

4. Metode og evaluering

intervjuspørsmålene ble lagt ved som vedlegg, ble det også beskrevet nøye hvordan de skulle gå frem for å utføre evalueringen på egen hånd. Dette var for at alle evalueringene skulle bli så like som mulig.

4.2.3.1 Brukeroppgaver

Hensikten med brukeroppgavene var hovedsaklig at respondentene skulle få en oversikt over systemet i forkant av intervjuet, og da spesielt brukergrensesnittet for relevante innlegg. På den måten kunne de gjøre seg opp en mening angående brukbarheten og nytteverdien av AnnForum før intervjuspørsmålene. Før den første evalueringen ble det lagt inn to «hovedinnlegg» relatert til kunstig intelligens, hvor et av innleggene omhandlet Turing Test og et omhandlet Machine Learning. Disse to innleggene var utformet som «Problem» i FLE3 og ble brukt som utgangspunkt for alle evalueringene. Formuleringene på disse to innleggene var så enkle som mulig, slik at respondentene kunne lese og svare på disse uten å måtte ha inngående kunnskap om temaene.

Opgavene som respondentene skulle utføre, innebar blant annet at de skulle skrive to poster i FLE3 som svar på eksisterende innlegg, i tillegg til å benytte brukergrensesnittet for relevante innlegg for å finne relevante ressurser. Se appendiks C for en oversikt over brukeroppgavene. Respondentene ble også bedt om å avgi stemmer på om de syntes innleggene de leste, var relevante eller ikke. Svarene som ble lagt til i FLE3 av respondentene under hver evaluering, ble liggende, slik at neste respondent skulle få flere innlegg å velge i. Begrunnelsen for dette valget var et forsøk på å simulere en kunnskapsbyggingsprosess, slik at den første respondenten hadde to mulige innlegg å svare på, andre respondent hadde fire mulige innlegg og så videre. Denne fremgangsmåten gjorde også at testingen av den dynamiske klassifiseringen av nye innlegg ble mer realistisk, fremfor om man for hver evaluering skulle ha startet med en tom FLE3-database.

4.2.3.2 Datainnsamling – intervju

Datainnsamlingen ble gjort ved å gjennomføre strukturerte intervju av de seks respondentene. Intervjuene ble gjort etter at brukerne hadde utført brukeroppgavene,

4. Metode og evaluering

og spørsmålene som ble stilt, var direkte relatert til nytteverdien og brukbarheten av AnnForum. For å få en oversikt over bakgrunnskunnskapene ble respondentene også spurt om hvor mye kunnskap de følte at de hadde om kunstig intelligens, samt om de hadde brukt FLE3 i en utdanningssammenheng før. For noen av spørsmålene ble de også bedt om å rangere svaret sitt, for eksempel «På en skala fra 1–5 der 5 er mest relevant, hvor relevant?».

Spørsmålene kan grovt fordeles i fire hovedkategorier:

- Hensikten med brukergrensesnittet for relevante innlegg og relevansen til de foreslåtte innleggene ved første øyekast, sett ut fra tittel og relevansverdi i prosent (figur 3.6).
- Relevansen for de innleggene som faktisk ble lest – hvor relevante syntes respondenten innlegg(ene) virket for den nåværende diskusjonen.
- Positive og negative aspekter ved systemet samt forslag til endringer.
- Nytteverdi – respondentens oppfatning og subjektive meninger angående systemet.

Med *relevansverdi* menes det hvor relevante respondenten syntes at de foreslåtte innleggene som kom opp, var for den nåværende diskusjonen. Med *positive/negative aspekter* menes det om det var noe respondenten syntes var særdeles bra/dårlig med brukergrensesnittet som helhet, mens *nytteverdi* refererer til om respondenten syntes at et system som annoterer og gjenbraker innlegg i diskusjonsforumer, har noe for seg.

Et utdrag av spørsmålene som ble stilt, er listet opp nedenfor. For en fullstendig oversikt over alle intervjuene med tilhørende spørsmål, se appendiks D.

- Skjønte du hensikten med «Show related notes»-knappen?
- Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

4. Metode og evaluering

- Hva var mest interessant ved de foreslåtte innleggene?
- Synes du relevansverdien i prosent var riktig for de innleggene du leste?
Hvorfor?

Et par dager etter at intervjuene var gjennomført, fikk respondentene tilsendt sitt transkriberte intervju på e-post for at de skulle kunne se over svarene sine, og eventuelt gjøre endringer dersom det var noe som de mente var galt. Det å gi respondentene denne muligheten ble gjort for at intervjuene skulle bli så korrekte som mulig. I tillegg var tanken at respondentene etter et par dager ville ha fått litt tid til å tenke over svarene sine, for på den måten å bedre kunne utdype informasjonen om de skulle ønske det. Alle respondentene valgte å benytte seg av denne muligheten ved å presisere formuleringer og endre på svar som de mente var uklare.

4.3 Analyse og oppsummering av data

Denne delen analyserer og oppsummerer de data som ble samlet inn i løpet av de seks evalueringene. Hver av de seks intervjuene vil analyseres med tanke på relevansverdi for de foreslåtte innleggene («Show related notes»), positive aspekter, negative aspekter og nytteverdi. I tillegg vil resultatene fra alle intervjuene oppsummeres og sammenliknes mot slutten av kapittelet. Transkripsjon av alle intervjuene finnes i sin helhet i appendiks D.

4.3.1 Første evaluering

Respondent nummer én var stort sett positiv til systemet og svarte at hensikten med brukergrensesnittet for relevante innlegg var å kunne se hva andre mente om et valgt tema. Innleggene som ble listet opp, virket relevante, hovedsaklig som følge av at titlene på de foreslåtte innleggene omhandlet temaet («Turing Test»).

Relevansverdien i prosent for de forskjellige innleggene var også nyttig, fordi de ga et inntrykk av hvor relevante ressursene var (høyere verdi = større relevans).

Om de innleggene som faktisk ble lest, svarte respondenten at de til dels var relevante. Ikke alle de foreslåtte innleggene omhandlet Turing Test direkte, selv om de kunne

4. Metode og evaluering

inneholde relevant informasjon om temaet. Prosentverdien for innleggene var stort sett riktig, innlegg med høy relevansverdi var stort sett nyttige. De leste innleggene ble valgt på bakgrunn av både relevansverdi og tittel, og innlegg som var bra skrevet, ble rangert positivt. Respondenten unnlot imidlertid å rangere dårlige innlegg negativt, dette for ikke å foreta en negativ rangering basert på subjektive meninger.

I et velskrevet innlegg har man gjerne mye fakta og man knytter det opp mot sine meninger. For mye subjektivitet i et innlegg kan lett oppfattes som mindre relevant enn et faktabasert innlegg. Dermed er det ikke sagt at de subjektive innleggene er irrelevante.

Det som var mest positivt ved de relevante innleggene, var ifølge respondenten muligheten for å få en mer utdypende kunnskap om det valgte temaet, men også å kunne se hva andre har ment om temaet, samt å kunne bruke innleggene som referanser. Et negativt aspekt ved systemet var at stemmegivningsfunksjonen kunne misbrukes ved at det var mulig å stemme sitt eget synspunkt høyt opp på listen. Det kunne også ha vært nyttig å knytte de enkelte innleggene til referanser for kunstig intelligens, som for eksempel vitenskaplige artikler.

På spørsmålene om nytteverdien av å vise relevante innlegg svarte respondenten at et slikt system i utgangspunktet virket som en god idé, men at det nok ville være mest til nytte i situasjoner der brukerne var høyt motiverte, som for eksempel en studiesituasjon. Innenfor offentlige eller mer overfladiske diskusjoner vil kanskje ikke brukerne være motiverte nok til benytte funksjonen. Det ble kommentert at brukergrensesnittet var oversiktlig.

4.3.2 Andre evaluering

Respondent nummer to var positiv til systemet og svarte at hensikten med brukergrensesnittet for relevante innlegg var å tolke innholdet i det nåværende innlegget, samt å hente frem andre innlegg som hadde lignende innhold. Han svarte at flesteparten av de relevante innleggene som ble foreslått, var nyttige for diskusjonen, og at innleggenes overskrifter raskt viste hvorfor systemet mente de var relevante. Relevansverdien i prosent var imidlertid det som var viktigst da prosentverdien ga et

4. Metode og evaluering

inntrykk av hvor mye man kunne forvente av innlegget.

Om de innleggene som faktisk ble lest i sin helhet, svarte respondenten at de som oftest var meget relevante, men med unntak av et testinnlegg som kom ganske høyt opp på listen. Innleggene ble hovedsaklig lest på bakgrunn av relevansverdi. De relevante innleggene påvirket også svarene som respondenten la inn i FLE3, ved at de ga et inntrykk av at det var et aktivt og levende forum, noe som økte motivasjonen for å formulere et godt svar. Men det var også noen innlegg som ble lest, som ikke var relevante.

Det som var mest positivt ved de relevante innleggene, var ifølge respondenten at «[...] de virket reelt relaterte og inneholdt ikke bare det samme ordet et par ganger som slike systemer vanligvis synes å resultere i.» En ting som ble kommentert som negativt, var at det ved et tilfelle kom et testinnlegg øverst på listen. Respondenten sa også at det kunne vært nyttig med en liten ingress til hvert av de relevante innleggene som ble listet opp, slik at det kunne vært lettere å se hva de handlet om.

