



Sluttrapport læringsprosjekt generativ KI

Statens vegvesen 2024



Illustrasjonen er KI-generert. Copilot / Bjørn Terje Bakken

Innholdsfortegnelse

1	Innledning.....	3
2	Hva er generativ KI?.....	3
2.1	Bruksområder for generativ KI	3
2.2	Eksempler på tjenester	4
3	Prosjektgjennomføring.....	5
3.1	Hvordan ble læringsprosjektet gjennomført?	5
3.2	Hvem var deltakerne?	6
3.2.1	Deltagernes holdning til generativ KI før oppstart	6
3.2.2	Deltagernes erfaringer med generativ KI før oppstart	6
3.3	Hva var deltagernes motivasjon for å delta i prosjektet?	7
4	Teknologivalg: Copilot Free (Copilot i nettleser)	9
4.1	Om verktøyet Copilot Free	9
5	Resultater fra spørreundersøkelsen	11
5.1	Evaluerings av prosjektgjennomføringen	11
5.1.1	Informasjonsflyt	11
5.1.2	Samarbeid med andre	12
5.1.3	Nytte av å delta	12
5.1.4	Generelle tilbakemeldinger på prosjektgjennomføringen	13
5.2	Evaluerings av valgt teknologi.....	14
5.2.1	Hvor ofte brukte du Copilot til jobbrelevante oppgaver i prosjektperioden?	14
5.2.2	Hvilke arbeidsoppgaver så du for deg at KI kunne hjelpe deg med?	14
5.2.3	Hvilke jobbrelevante oppgaver brukte du Copilot til?.....	15
5.2.4	Hvordan har Copilot påvirket tidsbruken din på dine daglige arbeidsoppgaver?	16
5.2.5	Har du møtt begrensninger ved bruk av Copilot?	18
5.2.6	Har du møtt på faktafeil i svarene?.....	20
5.2.7	Har du møtt etiske problemstillinger i svarene du har fått?	22
5.2.8	Har du møtt personvernmessige utfordringer med bruken?.....	23
5.2.9	Kommer du til å fortsette å bruke Copilot etter deltagelsen din i prosjektet?	25
6	Erfaringer fra 8 prosjektdeltagere	26
6.1	Elin Dalen-Rasmussen (Drift og vedlikehold – Plan og utbygging vest).....	26
6.2	Gina Ytteborg (Utbygging - Bærekraft)	26
6.3	Tone Lise Dahl (Fellesfunksjoner og HR - Kompetanse).....	26
6.4	John Ole Qvale (IT - Brukernær)	27
6.5	Thomas Bø Andersen (Utbygging - Sotrasambandet)	27
6.6	Marit Due (Vegdirektoratet – Strategi)	28
6.7	Hilde Austlid (Vegdirektoratet – Data, innsikt og analyse)	28

6.8	Einar Aasprong (Drift og vedlikehold)	29
6.8.1	Førsteintrykket	29
6.8.2	Språksvakheten i Copilot Free	30
6.8.3	Eksempler på god språkforståelse	31
6.8.4	Generell kvalitet på innholdet	31
6.8.5	Bildegenerator.....	32
6.8.6	Konklusjon (Faglig bruksverdi)	34
7	Oppsummering og avslutning	35
7.1	Hva har vi lært?.....	35
7.1.1	Ble deltagernes forventninger tilfredsstilt?	36
7.2	Potensielle gevinster ved bruk av Copilot Free.....	36
7.3	Veien videre	37

1 Innledning

Denne sluttrapporten oppsummerer et læringsprosjekt som ble gjennomført i Statens vegvesen våren 2024, og som handlet om bruk av generativ KI.

Rapporten er strukturert slik:

- Kapittel 2 gir en introduksjon til generativ KI, hva det er, hvordan det virker, og hvilke bruksområder det har.
- Kapittel 3 handler om prosjektgjennomføringen, hvordan prosjektet ble organisert og praktisk gjennomført, hvem som deltok og deltagerens motivasjon og forventninger.
- Kapittel 4 handler om verktøyet Copilot Free og hva det kan brukes til.
- Kapittel 5 presenterer vi resultatene fra spørreundersøkelsen vi gjennomførte i etterkant av læringsperioden.
- Kapittel 6 presenterer erfaringer fra 8 av prosjektdeltagerne.
- Kapittel 7 oppsummerer hovedfunnene fra læringsprosjektet.

