

Computing in Science Education (CSE) på IBV: status og veien videre

Notat skrevet av Lex Nederbragt, koordinator implementering av CSE på IBV, februar 2020. Med stor takk til de involverte undervisere for alle deres bidrag i skriveprosessen.

Høsten 2017, med innføring av det nye studieprogrammet Biovitenskap, har implementering av Computing in Science Education (CSE) begynt for alvor på IBV. Målet er å inkorporere beregningsorientert biologi som en del av bachelorløpet for alle biovitenskapsprogramstudenter. Arbeidet skjer i tett samarbeid med Centre for Computing in Science Education (CCSE)¹. Prosjektet er så vidt oss bekjent det første av sitt slag i Norden, kanskje til og med i hele verden.

Hensikten med dette notatet er å gi en status oppdatering på hvordan CSE så langt er implementert, og skissere noen mulige veier for videre implementering.

Oppsummering

IBV er i gang med implementering av beregningsorientert undervisning som en vesentlig komponent i alle kurs. Det gjenstår imidlertid mye arbeid før vi har oppnådd en helhetlig integrering av CSE på tvers av kursene. Til det trengs det både en videreføring av den kulturendringen som startet med det nye studieprogrammet 2017, mer ambisiøse (lærings)mål, koordinering av kurs, og kompetansebygging blant lærerkreftene ved instituttet. For å få nødvendig momentum er en større satsning nødvendig. Spesielt i spesialiseringsemnene fra og med 4. semester gjenstår mye innsats for å få CSE implementert. Generøs bistand er nødvendig for å senke terskelen emneansvarlige i dag har for å prioritere dette arbeidet. Emneansvarlig og undervisere må gis bistand i form av kompetanseheving og hjelp til utvikling og implementering av CSE i lab og kollokvieundervisningen. For å få CSE undervisningen i gang trengs PhD/postdoktorer med nødvendig kompetanse.

Mål

Jeg kan ikke finne at IBV har et uttalt **mål med CSE-implementeringen**. Jeg går ut fra at målet vil være å få integrert beregningsperspektivet i hele bachelorløpet slik at våre ferdigutdannede kandidater har de beregningsferdighetene som kreves av en moderne biolog, og dermed er godt rustet til (morgen)dagens arbeidsmarked. Våre kandidater vil med slike ferdigheter bli bedre rustet til å gå videre til masterprogrammet ved IBV eller andre steder der beregninger, analyser og programmering etterhvert er en viktig del av de fleste forskningsprosjekter. For å få dette til må **studentene ha et helhetlig tilbud hvor beregningsperspektivet er inkludert i de fleste kurs**. Vi må sørge for at vi gjennom våre kurs har en gradvis utvikling i beregningsferdighetene gjennom hele bachelorløpet.

¹ <https://www.mn.uio.no/ccse>

Status februar 2020

1.-3. semester

Det er blitt gjennomført en uformell spørreundersøkelse blant emneansvarlige og andre involverte for kursene i 1., 2., og 3. semester av studieprogrammet. Besvarelsene, som er sammenfattet i oversikten nederst i dette dokumentet, viser at det er CSE-implementering **på nesten alle kurs de første 3 semestrene**.

I **1. semester** er BIOS1100 innføringskurset i CSE. Her undervises grunnleggende programmering i Python og matematikk/modellering. Jupyter Notebooks med hjelp av UiO sin sky-løsning (JupyterHub) benyttes som plattform. CSE er en del av den summative vurderingen i kurset, både for obligatoriske innleveringen og slutteksamen. Høsten 2019 ble det mulig, for første gang for noen digital eksamen i Silurveien, for studentene å levere oppgaver som Jupyter Notebooks. Dette innebærer at studenten også kunne kjøre og teste koden sin før innlevering. Dette førte til at undervisningsformen og eksamensformen var mye mer samstemt enn før, noe som er viktig i et 'samstemt undervisning' ('Constructive Alignment') perspektiv. I samme semester tar studentene BIOS1110 og KJM1002 eller KJM1101. Noe av stoffet i førstnevnte kurs brukes i øvelser i BIOS1100. Kjemifagene benytter programmering i forbindelse med laboratoriearbeidet. På tre av fire kjemilaber må studentene gjøre enkle beregninger og noe databehandling for å besvare spørsmål i tilknytning til laboratorierapport.