På spørsmålene om nytteverdien av å vise relevante innlegg svarte respondenten at et slikt annoteringssystem hadde et potensial til å bli mye brukt, ikke bare i en studiesituasjon, men også på for eksempel sosiale nettsteder.

4.3.3 Tredje evaluering

Respondent nummer tre var også positiv til systemet og svarte at hensikten med brukergrensesnittet for relevante innlegg var å vise relevante innlegg som handler om samme tema, samt hvor relevante de var for det valgte FLE3-innlegget. Innleggene som kom opp i listen, virket relevante da de omhandlet det teamet som det ble lest om.

Om de innleggene som faktisk ble lest, svarte respondenten at de stort sett virket relevante for diskusjonen, selv om kvaliteten på innleggene ikke nødvendigvis var like bra. Innleggene ble valgt på bakgrunn av relevansverdi, de med høy rangering ble lest først, men også noen av de med lav rangering ble lest for å kunne sammenlikne. De med høy relevansverdi syntes respondenten hadde høy faglig kvalitet og mest

4. Metode og evaluering

tilknytting til temaet, og de ble derfor rangert positivt.

Det som var mest positivt ved de relevante innleggene, var ifølge respondenten å kunne se hva andre hadde diskutert og sagt om et tema. Det var ikke noe spesielt som ble trukket frem som negativt, men det ble kommentert at innleggene med lavest relevansverdi var minst interessante.

På spørsmålene om nytteverdien av å vise relevante innlegg ble det svart at det virket fornuftig å gjenbruke gamle innlegg fordi «[...] det er dumt at det samme skal skrives om igjen flere ganger, og man kan også finne lærdom i gamle innlegg.» Respondenten sa også at hun nok selv ville ha brukt en slik funksjon i en virkelig studiesituasjon.

4.3.4 Fjerde evaluering

Respondent nummer fire var positivt innstilt til systemet og svarte at hensikten med brukergrensesnittet for relevante innlegg var å vise innlegg som var relevante for nåværende innlegg. De innleggene som ble listet opp, var nyttige for diskusjonen da de aller fleste hadde titler som handlet om det aktuelle temaet.

Om de innleggene som faktisk ble lest, svarte respondenten at også disse for det meste var relevante, fordi «de innleggene med høy relevansverdi faktisk var mer relevante for diskusjonen, mens de med lav relevansverdi ofte ikke var så godt skrevne / veldig relevante». Det var hovedsaklig de innleggene med høy relevansverdi, og litt i kombinasjon med interessant tittel, som ble lest. Et par innlegg med lav relevans ble lest for å sammenlikne. De relevante innleggene påvirket også til dels svarene som respondenten selv la inn i FLE3.

Det som var mest positivt ved de relevante innleggene, var ifølge respondenten muligheten for avgi en stemme for om et foreslått innlegg var nyttig eller ikke. Det var også lettere å kunne navigere i en sortert liste fremfor et hierarki. En ting som imidlertid ble trukket frem som negativt, var at de relevante innleggene kom opp i et popup-vindu; en bedre løsning hadde vært om listen hadde vært direkte integrert i FLE3. Brukergrensesnittet for relevante innlegg kunne også hatt litt samme «look-and-feel» som grensesnittet til FLE3.

4. Metode og evaluering

På spørsmålene om nytteverdien av å vise relevante innlegg var respondenten positiv og svarte at en slik funksjon nok ville ha blitt brukt i en virkelig studiesituasjon fordi: «Jeg selv ville ha brukt den, og jeg er lat. Funksjonen er lettere tilgjengelig og mer nyttig enn den tradisjonelle SEARCH-knappen».

4.3.5 Femte evaluering

Respondent nummer fem forstod hensikten med systemet, men utdypet ikke dette noe videre. For de innleggene som ble listet opp så svarte han at de fleste var relevante, men at det kanskje var litt for lite brukerdata til å korrekt kalkulere relevans, blant annet fordi at innlegg som ble gitt en positiv stemme gikk fra langt nede på listen til høyt opp.

For de innleggene som faktisk ble lest så svarte respondenten disse var relevante, og at med flere innlegg i databasen så kunne det blitt enda mer interessant. Relevanseverdien i prosent virket også riktig for innleggene. Det som avgjorde hvilke foreslåtte innlegg som ble lest var en kombinasjon av både relevanseverdi og tittel. Av de relevante innleggene som ble lest så påvirket ikke de svarene som respondenten selv utformet og la til i FLE3.

Det som var mest positivt ved de relevante innleggene var i følge respondenten å kunne lese andres vinklinger på samme sak. Ingen spesielle aspekter ble trekt frem så negativt, men det ble kommentert at det kunne ha vært nyttig å se mer av innleggene som kom i listen, dette da det ikke alltid er slik at tittelen på innlegget reflekterer innholdet. Det ble også nevnt at relevanseverdien kanskje kunne bli litt upålitelig når få har stemt.

På spørsmålene om nytteverdien av å vise relevante innlegg så svarte respondenten at det virket som en fornuftig idé, da mye kunnskap går tapt mellom hvert kull, og at han selv nok ville ha brukt en slik funksjon. I tillegg så svarte han at det «...ofte finnes relevante innlegg, men på grunn av en enorm informasjonsmasse blir det nesten umulig å finne frem til det man ønsker».

4. Metode og evaluering

4.3.6 Sjette evaluering

Respondent nummer seks var positivt innstilt til systemet og svarte at hensikten med brukergrensesnittet for relevante innlegg var å søke etter tidligere innlegg basert på nøkkelord. De innleggene som ble listet opp i brukergrensesnittet, virket relevante, selv om det ikke fantes informasjon om undertråder og titlene på innleggene var dårlige.

Om de innleggene som faktisk ble lest, så syntes respondenten at innholdet her som oftest var bedre enn tittelen, og relevansverdien var bra nok. Innlegg ble valgt ut på bakgrunn av både tittel og relevansverdi, men det var tittelen som var viktigst. Når det kom til å avgi stemmer, ble ett av innleggene gitt en positiv stemme og ett ble stemt på som ikke relevant. De innleggene som ble lest, hadde ikke noen påvirkning på svarene som respondenten selv la inn, da han av gammel vane så på de foreslåtte innleggene først etter at svarene var lagt til.

Det som var mest positivt ved de relevante innleggene, var ifølge respondenten at det var et godt alternativ til å søke, da det på en måte fungerte som et «automatisk søk», samt at det også ble vist mer informasjon om et spesifikt tema. En ting som ble oppgitt som negativt, var at stemmefunksjonen kunne misbrukes ved å stemme sine egne synspunkt høyt opp på listen, og at det kanskje burde vært en sperre mot å stemme mer enn én gang på samme innlegg. Respondenten hadde flere forslag til endringer/fremtidig funksjonalitet; et av disse var at systemet burde oppfordre brukeren til å skrive mer relevante titler på innleggene. Et annet forslag var muligheten for å se hele tråden til et utvalgt relevant innlegg. Det siste var at brukergrensesnittet for relevante innlegg burde hatt en «look-and-feel» som var mer likt FLE3 sitt brukergrensesnitt.

På spørsmålene om nytteverdien av å vise relevante innlegg svarte respondenten at det virket meget nyttig, spesielt for generelle spørsmål, da man «[...] unngår å få dupliseringer av samme type spørsmål/svar. De relevante innleggene gjør deg og oppmerksom på lignende tema.». Det ble også svart at det er dumt at informasjon går tapt, enten når et nytt semester begynner, eller ved at en tråd med nyttig informasjon forsvinner i mengden når nye innlegg blir lagt til. Respondenten mente også at

4. Metode og evaluering

systemet ville bli brukt i en virkelig studiesituasjon, blant annet fordi diskusjonsforumer for læring kan bli litt mangelfulle da det «...nødvendigvis ikke er så mange tråder som startes i løpet av et semester [...]» .

4.3.7 Oppsummering av resultatet

Resultatet fra den eksperimentelle evalueringen var meget positiv: Av de seks respondentene som var med og evaluerte AnnForum, forstod alle sammen hensikten med systemet, i tillegg til at alle seks så nytteverdien av å ha et brukergrensesnitt som viser relevante innlegg i et diskusjonsforum. Alle respondentene hadde tatt kurs relatert til kunstig intelligens ved Universitetet i Bergen, og på en skala fra én til fem lå snittet på rundt tre. Ingen av respondentene hadde brukt gruppevaresystemet FLE3 i en praktisk situasjon før, men et par stykker hadde lest om det i forbindelse med kurs de hadde tatt. Den første respondenten sine brukeroppgaver og intervju spørsmål varierte også litt fra de andre fem sine spørsmål. Grunnen til dette var at det ble avdekket et par momenter ved oppgavene som var uklare, samt at noen av spørsmålene var litt tvetydige. For eksempel var det ikke hensiktsmessig at respondenten skulle svare på sitt eget spørsmål. Men forskjellene var ikke veldig store og bør derfor ikke påvirke resultatet i nevneverdig grad.