2 Hva er generativ KI?

Generativ kunstig intelligens er en type teknologi som gir datamaskiner evnen til å forstå og generere naturlig språk. Generativ KI benytter ulike språkmodeller, der en språkmodell er en statistisk modell som har blitt opplært (og i mange tilfeller lærer fortløpende) fra store mengder tekstdata. Den forutsier de mest sannsynlige ordene i en setning basert på konteksten. Språkmodeller brukes i søkemotorer, oversettelsesverktøy og tekstgenerering. Videre bruker generativ KI maskinlæring. Dette betyr at datamaskinen lærer fra eksempler.

Språkmodeller er en type kunstig intelligens (KI) som er utviklet for å forstå og generere naturlig språk. Disse modellene blir trent ved å analysere store mengder tekstdata, ofte hentet fra Internett, for å lære hvordan mennesker vanligvis uttrykker seg i forskjellige situasjoner og kontekster.

En språkmodell fungerer som en statistisk modell som er matet med informasjon slik at den kan bruke maskinlæring til å forutsi de mest sannsynlige ordene i en ordrekke. Den lærer ved å prøve og feile, akkurat som et menneske som ønsker å lære å spille piano blir bedre ved å øve. I starten vil ikke maskinen være spesielt flink, men etter hvert vil den justere seg selv og bli bedre for hver gang den øver. Dette gjør det mulig for språkmodeller å forstå og generere tekst som ligner på hvordan mennesker kommuniserer

2.1 Bruksområder for generativ KI

- Å skape nytt innhold
 - Skrive mail, presentasjoner, rapporter, blogger, artikler, klarspråk
 - Lage bilder
 - Lage video
 - Sparer tid, øker kvalitet, økt kreativitet
- Kodeutvikling
 - Skrive kode raskere, automatisere oppgaver, debugging, optimalisering
- Data-analyse
 - Analysere data og generere innsikt.
 - Oppdage mønstre, trender, og korrelasjoner.
- Prosessdokumentering
 - Dokumentere prosesser og prosedyrer
 - Workflows, redusere feil, samsvar

2.2 Eksempler på tjenester

Det finnes en rekke tilgjengelige tjenester innenfor dette området:

- **ChatGPT:** Chatbasert tjeneste utviklet av OpenAI, som kan generere tekst- og bildebaserte svar og samtaler. Benytter språkmodellen GPT-3.5 i gratisversjon og GPT-4 i betalbar versjon.
- **Dialogflow:** En plattform utviklet av Google som gir verktøy for å bygge chatbots og samtalegrensesnitt. Den støtter naturlig språkbehandling og kan integreres med andre Google-tjenester.
- **Rasa:** Et open-source rammeverk for å bygge chatbots og virtuelle assistenter. Rasa gir muligheten til å lage egendefinerte samtaleflyter og har en aktiv utviklerfellesskap.
- **IBM Watson Assistant:** En KI-plattform fra IBM som lar deg bygge chatbots og virtuelle agenter. Den støtter også integrasjon med andre IBM-tjenester og har avanserte funksjoner som naturlig språkforståelse og kontekstbevaring.
- **Wit.ai:** En annen open-source plattform for naturlig språkbehandling. Wit.ai gir verktøy for å bygge chatbots og kan brukes til å forstå brukerinntut i naturlig språk.
- **Copilot Free (tidligere Bing Chat Enterprise):** Chatbasert tjeneste som benytter OpenAI sin GPT-4-modell for å generere svar.

I dette læringsprosjektet bestemte vi oss for å benytte Copilot Free. Vi skriver mer om dette teknologivalget i [kapittel 4: Teknologivalg: Copilot Free \(Copilot i nettleser\)](#).

3 Prosjektgjennomføring

I dette kapittelet gjør vi rede for hvordan læringsprosjektet ble organisert, administrert og gjennomført.

3.1 Hvordan ble læringsprosjektet gjennomført?

Prosjektet ble initiert av avdelingsdirektørene i Produkter og tjenester IT og Utvikling IT.