I **2. semesteret** brukes Jupyter Notebooks i alle kurs, BIOS1120, BIOS1130 og FYS1001. Studenter får utlevert notebooks med ferdig utviklet kode som de i noen tilfeller må justere. Usikkerhetsberegning, nedlasting og fremstilling av data blir også videreutviklet. Programmering og beregninger går inn i innleveringer av notebooks eller laboratorierapporter i BIOS1120 og BIOS1130.

3. semester	BIOS1140 – Evolusjon og genetikk	BIOS1150 – Biologisk mangfold	STK1000 – Innføring i anvendt statistikk
2. semester	BIOS1120 – Fysiologi	BIOS1130 – Biokjemi I	FYS1001 – Innføring i fysikk
1. semester	BIOS1100 – Innføring i beregningsmodeller for biovitenskap	BIOS1110 – Celle- og molekylærbiologi og HMS-emner	KJM1002 – Innføring i kjemi / KJM1101 – Generell kjemi
	10 studiepoeng	10 studiepoeng	10 studiepoeng

Oppbygging 1., 2. og 3. semester Biovitenskapsprogrammet

I **3. semester** går studenter over til programmeringsspråket R. Dette er naturlig for STK1000, men også BIOS1140 bruker dette språket. Biologer trenger kunnskap om R siden det er mange viktige analysemetoder relevant for biologiske data som er tilgjengelig i dette

språket. BIOS1140 har mye modellering og beregning i labbene med en kombinasjon av at studenten utvikler egne og utdelte koder. Dette går også inn i de obligatoriske innleveringer. BIOS1150 har øvelser som bygger på populasjonsmodeller fra 1. semester som utdeles som ferdigskrevet R kode. STK1000 har 2 obligatoriske oppgaver der bruk av R er helt sentralt.

Spesialiseringsemner 4.-6. semester

CSE er en del av BIOS2100 - Generell økologi, BIOS3000 - Design og analyse av biologiske studier og BIOS3900 - Biokjemi II. CSE-delen består stort sett av ferdigutviklet kode som studenter skal jobbe med, eller en blanding av ferdig utviklet kode og kode som studenter skal skrive selv. BIOS3000 har innleveringer som krever CSE-kunnskap. Våren 2020 går BIOS3010 - Bioinformatikk, for første gang. Det ligger i planene å bygge ut programmeringsferdighetene i Python i dette kurset, og muligens blir det en del av den summative vurderingen for dette emnet også på sikt.

Matematikk

De tidligere studieprogrammene Biologi og Molekylærbiologi og biologisk kjemi hadde begge et 10 studiepoengs matematikk kurs i 1. semester, MAT1001. Dette kurset ble ikke videreført i Biovitenskap-programmet. Fra høsten 2018 ble det innført ett matematikkpensum som en del av undervisningen i BIOS1100, med Arne Sletsjøe fra Matematisk Institutt som foreleser. Stoffet, omtrent en tredjedel av MAT1001, ble i den grad det var mulig gjort relevant for resten av emnet. I 2. semester undervises studentene i bruk av likninger, og gjennomfører de kurvetilpasning til lineære og ikke-lineære data, mm. I 3. semester er det undervisning i dynamisk og stokastisk modellering med differenslikninger, og i matematikk som er relevant for statistikkfaget og for evolusjonsmodeller. Ytterligere undervisning i matematikk er så vidt jeg vet ikke planlagt for disse programstudentene.

Computational Science: Bioscience studieløp

Sammen med studieseksjonen ble det etablert et løp for Biovitenskapstudenter som ønsker å kvalifisere seg til den beregningsorienterte masterretningen CS:Bioscience². I dette løpet tar studentene MAT1100, MAT1110 og MAT1120 fra og med 3. semester. De første studenter starter på dette løpet høsten 2020.