4.3.7.1 Førsteintrykket av de foreslåtte innleggene

For innleggene som først ble listet opp når brukeren trykket på «Show related notes»-knappen, var det generelle førsteinntrykket at innleggene virket relevante for diskusjonen. Alle respondentene svarte at innleggene virket relevante hovedsaklig på grunn av rangeringen og relevansverdien i prosent, men også på grunn av at titlene på innleggene omhandlet det teamet som ble diskutert. En av respondentene sa imidlertid at han synes titlene på innleggene var dårlige, mens en annen stilte spørsmål om måten relevansverdiene i prosent ble regnet ut på. Det ble også uttrykt et ønske om å kunne se subtråder til de enkelte innleggene.

4.3.7.2 Relevansen for de innleggene som faktisk ble lest

De innleggene som respondentene faktisk gikk inn på og leste, var i de aller fleste

4. Metode og evaluering

tilfellene relevante for diskusjonen. Om hva som avgjorde hvilke innlegg som ble lest, svarte tre av respondentene at det var relevansverdien i prosent som var viktigst, to av respondentene svarte at det var en interessant tittel i samsvar med relevansverdien i prosent som avgjorde valget, mens én respondent oppga at det var tittelen som var viktigst. Alle sammen syntes at innleggene som var rangert øverst, var mest relevante, mens innleggene som var på bunnen av listen, som regel var mindre aktuelle. Men det ble kommentert at man kunne sammenlikne de mindre aktuelle med de høyere på listen med tanke på kvalitet. De foreslåtte innleggene som ble lest, påvirket også i to av tilfellene måten respondentene formulerte svarene sine på, to stykker svarte at innleggene ikke påvirket svarene, mens de to siste ikke nevnte dette. Flere av respondentene benyttet seg av muligheten til å stemme på de foreslåtte innleggene, selv om de hadde forskjellige argumenter for hvorfor de stemte som de gjorde. Det som er verdt å merke seg, er at noen innlegg ble rangert positivt, men ingen ble rangert negativt. Den første respondenten sa at dette ble gjort «[...] for ikke å rangere et innlegg negativt basert på subjektive meninger [...]».

4.3.7.3 Positive og negative aspekter – forslag til endringer

Det som generelt ble trukket frem som mest positivt ved systemet, var muligheten for å kunne få en mer utdypende kunnskap om et aktuelt tema, samt å se hva andre hadde skrevet. Innleggene virket også relevante, samt at de ga et inntrykk av at det var et levende diskusjonsforum man var inni, noe som blant annet økte motivasjonen for å formulere gode svar. Stemmegivningsfunksjonen falt også i smak hos en av respondentene. Noe som ble kommentert flere ganger, var den rangerte listen med innlegg. Respondentene likte denne, blant annet da det var et godt alternativ til å søke i diskusjonsforumet. Det var også bedre å kunne navigere i en sortert liste fremfor å bla i et hierarki av innlegg.

Når det kom til negative aspekter, ble det kommentert to ganger at stemmegivningsfunksjonen kunne misbrukes til å stemme sitt eget synspunkt høyt opp på listen. En annen ting som også ble kommentert som å telle negativt, var at et testinnlegg hadde havnet høyt opp på listen. En av respondentene likte heller ikke at de relevante innleggene kom opp i popup-vindu, og syntes heller at disse burde ha

4. Metode og evaluering

vært integrert i FLE3 for ikke å forstyrre arbeidsflyten.

Respondentene hadde også flere forslag til endringer. Dette var relatert til ting som ikke var direkte negative, men som de enten savnet eller kunne tenke seg å se i en eventuell ny versjon. Forslagene var:

- å ha muligheten for å se mer av innleggene i listen av foreslåtte innlegg, på en måte som en slags ingress eller introduksjon. Dette ble begrunnet med at det ofte ikke er lett å se hva et innlegg handler om bare på bakgrunn av en tittel.
- at brukergrensesnittet for relevante innlegg burde hatt litt mer samme «look-and-feel» som grensesnittet til FLE3.
- å ha muligheten for å lenke de enkelte innleggene opp mot referanser for temaet, som for eksempel vitenskaplige artikler.
- å ha muligheten for å se hele tråden til et utvalgt, relevant innlegg.

4.3.7.4 Generell nytteverdi av systemet

Alle respondentene så nytteverdien av å gjenbruke innlegg fra tidligere semester som relevante ressurser i nåværende semesters diskusjon. Hovedgrunnen til at de likte ideen, var at det ville føre til mindre duplisering av informasjon, noe som er et vanlig problem i tradisjonelle diskusjonsforumer. En annen ting som flere av respondentene så på som positivt, var at brukergrensesnittet for relevante innlegg kunne brukes som et alternativ til søkefunksjonen, blant annet fordi de syntes det var lettere å navigere i en sortert liste fremfor å måtte søke eller navigere i en hierarkisk tråd. Generelt syntes de at det var lettere å få oversikt over relevant informasjon og andres synspunkter, i tillegg til at relevante innlegg ikke ville gå tapt fra semester til semester.

Respondentene hadde litt forskjellige oppfatninger om for hvilke bruksområder det ville være nyttig å gjenbruke innlegg. Noen av respondentene mente at et slikt system ville være mest nyttig i situasjoner der brukerne er høyt motiverte, som for eksempel studiesituasjoner, og at i mer generelle diskusjoner ville ikke brukerne være motiverte nok til å bry seg med en slik funksjon. Andre igjen mente at også på sosiale nettsteder

4. Metode og evaluering

kunne systemet være nyttig. Noen av utsagnene angående nytteverdi har blitt listet opp nedenfor:

- «Det er dumt at det samme skal skrives om igjen flere ganger, og man kan også finne lærdom i gamle innlegg.»
- «Det er lettere med en liste enn å måtte navigere i et hierarki, og sorteringen gjorde det lettere å unngå unyttige innlegg.»
- «Dette er svært nyttig i et diskusjonsforum. Ofte finnes relevante innlegg, men på grunn av en enorm informasjonsmasse blir det nesten umulig å finne frem til det man ønsker.»
- «[...] det er synd at viktig informasjon går tapt fra semester til semester. Eventuelt at en diskusjonstråd med nyttig informasjon blir ”begravd” fordi den har vært inaktiv.»
- «I offentlige og/eller overfladiske diskusjoner vil brukerne kanskje ikke være motiverte nok til å bruke en slik funksjon.»

4.3 Diskusjon av resultatet

Resultatet fra evalueringen visste at respondentene var fornøyde med funksjonaliteten i systemet som helhet. De likte muligheten til å kunne få en oversikt over innlegg fra tidligere semester, og syntes i flere tilfeller at det å kunne se relaterte innlegg var et godt alternativ til å søke etter innlegg. Brukergrensesnittet for relevante innlegg var oversiktlig, og ressursene som kom opp, var i de fleste tilfellene relevante for den nåværende diskusjonen. Dette kan tyde på at det å annotere innlegg i diskusjonsforumer kan være nyttig, i hvert fall når det kommer til en læringssituasjon.

Den automatiske annoteringen av innleggene fungerte derfor meget bra, og en av grunnene for dette kan være at grensesnittet og annoteringen av innleggene var enkle å bruke. Funksjonen førte ikke til mer arbeid for respondentene, men lå heller tilgjengelig slik at de kunne se relevante innlegg dersom de ønsket det. Dette henger blant annet sammen med det som ble nevnt i innledningen og i kapittel 2.3.1: Dersom annoteringsprosessen medfører merarbeid for brukeren, for eksempel ved at de må

4. Metode og evaluering

bruke tid på å manuelt tilordne innlegg til konsepter, så vil funksjonen mest sannsynlig ikke bli brukt. En annen fordel er at innlegg som kan være nyttige fra tidligere semester ikke forsvinner, men blir tilgjengelige for nye studenter som nye ressurser. Dette innebærer igjen at man sparer tid da man slipper å formulere spørsmål og svar som allerede har blitt stilt.

Som allerede nevnt i del 4.2.2 var den opprinnelige planen å gjennomføre evalueringen over et helt semester opp mot faktiske studenter. Dersom det hadde blitt gjort kunne nok resultatet ha blitt annerledes. Man bør nok derfor i fremtiden forsøke å få testet ut systemet i et ideell situasjon, hovedsaklig for å se om dette påvirker resultatet. Men frem til man får gjennomført en slik evaluering så virker det som om systemet kan ha en positiv innvirkning på studentenes læringsprosess.

5. Konklusjon

5. Konklusjon

Målet med denne oppgaven har vært å videreutvikle AnnForum som et verktøy for semantisk annotering av innlegg i et diskusjonsforum, for så å teste hvorvidt verktøyet kunne være til nytte for studentene i deres kollaborative læringsprosess. Hypotesen var at et slikt system ville gjøre det enklere for studentene å få tilgang til informasjon som har blitt diskutert i tidligere semester, og dermed gjenbruke denne informasjonen fremfor å starte på nytt hvert semester. AnnForum ble deretter evaluert og testet ved å gjennomføre en eksperimentell evaluering med seks studenter ved Universitet i Bergen.