Prosjektet ble ledet av en prosjektgruppe med følgende medlemmer:

- Bjørn Terje Bakken (prosjektleder, Forretningsutvikling og innovasjon IT)
- Bjørnar Aassveen (Forretningsutvikling og innovasjon IT)
- Christian Kibsgaard Lunde (Utvikling IT)
- Even Magnus Fjeldstad (Utvikling IT)
- Kine Linde (Utvikling IT)
- Kjetil Ringstad (Utvikling IT)

Prosjektgruppa hadde oppstartsmøte for planlegging av prosjektet i januar 2024. I løpet av januar ble det gjennomført flere planleggingsmøter som resulterte i en overordnet plan for prosjektet med følgende føringer:

- Vi skal ikke legge begrensninger på antall deltagere i læringsprosjektet
- Læringsprosjektet skal omtales og markedsføres gjennom en etatsfelles nyhetssak på Vegveven. Nyhetssaken ble publisert fredag 26. januar 2024.
- Læringsprosjektet skal la deltagerne få benytte **Copilot Free** (nettbasert Copilot, tidligere kalt Bing Chat Enterprise) i 30 dager.
- 30 dagers læringsperiode skal være gjennomført før påske 2024, og skal starte med et felles kickoff-møte på Teams.
Kickoff ble gjennomført fredag 16. februar. Dette var et totimers møte der prosjektgruppa først presenterte opplegg og rammer/kjøreregler for bruk av Copilot, før representanter for Microsoft presenterte Copilot og beste praksis for bruk av dette verktøyet.
- Gjennom læringsperioden skal alt av kommunikasjon mellom deltagere og prosjektgruppe skje i læringsprosjektets samarbeidsrom på Teams.
Det ble etablert et eget Teamsrom for læringsprosjektet med en felleskanal for alle deltagere, samt divisjonskanaler for samarbeid og diskusjoner internt i hver divisjon.
- Prosjektgruppa skal så langt det er mulig poste innlegg daglig på Teams for å opprettholde fokus og læringsaktivitet.
Dette ble gjennomført med et snitt på 4 innlegg pr. uke.
- Læringsperioden skal avsluttes med et felles evalueringsmøte på Teams.
Evalueringsmøtet ble avholdt mandag 18. mars.
- Deltagerne skal få tilsendt et evalueringsskjema/spørreundersøkelse i etterkant av 30-dagerperioden der de blir bedt om å evaluere læringsprosjektet og dele erfaringer.
Spørreundersøkelsen ble sendt ut via innlegg i Generelt-kanalen på Teams 19. mars (dagen etter avslutningsseminaret). Det ble purret på å svare på spørreundersøkelsen 2 ganger etter dette, og pr. 11. april hadde 128 deltagere svart på spørreundersøkelsen. Denne svarprosenten tror vi utgjør et representativt utvalg av deltagerne, siden det virker som om et antall av deltagerne var lite aktive.
- Læringsprosjektet skal resultere i en sluttrapport som skal være ferdig innen utløpet av juni 2024.

3.2 Hvem var deltakerne?

Invitasjonen til å delta i læringsprosjektet vakte stor interesse. Selve nyhetssaken er i skrivende stund likt av 88 personer, og har fått 18 kommentarer. Nyhetssaken er blant sakene med høyest engasjement på Vegveven.

I nyhetssaken ble interesserte deltagere invitert til å sende en e-post til prosjektleder med en kort tekst om egen motivasjon for å delta. Her fikk vi mye nyttig innsikt som vil omtales grundigere i neste delkapittel.

Totalt har læringsprosjektet hatt 275 mer eller mindre aktive deltagere. Disse er fordelt slik på de ulike divisjonene:

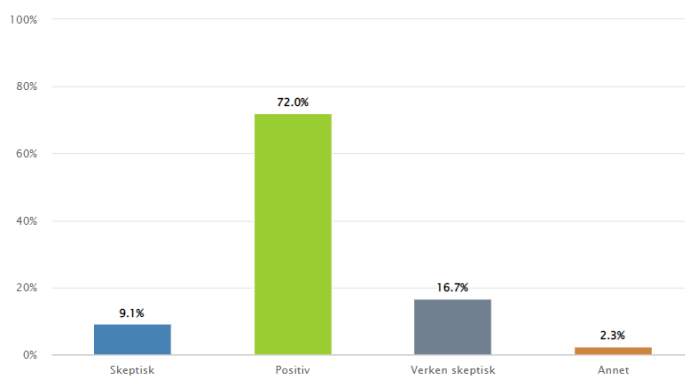
- 26 fra Drift og vedlikehold
- 39 fra Fellesfunksjoner og HR
- 39 fra IT
- 32 fra Trafikant og kjøretøy
- 47 fra Transport og samfunn
- 51 fra Utbygging
- 41 fra Vegdirektoratet