Effekter på studentenes studievalg

Studieseksjonen melder at studentene er mer interessert i bioinformatikk (basert på studieveiledninger) og påmeldingen til BIOS3010 er stor (selv om dette delvis skyldes større kull enn før). Noen biovitenskapstudenter tenker å gå CS:Bioscience retningen og noen har byttet studieprogram til informatikk.

Videre implementering

Det er en stor styrke at alle 2. og 3. semesterkurs bruker programmeringsferdighetene studentene får med seg fra 1. semester. Det ville styrke studentenes CSE og

² <https://www.uio.no/studier/program/computational-science-master/studieretninger/bioscience>

beregningsferdigheter om alle 2. og 3. semesterkurs **bygger videre på programmeringsferdighetene** studentene får med seg fra 1. semester, **samt å teste studenter i det som en del av eksamen.**

Den nåværende situasjonen er at CSE-komponenten etter 1. semester er oppgaver som del av gruppeundervisningen eller laboratorieundervisningen der studenter skriver begrenset med egen kode, men får det meste utlevert. Dette kan vedlikeholde ferdighetene som studenter tilegner seg i første semester, men vil i begrenset grad videreutvikle dem.

For studentene ville det være en fordel at programmeringskompetanse i Python aktiviseres i 3. semester under innføringen i R for å gjøre overgangen enklere. Med dette menes at studentene får vist hvordan ting de har lært å gjøre i Python gjøres i R. Dette ville både gjøre overgangen lettere samt styrket Python-kunnskapene.

For at studieprogrammet skal oppleves som helhetlig er det **essensielt at CSE-komponenten integreres i hele utdanningsløpet.** Det er derfor viktig at flere spesialiseringsemner aktivt implementerer en CSE-komponent. Det bør lages et eller flere spesialiseringsemne som utdyper kunnskap om beregningsmodeller i biovitenskap.

I og med at matematikk er en vesentlig komponent av modellering er det **viktig at det blir mer matematikkundervisning, med en klar progresjon, i programmet,** spesielt i de tre første semestrene.

CSE-arbeidet på IBV oppleves å ha en bred støtte i ledelse og blant fast vitenskapelige ansatte. Det er forholdsvis lite koordinering mellom emnene, både på tvers av semestrene og innad i samme semester. Tiden er moden for å samle de involverte undervisere og utveksle erfaringer, noe jeg kommer til å ta initiativ til.

Utfordringer

Det å integrere beregninger i et kurs er et stort utviklingsarbeid. Å implementere CSE i et kurs krever undervisningstid, noe som er et spørsmål om prioritering. Det er noe begrenset CSE-kompetanse blant emneansvarlige, noe som gjør terskelen for å ta grep med CSE høy. CSE-utviklingen er for en stor del avhengig av PhD-studenter med pliktarbeidstid og riktig kompetanse, og 'CSE sommerstudenter', som er masterstudenter lønnet fra CCSE.

En ikke-uvesentlig, praktisk utfordring er ikke-programstudenter (bl. a. lektorstudenter) som ikke har den samme CSE-, og spesielt programmerings-bakgrunnen som programstudenter ved kursstart i 2., 3. semester, samt 4.-6. semester. Dette fører til at CSE-delen blir lagt opp som et ekstra tilbud som ikke er en del av den summative vurderingen i kurset. Dette er med på hindre at programmet har et helhetlig tilbud med beregningsperspektiv i alle kurs.

Det er et begrenset antall PhD-studenter og postdoktorer med pliktarbeid på instituttet, og relativt få av disse har kompetanse som er relevant i utvikling av CSE. Noen forskningsgrupper har, grunnet deres tverrfaglighet, PhD-studenter og/eller postdoktorer med relevant CSE-kompetanse, men disse er ofte ansatt ved andre institutter og har derfor

ofte sitt pliktarbeid der. Disse instituttene synes kanskje ikke det er rimelig at de avser pliktarbeidstid for arbeid med CSE ved IBV uten tilsvarende motytelser. Et alternativ er at IBV i større grad kunne vurdere å tilby pliktarbeid til slike kandidater som likevel har sitt arbeidssted (men ikke ansettelsesforhold) ved IBV.