Hovedsaklig var det teori fra semantisk web og datastøttet læring som ble brukt som utgangspunkt for utviklingen av systemet, og da særlig med vekt på annoteringer, ontologier og domenekart. Det ble gitt en beskrivelse av svakhetene ved tradisjonelle systemer for kunnskapshåndtering (blant annet diskusjonsforumer), hvor en underliggende mangel på konseptuell struktur flere ganger ble dratt frem som det største problemet. I tillegg ble lignende applikasjoner for annotering og klassifisering av informasjon diskutert, samt at det ble argumentert rundt fordelene og ulempene rundt ved disse systemene.

5.1 Tilbakeblikk på evalueringen

Resultatet fra de seks evalueringene var som nevnt meget positivt og viste at respondentene likte å ha tilgang til læringsressurser fra tidligere semesters diskusjoner. Grunnen for dette var at det ble lettere å finne mer utfyllende informasjon om et tema, noe som resulterte i redusert tidsbruk, i tillegg til at de syntes det var interessant å se hva tidligere studenter hadde skrevet. Særlig likte respondentene å få de relevante innleggene opp i en rangert liste da dette med en gang ga et inntrykk av hvilke innlegg som var mest relevante. Et par av respondentene sa også at en rangert liste over relevante innlegg kunne være et godt alternativ til å foreta et tradisjonelt tekstbasert søk. Muligheten for å stemme på om et innlegg var positivt eller negativt ble også stort sett godt mottatt.

5. Konklusjon

5.2 Tilbakeblikk på utviklingen

Kombinasjonen av en automatisk klassifisering og en konseptuell domenemodell viste seg å fungere bra. Respondentene ga uttrykk for at innlegg som kom høyt opp på listen, virket mer relevante enn de lenger nede, noe som kan tyde på at et system som AnnForum er riktig vei å gå i en semantisk web-kontekst. Den dynamiske annoteringen av nye innlegg viste seg også å fungere meget godt: Nye innlegg ble plassert under de riktige konseptene stort sett i alle tilfeller, og relevansverdien virket riktig (da med utgangspunkt i respondentene sine tilbakemeldinger, samt manuell inspeksjon av relevansverdiene i databasen). Annoteringsprosessen forstyrret heller ikke respondentene sin arbeidsflyt da klassifiseringen skjedde automatisk, noe som er en forbedring i forhold til Helic sitt system (Helic et al., 2003).

AnnForum kan nok også være aktuelt i forbindelse med Knowledge Management systemer sett i en organisasjonssammenheng, som for eksempel en intern-wiki. Dette krever selvsagt at AnnForum tilpasses det aktuelle Knowledge Management systemet; men implikasjonene er at ny informasjon vil bli dynamisk/automatisk annotert i henhold til en domenemodell for organisasjonen, noe som igjen kan forbedre søkemekanismene ved at man får frem relaterte ressurser sortert etter relevansverdi.

5.3 Forbedringer og fremtidig arbeid

Selv om det kom frem veldig mange positive aspekter gjennom evalueringen, ble det identifisert et par mangler og designmessige feil, samt at det også ble gitt flere forslag til forbedringer. En av de viktigste tingene som kom frem, var at funksjonen som ga studentene muligheten til å stemme på relevansen av et innlegg, kunne misbrukes ved å stemme flere ganger på samme innlegg. En annen negativ ting var at et testinnlegg kom veldig høyt på listen under konseptet «Turing Test». Ellers kommenterte to respondenter at utseendet, blant annet valg av farger, for brukergrensesnittet for relevante innlegg burde ha lignet mer på FLE3 sitt grensesnitt. Det ble også nevnt at det hadde vært nyttig å se hele konteksten/tråden som et innlegg hørte inn under, i tillegg til å kunne lenke de enkelte innleggene opp mot andre eksterne ressurser som referanser.

5. Konklusjon

Fremtidig arbeid bør derfor først og fremst dreie seg om de ovennevnte momentene, selv om det bør nevnes at funksjonen for å kunne se hele tråden som et innlegg hører til under, var planlagt, men det ble ikke implementert på grunn av tidsmangel. Det samme gjelder et brukergrensesnitt som lar studenten bla gjennom alle innlegg tilknyttet et spesifikt konsept. Mer arbeid bør også legges i å gjøre det lettere å importere og klassifisere hele diskusjonsforumer. For eksempel kunne man ha hatt et menyvalg for dette i lærerens annoteringsverktøy. I tillegg bør AnnForum bli mer generisk, slik at det blir lettere å integrere det i andre typer diskusjonsforumer. Ellers kan systemet nok refaktoreres en god del for å gjøre koden og databasen mer oversiktlig.

Neste steg vil bli å se på muligheten for å tilpasse systemet til andre domener. Selv om AnnForum først og fremst ble utviklet med tanke på diskusjonsforumer, så bør det ikke være umulig å bruke de samme prinsippene i lignende situasjoner der klassifisering av dokumenter og annen ustrukturert informasjon kan være hensiktsmessig. Dette kan for eksempel være for å kategorisere nyhetsartikler eller organiseringen av ressurser i et web-basert Content Management System. Noen av modulene fra AnnForum kan nok gjenbrukes i denne sammenhengen. For eksempel kan nøkkelordgjenkjenneren benyttes til å klassifisere innholdet, mens brukergrensesnittet for relevante innlegg kan tilpasses til å hente frem relaterte ressurser uavhengig av klassifiseringsalgoritmen. Fremgangsmåten ved bruk i andre systemer vil generelt være den samme som for AnnForum:

1. Opprett en konseptuell domenemodell for det aktuelle domenet.
2. Velg hvilke eksisterende ressurser som skal annoteres.
3. Identifiser og legg til rette for dynamisk annotering av nye ressurser.
4. Lag en funksjon som henter frem relevante ressurser.

6. Litteraturoversikt

- Antoniou, G. and Harmelen, F. A. (2004). *A Semantic Web Primer*. London: The MIT Press.
- Azouaou F., Chen, W. og Desmoulins, C. (2005). "Semantic Annotation Tools for Learning Material". *CaiSE'05*. 17th Conference on Advanced Information Systems Engineering, Portugal.
- Baeza-Yates, R., Ribeiro-Neto, B. (1999). *Modern Information Retrieval*. Essex, England: Paerson Education Limited.
- Berners-Lee, T. (1998) "Semantic Web Road Map".
[URL] <http://www.w3.org/DesignIssues/Semantic.htm> (14 oktober 1998).
- Berners-Lee, T. (2000). *Weaving the Web*. New York: HarperCollins Publishers inc.
- Berners-Lee, T., Hendler, J., Lassila, O. (2001). "The Semantic Web". *Scientific American*.
[URL] <http://www.sciam.com/article.cfm?id=the-semantic-web> (mai 2001).
- Berners-Lee, T. (2007). "Interview with Tim Berners-Lee – The Semantic Web of data". *Technology Review*.
[URL] <http://www.technologyreview.com/Infotech/18451/> (27 mars 2007).
- Chen, W. (2004). "Reuse of Collaborative Knowledge in Discussion Forums", Department of Information Science and Media Studies, University of Bergen.
- Craven, M., et al. (1998). "Learning to Extract Symbolic Knowledge from the World Wide Web", Proceedings of AAAI-98, 15th Conference of the American Association for Artificial Intelligence.
- Dillenbourg, P. (1999). "What do you mean by 'collaborative learning'?"
P. Dillenbourg (Ed): *Collaborative Learning: Cognitive and Computational Approaches*: 1-9. Oxford: Elsevier.
- Dix, A., Finlay, J., Gregory, D., Russel, B. (2004). *Human-Computer Interaction, third edition*. Essex, England: Pearson Education Limited.
- Euzenat, I. (2002). "Eight Questions about Semantic Web Annotations". *IEEE Intelligent Systems*. IEEE Educational Activities Department. 55–62.
- Garshol, M.L. (2002). "What are Topic Maps". XML.com.
[URL] <http://www.xml.com/pub/a/2002/09/11/topicmaps.html> (11 september 2002).

6. Litteraturoversikt

- Gourlay, S. N. (2002). "Tacit Knowledge, tacit knowing or behaving?". 3rd European Organizational Knowledge, Learning, and Capabilities conference, Athens, Greece, 5.–6 april.
- Gruber, T.R. (1993). "A translation approach to portable ontology specifications". *Current issues in knowledge modeling*, volume 5 , issue 2. 199–220.
- Gruber, T.R. (2007). "Ontology of Folksonomy: A Mash-up of Apples and Oranges". *Int'l Journal on Semantic Web & Information Systems*. 3(2).
- Guarino, N. (1998). "Formal Ontology and Information Systems". Proceedings of the 1st International Conference on Formal Ontologies in Information Systems, FOIS'98. Trento, Italy: IOS Press. 3–15.
- Grønmo, S. (2004). *Samfunnsvitenskapelige Metoder, første utgave*. Bergen: Fagbokforlaget Vigmostad & Bjørke AS.
- Handschuh, S., *Annotation for the Semantic Web*, volume 96 of Frontiers in Artificial Intelligence and Applications, 2003.
- Helic, D., Maurer, H. and Scerbakov, N. (2003). "Reusing discussion forums as learning resources in WBT systems". Proceedings of the IASTED Int. Conf. on Computers and Advanced Technology in Education. Rhodes, Greece. 223-228.
- Helic, D., Maurer H. and Scerbakov, N. (2004). "Discussion Forums as Learning Resources in Web-based Education". *Advanced Technology for Learning*. Vol., 1, No. 1. Anaheim: Acta Press. 8–15.
- Koschmann, T. mfl. (1996). "Computer Supported Problem-Based Learning: A Principled Approach to the Use of Computers in Collaborative Collaborative Learning". *CSCL: Theory and Praticice of an Emerging Paradigm*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Lipponen, L. (2002). "Exploring foundations for computer supported collaborative learning". G. Stahl, (Ed), 4th CSCL: Foundations for a CSCL Community, (CSCL), Colorado, LEA, NJ. USA. 72—81.
- Maier, R. (2007). *Knowledge Management Systems - Information and Communication Technologies for Knowledge Management* (3rd ed). Berlin: Springer.
- Nilsson, N. J. (1998). *Artificial Intelligence: A New Synthesis*. San Fransisco, California: Morgan Kaufmann Publishers Inc.
- Shadbolt, N., Hall, W. and Berners-Lee, T. (2005). "The Semantic Web Revisited". *IEEE Intelligent Systems* 21. 90-101.