Fordelingen på divisjonene er med andre ord relativt jevn, og vi kan dermed si at vi har jevnt god deltagelse fra hele etaten på dette læringsprosjektet.

3.2.1 Deltagernes holdning til generativ KI før oppstart

I den avsluttende spørreundersøkelsen spurte vi om hvilken holdning deltagerne hadde til generativ KI før testperioden. Vi fikk følgende svar:

Hvilken holdning hadde du til språkbasert/generativ KI før testperioden?

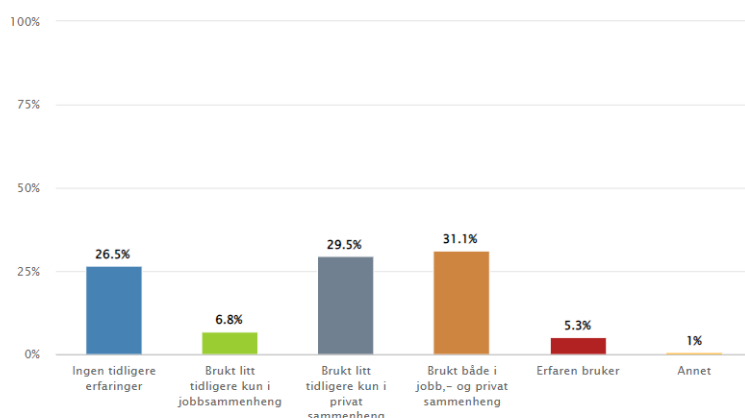


Vi ser at nesten $\frac{3}{4}$ av deltagerne var positive til teknologien i utgangspunktet. Ca. 10 % var skeptiske til teknologien, mens ca 17 % var verken positive eller skeptiske. Disse tallene er det interessant å ha i mente når vi evaluerer utbyttet av læringsprosjektet får tilbakemelding på om dette er teknologi som deltagerne vil fortsette å benytte også etter at læringsprosjektet er ferdig.

3.2.2 Deltagernes erfaringer med generativ KI før oppstart

I den avsluttende spørreundersøkelsen spurte vi også om hvilken erfaring deltagerne hadde til generativ KI før testperioden. Vi fikk følgende svar:

Hvilke erfaringer hadde du med språkbasert/generativ AI før oppstart?



Vi ser at litt over ¼ av deltagerne var helt uten erfaringer med teknologien i utgangspunktet. De øvrige deltagerne hadde mer eller mindre erfaring med denne eller tilsvarende teknologier fra før. Ca. 30 % av deltagerne hadde kun privat erfaring. Ca. 40 % av deltagerne hadde benyttet denne eller tilsvarende teknologier i jobbsammenheng tidligere.

3.3 Hva var deltagerens motivasjon for å delta i prosjektet?

Alle deltagere ble bedt om å sette ord på egen motivasjon for å delta i prosjektet. De aller fleste som sendte e-post med forespørsel om å delta gjorde dette. Her er noen utvalgte sitater fra påmeldingene:

«Dels synes jeg KI virker veldig spennende og så tror jeg at jeg hadde hatt stor nytte av det. Jeg er svensk og jeg synes jeg bruker unødig lang tid på meg for å formulere e-poster på norsk.»

«Jeg er veldig nysgjerrig på hvordan KI kan utnyttes i min arbeidshverdag og ønsker å være litt fremoverlent på dette området (...selv om jeg må innrømme at jeg synes det er litt (les: veldig) skummelt og kan være en utfordring i forhold til personvern).»

«Jobber i styringsstab TOS og koordinerer mye høringer som skal sammenfattes til et felles innspill. Skriver også en del egentekst basert på innspill fra transportavdelinger på en tilskuddsordning (for å nevne en sak).»