Den begrensede CSE-kompetansen blant emneansvarlige fører til at en god del av dette arbeidet må overlates til PhD-studenter med pliktarbeid, samt masterstudenter som lønnes for dette (via CCSE). IBV har i de senere år hatt med en formulering om ønsket relevant kompetanse innen CSE i utlysningstekstene for fast vitenskapelig ansatte, noe som på sikt kan bidra til en kompetanseøkning. Dette forutsetter imidlertid at slik kompetanse finnes blant kandidatene og prioriteres ved tilsetting.

Veien videre

I IBVs strategi av 2020 står følgende om CSE-innsatsen: "Computing in Science Education" (CSE-prosjektet) ved MN-fakultetet har utviklet en prisbelønnet læringsmåte der beregninger tas naturlig inn i realfagsutdanningene tilpasset de ulike studentenes interesser. IBV skal dra nytte av CSE i sine bachelorprogram for topp utdanning av morgendagens arbeidstakere.³ Målet jeg beskrev under avsnittet 'Mål' mener jeg er en konkretisering av dette. Jeg ser behov for å **forankre en slik konkretisering av målet i instituttets strategi, samt gi CSE-koordinatoren mandat, rammevilkår og delegert myndighet for arbeidet.**

Til nå har integreringen av CSE på IBV blitt drevet av emneansvarlige som mener dette er et viktig tiltak. Jeg mener at det er på tide å **skru opp ambisjonsnivået.** Det bør, i fellesskap, lages **et sett med læringsutbyttebeskrivelser rettet mot beregningsferdighetene for hvert semester**, inkludert 4.-6. semester.

Matematikkunnskap er en vesentlig del av dette, og det bør tas grep med å **utvide matematikkundervisning i det i 2. og 3. semester. Testing av beregningsferdighetene bør bli en del av den summative vurderingen for alle relevante kurs.** Det at det nå er mulig å kjøre Python-kode på eksamen med hjelp av Jupyter Notebooks viser at det er praktisk mulig å teste slike ferdigheter på eksamen også i andre fag og andre programmeringsspråk. **Kunnskap i programmering og modellering bør være obligatorisk forkunnskap for kurs etter 1. semester**, noe som krever et lynkurs som sikrer relevant kompetanse hos studenter som ikke har hatt BIOS1100.

Det er et klart **behov for jevnlig erfaringsutveksling mellom undervisere i de involverte kursene.**

Ressurser

Et virkemiddel for å støtte opp under utvidelsen av CSE-implementering er å øke kompetansen blant eksisterende ansatte involvert i undervisning på biovitenskapsprogrammet. En mulig tilnærming er et samarbeid med KURT –

³ <https://www.mn.uio.no/ibv/om/strategi/strategi-2020>

Kompetansesenter for undervisning i realfag og teknologi⁴ på fakultetet. KURT organiserer ProFag, etterutdanning for lærere i realfaglig PROgrammering for FAGenes skyld⁵. Det har vært diskutert i KURT å utvide dette tilbudet til ansatte på de forskjellige instituttene på fakultetet. **IBV bør ta initiativ til et samarbeid med KURT for å realisere kompetanseheving for ansatte involvert i undervisningen**, og prioritere ressurser til et slikt tilbud.

BIOS1100 har i perioden 2015-2018 hatt en fantastisk ressurs i fire PhD-studenter fra CINPLA-miljøet, som kunne bruke en stor del av sitt pliktarbeid på å lage en lærebok for kurset⁶. Kun en av disse studentene hadde sitt pliktarbeid dekket av IBV. Disse kandidatene har nå disputert og er i jobb hos andre arbeidsgivere. Ved søknad og innvilgelse av CCSE forpliktet IBV seg til en egenandel i CCSE: "IBV har forpliktet seg til å bidra med 20% av undervisningstiden til 4 PhD-studenter/post-docs til utvikling av CSE-undervisning." Etter at CINPLA PhD-studentene avsluttet sin periode har instituttet prioritert å gi 25% pliktarbeid for en postdoktor til CSE-arbeidet. Det er ikke avklart **hvilke 3 andre PhD-studenter som skal ha 20% av sin undervisningstid knyttet til CSE-arbeid ved instituttet**. Dette er avgjørende for det videre arbeidet og må avklares raskt. Uten at dette gjøres vil det ikke være mulig å starte et mer langsiktig utviklingsarbeid for de kursene som ikke har integrert beregninger.