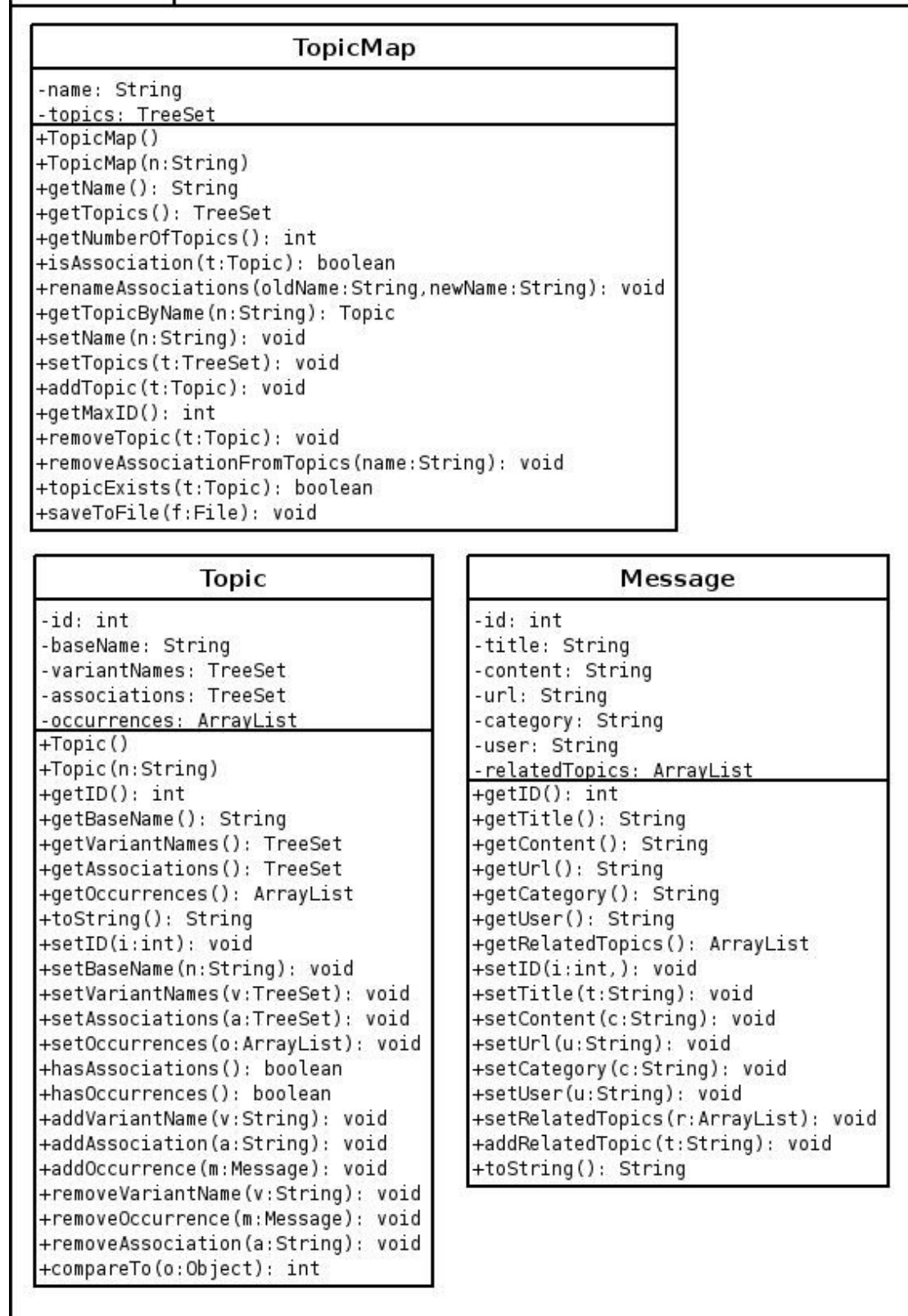
6. Litteraturoversikt

- Smoulders, W. M. mfl. (2000). "Content-Based Image Retrieval at the End of the Early Years". *IEEE Transactions on Pattern Analysis and Machine Intelligence*. 1349–1380.
- Stahl, G., Koschmann, T., Suthers, D. (2006). "Computer Supported Collaborative Learning: An historical perspective". R. K. Sawyer (Ed.). *Cambridge handbook of the learning sciences*. Cambridge, UK: Cambridge University Press. 409-426.
- Stanford Encyclopedia of Philosophy, Model Theory (2005).
[URL] <http://plato.stanford.edu/entries/model-theory/> (12 july 2005).
- Usability First (1998). "Groupware: Introduction".
[URL] <http://www.usabilityfirst.com/groupware/intro.txt> (2005).
- Vander Wahl, T. (2007). "Folksonomy".
[URL] <http://vanderwal.net/folksonomy.html> (2 februar 2007).

Appendiks

Appendiks A: Klassediagram

AnnForum: model



AnnForum:mysql

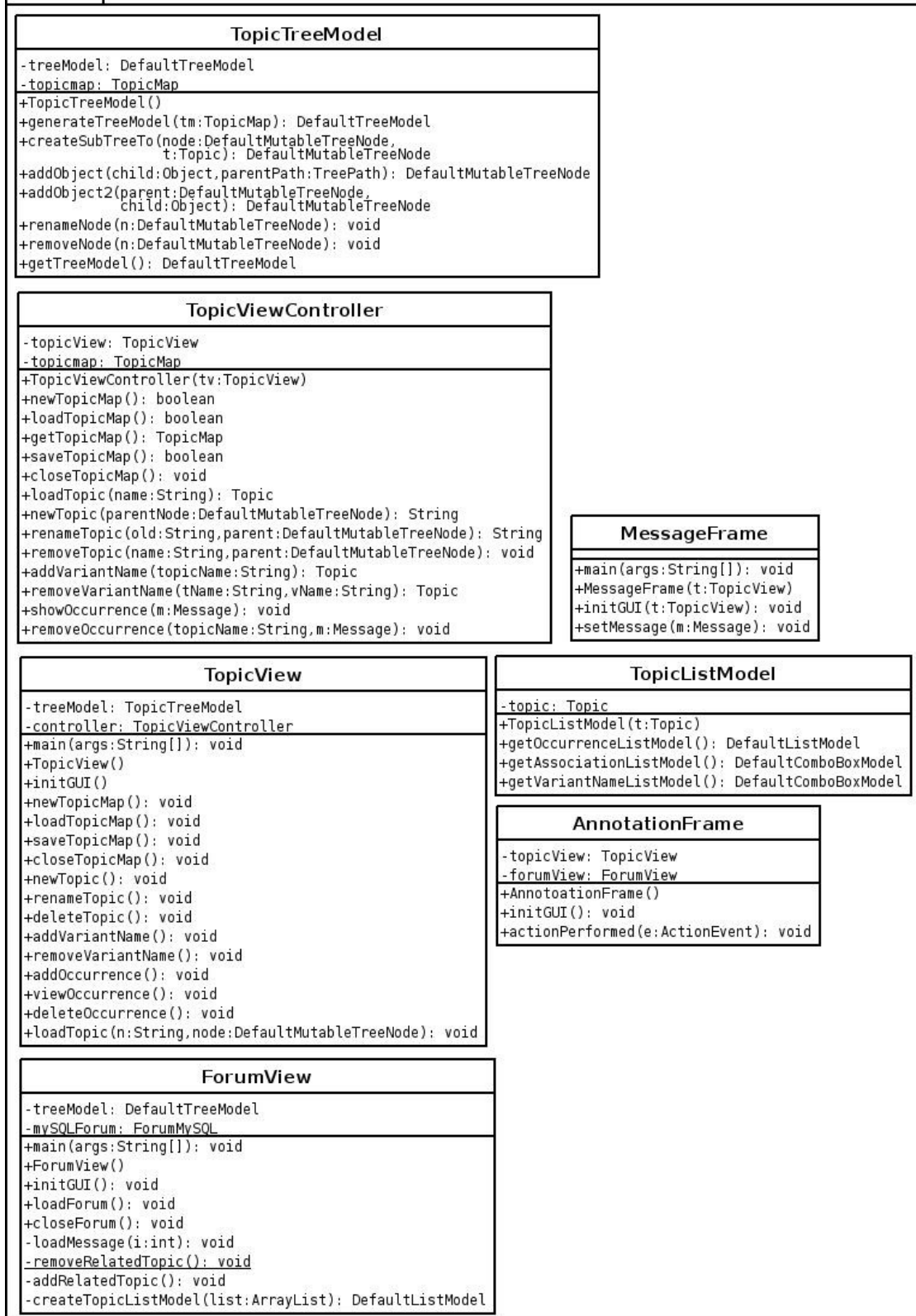
ForumMySQL

```
-dbClass: String
-database: String
-url: String
-user: String
-password: String
-messageTable: String
+ForumMySQL(dbTable:String)
+insertMessage(id:int,reply_to_id:int): void
+convertString(source:String): String
+getMessage(id:int): Message
+updateRelatedConcepts(id:int,a:ArrayList): void
+getTreeModel(): DefaultTreeModel
+createSubNodeTo(n:DefaultMutableTreeNode): DefaultMutableTreeNode
+getReplySet(id:int): ResultSet
+getAllMessages(): ArrayList
+topicArrayListToString(a:ArrayList): String
+stringToTopicArrayList(s:String): ArrayList
+getLowestReplyToID(): int
```

AnnForum:xml

TopicMapXML

```
-topicmap: TopicMap
-topicArrayList: ArrayList
+TopicMapXML()
+TopicMapXML(tm:TopicMap)
+loadTopicMapFromFile(f:File): TopicMap
+saveTopicMapToFile(f:File): void
-createTopicElement(d:Document,t:Topic)
-createTopicOccurrenceElement(d:Document,
m:Message): Element
-printToFile(file:File,d:Document): void
-parseTopic(topicElement:Element): Topic
-parseTopicVariantNames(a:NodeList): TreeSet
-parseTopicAssociations(a:NodeList): TreeSet
-parseTopicOccurences(o:NodeList): ArrayList
```



KeywordRecognizer:recognizer

Start

```
-stripQuotes(s:String): String  
+readStdIn(): String  
+main(args:String[]): void
```

Classifier

```
-vectorStorage: TermVectorStorage  
-vectorClassifier: VectorClassifier  
+Classifier()  
+getValueForTitle(title:String,t:Topic): double  
+getValueForContent(content:String,t:Topic): double
```

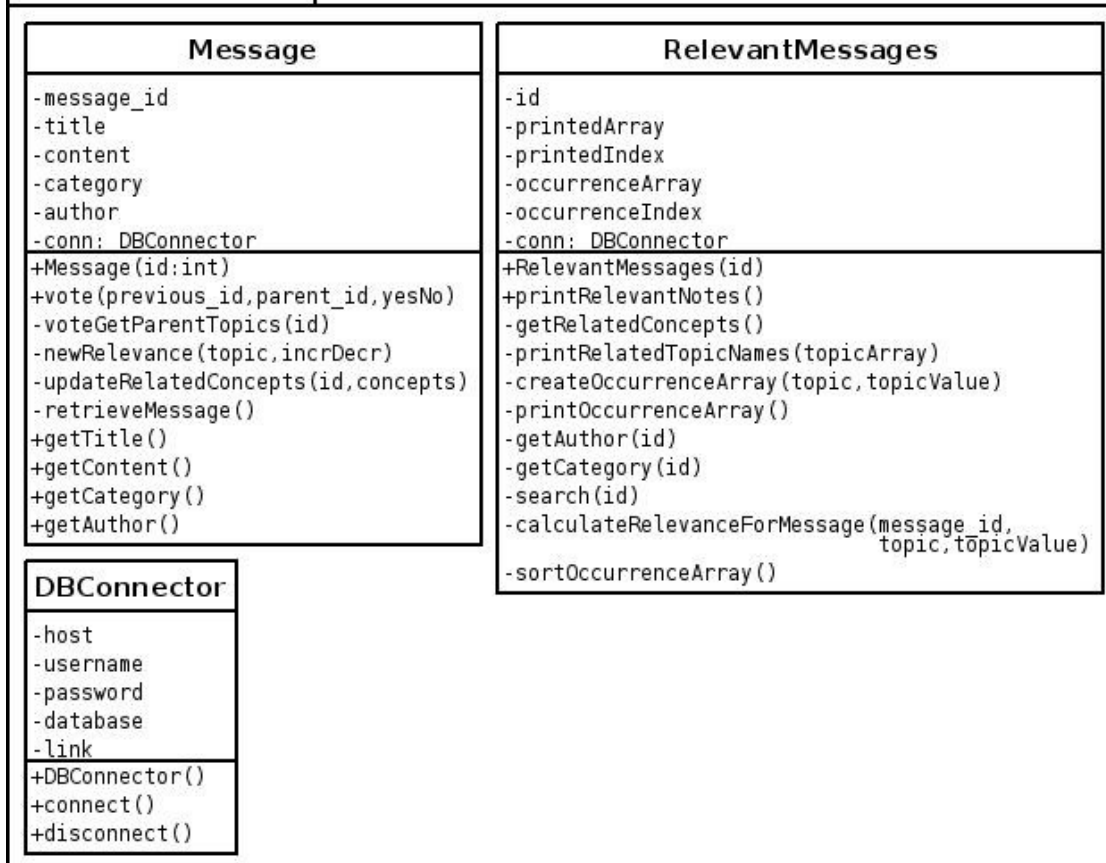
ClassifyDB

```
-topicMapPath: String  
-database: String  
-dbClass: String  
-url: String  
-user: String  
-password: String  
-messageTable: String  
-topicmap: TopicMap  
-classifier: Classifier  
+ClassifyDB()  
+startClassification(): void  
-classifyMessage(): void  
+decideConcept(): void  
-addOccurrence(): void  
-loadTopicMap(): void  
+getTopicNames(): ArrayList
```

ClassifyMessage

```
-topicMapPath: String  
-database: String  
-dbClass: String  
-url: String  
-user: String  
-password: String  
-messageTable: String  
-topicmap: TopicMap  
-topicNames: ArrayList  
-message: Message  
-classifier: Classifier  
+ClassifyMessage(m:Message,replyID:int)  
+setMessage(m:Message): void  
+classify(): void  
+decideConcept(topic:Topic): void  
-insertMessage(): void  
-loadTopicMap(): void
```

relevantMessageInterface



Appendiks B: Domenemodellen

```
<?xml version="1.0" encoding="utf-8"?>
<topicMap name="Artificial Intelligence">
  <topic id="1">
    <baseName>Agents</baseName>
    <association>
      <member name="Physical Agents"/>
      <member name="Temporal Agents"/>
    </association>
  </topic>
  <topic id="2">
    <baseName>Algorithm types</baseName>
  </topic>
  <topic id="3">
    <baseName>Applications</baseName>
  </topic>
  <topic id="4">
    <baseName>Bayesian Inference</baseName>
  </topic>
  <topic id="5">
    <baseName>Cognitive Science</baseName>
    <association>
      <member name="Computational modeling"/>
    </association>
  </topic>
  <topic id="6">
    <baseName>Computational modeling</baseName>
    <association>
      <member name="Data Preparation"/>
    </association>
  </topic>
  <topic id="7">
    <baseName>Data Preparation</baseName>
  </topic>
</topicMap>
```

```
<topic id="8">
  <baseName>Education</baseName>
</topic>
<topic id="9">
  <baseName>Ethical & Social implications</baseName>
</topic>
<topic id="10">
  <baseName>Expert Systems</baseName>
</topic>
<topic id="11">
  <baseName>Games & Puzzles</baseName>
</topic>
<topic id="12">
  <baseName>History</baseName>
</topic>
<topic id="13">
  <baseName>Human interaction</baseName>
  <association>
    <member name="Bayesian Inference"/>
  </association>
</topic>
<topic id="14">
  <baseName>Interfaces</baseName>
</topic>
<topic id="15">
  <baseName>Machine Learning</baseName>
  <association>
    <member name="Algorithm types"/>
    <member name="Human interaction"/>
  </association>
</topic>
<topic id="16">
  <baseName>Natural Language Processing</baseName>
</topic>
<topic id="17">
```


<baseName>Objections</baseName>
</topic>
<topic id="18">
 <baseName>Philosophy</baseName>
</topic>
<topic id="19">
 <baseName>Physical Agents</baseName>
</topic>
<topic id="20">
 <baseName>Reasoning</baseName>
</topic>
<topic id="21">
 <baseName>Reference Shelf </baseName>
</topic>
<topic id="22">
 <baseName>Representation</baseName>
 <variantName>KR</variantName>
</topic>
<topic id="23">
 <baseName>Robots</baseName>
</topic>
<topic id="24">
 <baseName>Science Fiction</baseName>
</topic>
<topic id="25">
 <baseName>Speech</baseName>
</topic>
<topic id="26">
 <baseName>Systems & Languages</baseName>
</topic>
<topic id="27">
 <baseName>Temporal Agents</baseName>
</topic>
<topic id="28">
 <baseName>Turing Test</baseName>

```
<association>
  <member name="Objections"/>
</association>
</topic>
<topic id="29">
  <baseName>Vision</baseName>
</topic>
</topicMap>
```