Det **bør utøves fleksibilitet når det gjelder bruk av pliktarbeidstid** fra PhD-studenter og postdoktorer med kompetanse til å bidra til CSE-implementering på IBV, og som jobber på forskningsprosjekter der IBV er vert, men som er formelt ansatt på andre institutter.

⁴ <https://www.mn.uio.no/kurt>

⁵ <https://www.mn.uio.no/kurt/livslang-lering/profag>

⁶ Læreboken er fortsatt under utvikling, koordinert av Lex Nederbragt

Appendiks A: detaljer for de relevante kursene

Denne oversikten er begrenset til de emnene som har en (vesentlig) komponent av Computing in Science Education - dvs det gjøres beregningsorientert undervisning med hjelp av et programmeringsspråk (Python eller R). Innholdet er basert på besvarelsen på den uformelle spørreundersøkelsen gjennomført i 2019.

1. semester

- BIOS1100 - Innføring i beregningsmodeller
 - Python i Jupyter notebooks via JupyterHub
 - alle obligatoriske innleveringer krever CSE-kunnskap
 - CSE del av stort sett alle eksamensoppgaver
 - forskning
 - H18: 2 MSc (lektor) som forsker på holdninger og arbeidsmåte
 - H19: 2 MSc (lektor) som fortsetter disse undersøkelser
 - bok
 - trengs en god slump TID og oppmerksomhet for ferdigstilling
 - har hatt hjelp av CCSE-sommerstudenter
- BIOS1110 - Celle- og Molekylærbiologi
 - mye potensiale i kobling mot BIOS1100, kan utnyttes bedre
 - H19 eksperimenteres det med Python-oppgaver (i Jupyter Notebook) i gruppeundervisning utviklet av en CSE-sommerstudent
 - Python i Jupyter notebooks via JupyterHub
 - har hatt hjelp av en CCSE-sommerstudent
- KJM1002 og KJM1101
 - Python i Jupyter notebooks via JupyterHub i KJM1101 (Generell kjemi), og KJM1002 (Innføring i kjemi) i forbindelse med labundervisningen
 - noe analyseresultater som inngår i labrapporten baseres på utregninger gjort med Python i notebook
 - studentene velger selv om de vil bruke ferdigskrevet kode eller skrive egen kode, til det kan de få orakelhjelp
 - underviser regresjonsanalyse

2. semester

- BIOS1120 - Fysiologi
 - Python i Jupyter notebooks via JupyterHub brukes i gruppetimene
 - målet er at studenter skal vedlikeholde og utvikle sine programmeringsferdigheter fra BIOS1100
 - underviser i å eksplorere fenomener og likninger ved bruk av programmering
 - studenter må skrive en del selvstendig kode, men får det meste utlevert
 - innlevering og godkjenning av notebooks er obligatorisk
 - programmering ikke en del av eksamen, men forståelse av tema som er arbeidet med i beregningsøvelsene blir gitt
- BIOS1130 - Biokjemi

- Python i Jupyter notebooks via JupyterHub
- brukes i gruppetimene og i forbindelse med laboppgaver
- studenter får stort sett utlevert kode, i noen tilfeller må koden tilpasses selvstendig
- studenter benytter kurvetilpasning til lineære og ikke-lineære data, least square fitting, og oppsett av egne lineær polynom og andre funksjoner
- ikke en del av eksamen
- har hatt hjelp av CCSE-sommerstudenter
- FYS1001
 - Python i Jupyter notebooks via JupyterHub brukes i gruppetimene
 - studenter må skrive litt selvstendig kode, men får det meste utlevert
 - bruker importering av data, plotting, kurvetilpasning, Python som kalkulator
 - ikke en del av eksamen