Appendiks C: Brukeroppgaver

1. Åpne Internet Explorer
2. Gå til adressen <http://129.177.34.67:8080/FLE>.
3. Logg deg inn (brukernavn: x, passord: x), du vil nå havne på siden «WebTop».
4. Fra toppmenyen, velg «Knowledge Building», trykk deretter på lenken «Artificial Intelligence discussion» for å komme til diskusjonsforumet for kunstig intelligens.
5. Her ligger det to hovedinnlegg, «Problem with the Turing test» og «Machine learning and intelligent robots», som hver har ett eller flere svar fra andre brukere.
6. Velg ett av hovedinnleggene eller ett av svarene, og se for deg at du ønsker å finne ut mer om dette temaet. Om du er usikker på hvilket innlegg du skal velge, bare les et av hovedinnleggene.
7. Bruk «Show related notes»-funksjonen for å se etter relevante ressurser for dette innlegget.
8. Svar på innlegget du valgte i 6 ved å gi din egen forklaring/definisjon (knowledge type: «My Explanation») på dette innlegget. Kom gjerne inn på andre AI-emner om du ser dette som nødvendig (machine learning, reasoning etc).
9. Bruk «Show related notes»-funksjonen for å se om du finner noen relevante ressurser for innlegget du nettopp la til.
10. Gå tilbake til «Artificial Intelligence discussion».
11. Velg et nytt innlegg på samme måte som i punkt 6, se for deg at du ønsker å finne ut mer om dette temaet. Igjen, om du er usikker på hvilket innlegg du skal velge, bare les et av hovedinnleggene.
12. Bruk «Show related notes»-funksjonen for å se etter relevante ressurser for dette innlegget.
13. Svar på innlegget du valgte i punkt 11, men denne gangen med et nytt spørsmål som videre utdyper problemet (knowledge type: «Problem»). Dette kan for eksempel være noe du selv lurer på angående temaet, eventuelt relatert til et annet tema innenfor kunstig intelligens (machine learning, reasoning etc).
14. Bruk «Show related notes»-funksjonen for å se om du finner noen relevante ressurser for innlegget du nettopp la til.
15. Om du ikke allerede har gjort det, bruk «Show related notes»-funksjonen til å

bla gjennom noen innlegg. Avgi stemme om du synes et innlegg er relevant eller ikke.

16. Fra toppmenyen, velg «WebTop».

17. Logg ut.

Appendiks D: Intervjutranskriberinger

Respondent 1

Kunnskap om AI-domenet (1–5): 3

Brukt FLE3 før (ja/nei)? Nei

1. Skjønte du hensikten med «Show related notes»-knappen?

Ja, det var fordelaktig å kunne se hva andre har sagt om temaet. I tillegg var det òg nyttig med en prosentverdi som sa noe om relevansen for de forskjellige innleggene.

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

Ja, innleggene som ble listet opp, virket relevante, dette på grunn av tittelen på de forskjellige innleggene som i stor grad handlet om Turing Test i en eller annen form. Relevansverdien i prosent ga òg et inntrykk av at disse ressursene var relevante desto høyere prosentverdien var.

Negativt: Relevansverdien kan òg misbrukes, for eksempel om en annen bruker stemmer sitt eget synspunkt høyt opp på listen.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

Ja, til dels. Ikke alle innleggene omhandlet Turing Test direkte, men det var mye relevant informasjon om temaet. Noen av innleggene referer til andre innlegg, noe som er vanlig i et diskusjonsforum.

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

Veldig nyttig, bra å kunne se andres synspunkter om temaet samt bruke innleggene som referanser.

5. Hva var mest interessant ved de foreslåtte innleggene?

Muligheten for å få mer utdypende kunnskap om temaet.

6. Hva var minst interessant ved de foreslåtte innleggene?

Allerede svart på (relevansverdien kan misbrukes om man ønsker å få sine egne meninger høyere opp på listen).

7. Var det noe du helst ville ha sett gjort på en annen måte?

Det kunne kanskje vært nyttig å knytte de enkelte innlegg opp mot referanser for AI-feltet, for eksempel vitenskaplige artikler og lignende.

8. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

Leste innlegg ut fra relevansverdi og tittel. Innlegg som var bra skrevet, ble rangert positivt. Unngikk bevisst å rangere innlegg som ikke var relevante, dette for ikke å rangere et innlegg negativt basert på subjektive meninger.

9. Synes du relevansverdien i prosent var riktig for de innleggene du leste?

Hvorfor?

Ja, for så vidt. Virket som om at innlegg med høy relevansverdi faktisk var relevante, men subjektive meninger kan òg spille inn på relevansen for innlegg.

Et velskrevet innlegg har man gjerne mye fakta, og man knytter det opp mot sine meninger. For mye subjektivitet i et innlegg kan lett oppfattes som mindre relevant enn et faktabasert innlegg. Dermed er det ikke sagt at de subjektive innleggene er irrelevante.

10. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Det virker som en god idé (med utgangspunkt i svarene som ble gitt ovenfor).

11. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Ja, men for det meste i situasjoner der brukerne er høyt motiverte, for eksempel i en studiesituasjon. I offentlige og/eller overfladiske diskusjoner vil brukerne kanskje ikke være motiverte nok til å bruke en slik funksjon.

12. Ville du brukt en slik funksjon i en studiesituasjon?

Ja, tror det.

13. Andre ting som var uklare eller som du ønsker å få frem?

Grensesnittet var oversiktlig.

Respondent 2

Kunnskap om AI-domenet (1–5): 5

Brukt FLE3 før (ja/nei)? Nei

1. Skjønnte du hensikten med «Show related notes»-knappen?

Ja. Den forsøker å tolke innholdet i innlegg og vise andre innlegg som er "like".

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 4

Innleggenes overskrifter viste raskt hvorfor programmet mente de var relevante.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 4

Som oftest var innleggene meget relevante og ga et inntrykk av gjennomtenkt diskusjon uten å vektlegge noen "sider" av diskusjonen mer enn andre. Noen av innleggene var irrelevante (Turing test ble lenket til et testinnlegg med kun ordet "test" i f.eks.)

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

På en skala fra 1–5 der 5 er veldig nyttig, hvor nyttig? 5

Meget nyttig. Helt klart en god idé.

5. Hadde de relevante innleggene du leste, noen påvirkning på hvordan du formulerte svarene dine?

Ja. De ga meg inntrykk av at det var et levende forum som folk virkelig brukte. Dette ga meg igjen inspirasjon til å skrive virkelige innlegg, ikke bare lire av meg en setning eller to.

6. Hva var mest interessant ved de foreslåtte innleggene?

Det mest interessante var at de virket reelt relaterte, de inneholdt ikke bare det samme ordet et par ganger som slike systemer vanligvis synes å resultere i.

7. Hva var minst interessant ved de foreslåtte innleggene?

Den gangen en test kom inn i lista mi.

8. Var det noe du helst ville ha sett gjort på en annen måte?

En enkel ingress under overskriftene i lista ville gjort det enda bedre!

9. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

Jeg leste alle med 100 % relevans, og stort sett alle med over 90 %. Rangerte de jeg syntes virkelig traff eller bommet på temaet.

10. Synes du relevansverdien i prosent var riktig for de innleggene du leste?

Hvorfor?

Relevansverdien er veldig viktig. Selv om jeg ikke kjenner metrikken bak, gir det meg et inntrykk av hvor mye jeg kan forvente av innlegget før jeg trykker på det.

11. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Ypperlig.

12. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Jeg tror den ville bli brukt mye i en studiesituasjon, og jeg tror også den ville vært veldig nyttig på "sosiale nettsider" generelt.

13. Ville du brukt en slik funksjon i en studiesituasjon?

Ja.

14. Andre ting som var uklare eller som du ønsker å få frem?

Alt greit.

Respondent 3

Kunnskap om AI-domenet (1–5): 2

Brukt FLE3 før (ja/nei)? Nei

1. Skjønte du hensikten med «Show related notes»-knappen?

Ja, den viser innlegg som handler om samme tema, og hvor relevante de var for det jeg leste.

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 4

Ja, jeg syntes det var mange relevante innlegg der. Innleggene handlet om det temaet jeg leste om.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 3

Samme som på spørsmål 2, innleggene handlet om samme tema, selv om kvaliteten på innleggene ikke nødvendigvis var like relevant.

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

På en skala fra 1–5 der 5 er veldig nyttig, hvor nyttig? 4

Nyttig.

5. Hadde de relevante innleggene du leste, noen påvirkning på hvordan du formulerte svarene dine?

Det er mulig jeg ble litt påvirket av det, ja.

6. Hva var mest interessant ved de foreslåtte innleggene?

Det var interessant å se hva andre hadde skrevet, og man kunne lese og lære mer om temaet enn det som var oppe i diskusjonen jeg var i.

7. Hva var minst interessant ved de foreslåtte innleggene?

De minst relevante innleggene var uinteressante.

8. Var det noe du helst ville ha sett gjort på en annen måte?

Nei.

9. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

De med høyest ranking leste jeg først, men jeg leste også noen med dårligere ranking for å se om jeg syntes de var dårligere. De som hadde høy relevans til temaet, rangerte jeg.

10. Synes du relevansverdien i prosent var riktig for de innleggene du leste?

Hvorfor?

Ja, jeg syntes innleggene med høyest relevansverdi hadde mest tilknytning til temaet, og jeg syntes innholdet i de innleggene hadde en faglig høy kvalitet.

11. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Jeg synes det er bra. Det er dumt at det samme skal skrives om igjen flere ganger, og man kan også finne lærdom i gamle innlegg.

12. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Ja, det tror jeg.

13. Ville du brukt en slik funksjon i en studiesituasjon?

Ja, jeg tror jeg ville benyttet meg av det.

14. Andre ting som var uklare eller som du ønsker å få frem?

Jeg er litt imot at samme ting skal skrives om igjen flere ganger ulike steder. Da er det litt bortkastet at man finner et poeng i en gammel diskusjon som man bruker i den nye. Samtidig ser jeg viktigheten av at de viktige poengene kommer klart frem i hver diskusjon også.

Respondent 4

Kunnskap om AI-domenet (1–5): 3, Har tatt et par fag, og noenlunde hobbyinteressert.

Brukt FLE3 før (ja/nei)? Nei.

1. Skjønte du hensikten med «Show related notes»-knappen?

Ja, den viser innlegg som er relevante for nåværende innlegg.

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 5

Ja, alle handlet om det aktuelle temaet, i dette tilfellet Turing Test.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 5

Ja (se forrige spørsmål).

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

På en skala fra 1–5 der 5 er veldig nyttig, hvor nyttig? 5

Nyttig! Det er lettere med en liste enn å måtte navigere et hierarki, og sorteringen gjorde det lettere å unngå unyttige innlegg.

5. Hadde de relevante innleggene du leste, noen påvirkning på hvordan du formulerte svarene dine?

Ja, leste en god del av de relevante innleggene, og da særlig før jeg skrev det andre innlegget (som da var basert på disse).

6. Hva var mest interessant ved de foreslåtte innleggene?

Synes det var interessant med muligheten for å avgi stemme for om et foreslått innlegg var relevant eller ikke.

7. Hva var minst interessant ved de foreslåtte innleggene?

Likte ikke popup-vinduene, synes heller at innleggene kunne kommet opp i samme vindu som listen med relevante innlegg, for eksempel til høyre for listen.

8. Var det noe du helst ville ha sett gjort på en annen måte?

Se over, samt at UI-en gjerne kunne hatt litt mer samme look-and-feel som FLE3.

9. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

Leste en god del av innleggene som handlet om Turing Test. Leste hovedsaklig først de innleggene med høy relevansverdi, men så òg litt på de på bunnen for å kunne sammenlikne. Var vel en kombinasjon av høy relevansverdi og interessant tittel som avgjorde hvilke relevante innlegg som ble lest. For det andre innlegget jeg skrev, fikk jeg ikke svar på spørsmålet mitt ved hjelp av de relevante innleggene.

10. Synes du relevansverdien i prosent var riktig for de innleggene du leste? Hvorfor?

Ja, fikk inntrykk av at de innleggene med høy relevansverdi faktisk var mer relevante for diskusjonen, mens de med lav relevansverdi ofte ikke var så godt skrevne / veldig relevante.

11. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Jeg er for.

12. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Ja, fordi jeg selv ville ha brukt den, og jeg er lat. Funksjonen er lettere tilgjengelig og mer nyttig enn den tradisjonelle SEARCH-knappen.

13. Ville du brukt en slik funksjon i en studiesituasjon?

Ja (se 12).

14. Andre ting som var uklare eller som du ønsker å få frem?

Nei.

Respondent 5

Kunnskap om AI-domenet (1–5): 4

Brukt FLE3 før (ja/nei)? Nei

1. Skjønte du hensikten med «Show related notes»-knappen?

Ja.

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 3

De fleste var relevant, men samtidig så virket det som at det var lite brukerdata for å beregne relevans. Etter at jeg hadde sagt at et innlegg var relevant, gikk det fra bunnen av listen og ganske langt opp.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 4

Ja, det var interessant å lese. Med mange innlegg kunne dette ha vært svært interessant.

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

På en skala fra 1–5 der 5 er veldig nyttig, hvor nyttig? 4

Det er svært nyttig i et diskusjonsforum. Ofte finnes relevante innlegg, men på grunn av en enorm informasjonsmasse blir det nesten umulig å finne frem til det man ønsker.

5. Hadde de relevante innleggene du leste, noen påvirkning på hvordan du formulerte svarene dine?

Nei.

6. Hva var mest interessant ved de foreslåtte innleggene?

Å lese andres vinklinger på samme sak.

7. Hva var minst interessant ved de foreslåtte innleggene?

8. Var det noe du helst ville ha sett gjort på en annen måte?

Det hadde vært fint om man hadde kunne sett mer av innleggene i listen som de kom opp i. Det er ofte slik at overskriften ikke reflekterer innholdet på en god måte.

9. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

Overskriften og rangeringen påvirket hvilke jeg leste.

10. Synes du relevansverdien i prosent var riktig for de innleggene du leste? Hvorfor?

Ja, men når få har stemt, blir prosentverdien litt upålitelig.

11. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Kjempebra. Mye kunnskap går tapt mellom hvert kull.

12. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Så absolutt.

13. Ville du brukt en slik funksjon i en studiesituasjon?

Ja.

14. Andre ting som var uklare eller som du ønsker å få frem?

Nei.

Respondent 6

Kunnskap om AI-domenet (1–5): 2

Brukt FLE3 før (ja/nei)? Nei

1. Skjønte du hensikten med «Show related notes»-knappen?

Ja, den søker etter tidligere innlegg om et tema basert på nøkkelordene.

2. Ved bruk av «Show related notes»-knappen, synes du innleggene som kom opp i listen, virket relevante for diskusjonen? Hvorfor?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 3

Ja, de virket relevante, men fant ikke informasjon for sub-tråder, titlene på de relevante innleggene var òg dårlige.

3. For de innleggene du faktisk var inne på, synes du de var relevante for diskusjonen? På hvilken måte?

På en skala fra 1–5 der 5 er mest relevant, hvor relevante? 4

Ja, innholdet i innleggene var stort sett bedre enn tittelen.

4. Hva synes du om muligheten for å se relevante innlegg? Nyttig eller unyttig?

På en skala fra 1–5 der 5 er veldig nyttig, hvor nyttig? 5

Veldig nyttig, spesielt for generelle spørsmål da man unngår å få duplikeringer av samme type spørsmål/svar. De relevante innleggene gjør deg òg oppmerksom på lignende tema.

5. Hadde de relevante innleggene du leste, noen påvirkning på hvordan du formulerte svarene dine?

Nei, av gammel vane så jeg på de relevante innleggene fra tidligere diskusjoner etter jeg formulerte svaret i nåværende diskusjon.

6. Hva var mest interessant ved de foreslåtte innleggene?

Et godt alternativ til å søke da det på en måte fungerte som «automatisk søk». Viste òg mer info om et tema. Ikke detaljerte nok med søkefunksjoner (kan du klarifisere denne litt?) Skulle likt å sett hvilke diskusjoner innleggene var hentet fra, innlegget de svarte på, og innlegg som er svar på innlegget. Er også det jeg prøver å si under i #8.

7. Hva var minst interessant ved de foreslåtte innleggene?

Det er fare for misbruk ved at man kan stemme sine synspunkt høyt opp på listen, bør være en sperre for dette.

8. Var det noe du helst ville ha sett gjort på en annen måte?

Kanskje oppfordre bruker til å skrive mer detaljerte / bedre titler til innleggene. Bør òg være mulig å vise hele tråden for de enkelte relevante innleggene. GUI-en burde hatt samme stil som FLE3.

9. Hvilke innlegg leste/rangerte du? Hva var det som gjorde at du valgte akkurat disse innleggene?

Leste innlegg basert på tittel og relevansverdi, tittel var det som var viktigst. Innleggene var relevante, og det var derfor jeg rangerte de, stemte ett innlegg opp på listen og ett innlegg ned.

10. Synes du relevansverdien i prosent var riktig for de innleggene du leste? Hvorfor?

Ja, som oftest var relevansverdien bra nok.

11. Hva synes du om ideen om å gjenbruke innlegg fra tidligere semesters diskusjoner?

Bra, det er synd at viktig informasjon går tapt fra semester til semester. Eventuelt at en diskusjonstråd med nyttig informasjon blir ”begravd” fordi den har vært innaktiv, mens nyere tråder er startet og dermed kommet først på listen av alle tråder.

12. Tror du at en slik funksjon ville blitt brukt i en virkelig studiesituasjon?

Ja, det tror jeg. Diskusjonsforumer kan bli mangelfulle fordi det nødvendigvis ikke er så mange tråder som startes i løpet av et semester. Flyktig info blir ofte ikke brukt, og da mener jeg tråder som diskuter ting utenom fagfeltet (for eksempel bytte av klasserom / fagfest en spesifikk dato) ikke vil komme opp i ”relevante innlegg”, altså disse kan bli begravd.

13. Ville du brukt en slik funksjon i en studiesituasjon?

Ja.

14. Andre ting som var uklare eller som du ønsker å få frem?

Nei.