3. semester

- BIOS1140 - Evolusjon og genetikk
 - bruker mange øvelser i Rstudio/R i gruppetimene, installert på egen maskin
 - simulering av dynamisk og stokastisk modellering med differenslikninger, PCA, maksimum likelihood, vektorisering, matriser
 - fil- og databehandling og analyse av store datasett, statistisk analyse, simuleringer, visualisering og tolkning av biologiske data
 - kombinasjon av at studenten utvikler egne koder og utdelte koder
 - R og RStudio på studentenes egen maskin
 - del av obligatoriske datalabber med innlevering
 - ikke en del av eksamen, men eksamen inkluderer analyse og tolkning av modeller
- BIOS1150 - Biomangfold
 - bruker noen øvelser med hjelp av ferdigskrevet R i gruppetimene
 - meningen var at studenter skulle skrive litt selvstendig kode, men det fungerte dårlig
 - R via shiny-baserte simuleringer via en sky-basert løsning og BYOD
 - bygger på populasjonsmodeller (eksponensiell vekst) og matematikk (matriseregning) fra BIOS1100
 - ikke en del av eksamen, en enkel Canvas-innlevering av svar på spørsmål som stilles i øvelsene (fritekst - ikke programmering)
- STK1000 - Statistikk
 - bruker R, hovedsakelig via RStudio installert på studentenes egen maskin
 - ukesoppgaver i gruppetimene der det er nødvendig å bruke R i mange av oppgavene.
 - 2 obligatoriske oppgaver der bruk av R er helt sentralt
 - R kode blir utdelt men studentene lærer å bruke eksisterende R-funksjoner til å tilpasse statistiske modeller
 - på eksamen må studentene tolke utskrifter fra R

Spesialiseringsemner 4.-6. semester

- BIOS2100
 - grunnet tidsmangel og at studentene har svært ulik bakgrunn i programmering brukes ferdigstilte programmer i R (Shiny apps)
 - fokus på modelleringsferdigheter heller enn programmering
- BIOS3000
 - en stor del av kurset består av at studentene gjør dataøvelser (simuleringer og analyser) på egenhånd
 - krever forholdsvis lite programmeringsferdigheter, blanding av ferdig utviklet kode og skrive kode selv
 - R via Rstudio på egen maskin
 - innleveringer skrevet i RMarkdown (markup språk som blander tekst, kode og resultater og kan gjøres om til html, pdf, mm)
 - har hatt hjelp av CCSE-sommerstudenter
- BIOS3010 - Bioinformatikk
 - må utvikles fra scratch
 - BIOS1100 bør være obligatorisk forkunnskap
- BIOS3900 - Biokjemi II
 - Python i Jupyter notebooks via JupyterHub brukes i 1 av 6 labøvelser
 - studenter må ikke skrive selvstendig kode, men får det utlevert
 - har hatt hjelp av CCSE-sommerstudenter

Appendiks B: alle spørsmål og svar

Spørsmålene:

- 1) integreres CSE i undervisningen og på hvilken måte? (pensum, gruppeundervisning, innleveringer, mm)
- 2) hvilket programmeringsspråk brukes og på hvilken måte jobber studenter med det (f. eks. 'Python i Jupyter notebooks via JupyterHub', eller 'R via Rstudio installert på studentenes egen maskin')
- 3) hva lærer studentene av nye programmerings/modelleringsferdigheter?
- 4) utvikler studenter programmer/modeller selv eller får de dem utdelt?
- 5) har kurset innleveringer som krever programmering/modellering?
- 6) hvilken andel av eksamensoppgaver krever programmering/modellering?

Alle svarene er tilgjengelig her:

https://docs.google.com/spreadsheets/d/1MR52cVTijtiIAZDXXZO70_X2PN2m5KChOtml3uJ_wpg/edit?usp=sharing