«Jeg jobber på Økonomi og virksomhetsstyring (ØKV), på avdelingen Mål og resultat. Jeg har ansvar for SVV sin årsrapport, og i disse dager jobber vi inn mot fristen den 15. mars for å levere til Samferdselsdepartementet. Jeg er nysgjerrig på hvordan jeg kan bruke KI i dette arbeidet, og ønsker derfor å være med å teste.»

«... Ellers skriver jeg planer og gjør analyser i jobben, samt lager en del presentasjoner og kanskje KI kan hjelpe meg å gjøre disse mer interessant? Hvis jeg begynner å lære meg litt om KI, så kan jeg kanskje lettere lære meg mer og dermed se om KI kan brukes som et verktøy i HMS-analyser. Der er det nemlig potensielt mye å hente for SVV.»

«Inntil nylig så jeg på KI som et litt nytteløst verktøy som kunne pumpe ut innhold uten substans og mening, men har jo da forstått at det er under rivende utvikling, og tok til meg en kommentar om at «du skal ikke være redd for at du mister jobben til KI, men du skal være redd for at du mister jobben til en som behersker KI». Etter dette har jeg lekt litt med ChatGPT, og har et visst innsyn i teknologien, men har bare fått økt interesse etter denne oppvåkningen i november.»

«Jeg er nysgjerrig på om dette kan være nyttig i forbindelse med regelverksutvikling.»

«Er nysgjerrig på hvordan jeg kan utnytte KI og KI-verktøy for å bli mer effektiv i arbeidshverdagen og øke kvaliteten på leveranser. Digitalisering og bruk av KI er også et satsingsområde i etaten, og tenker vi medarbeidere må bidra til å følge opp satsingen.»

«Vi utarbeider også ulike støttedokument, deriblant konteringsveiledning. Kanskje kan vi bruke det til å lage enklere oppsummeringer i stedet for å lage brukerveiledninger manuelt. Kan det være et alternativ til nanolæringene?»

«Jeg jobber med vegnormaler og digitale vegnormaler. Jeg ønsker å teste generativ KI for å finne ut om den måten vegnormaler er digitalisert kan tolke riktig og kan bruke i fremtiden KI/AI.»

«Skriver mye tekster, bl.a forenkling av rapporter og mer komplekse saker. Er en av etatens største produsenter av pressemeldinger, Vev-saker, talepunkter til vegdirektøren, m.m.»

«Jeg tenkte at dette kunne være en god introduksjon for en som i utgangspunktet er litt skeptisk i forhold til «hvor dataen blir av» når man bruker KI.»

«Fra å ha ignorert temaet (for) lenge ønsker jeg nå å oppdatere meg, lære å bruke fornuftig verktøy og metodikk og å lære om potensialet som finnes. Jeg ønsker også å gjøre dette i trygge rammer der datasikkerhet og etikk ivaretas.»

«Jeg ønsker å være med, fordi verktøyet kan være nyttig i min arbeidshverdag som jurist/kontraktsrådgiver, kontrakt og marked i Utbyggingsdivisjonen. Ved å delta vil jeg kunne lære å bruke verktøyet på en sikker måte og øke kunnskapen om KI.»

Vi kan oppsummere deltagernes motivasjon for å delta i læringsprosjektet i følgende fem punkter:

1. **Bruk av KI i tekstproduksjon og oversettelse:** KI kan spille en rolle i produksjon og oppsummering av rapporter og tekster, samt språkoversetting. Det kan også være et verktøy i arbeidet med produksjon av egne tekster/rapporter og kvalitetssikring av rapporter fra konsulenter.
2. **Effektivisering og forenkling av oppgaver:** KI kan potensielt forenkle oppgaver som utføres i dag og være et nyttig verktøy for å effektivisere en del arbeidsoppgaver, eksempelvis produksjon av nyhetsmeldinger, oppsummering av tekst, møtetreferat, planlegging av agenda og oppsett for arbeidsmøter, innlegg og presentasjoner.
3. **Interesse for KI i daglig arbeid:** Det er interesse for å se hvordan KI kan brukes i daglig arbeid, og for samspillet mellom data/teknologi og mennesket. Dette inkluderer også en interesse for å øke den “pedagogiske merverdien” i arbeidet.
4. **KI i konkret saksbehandling:** KI kan kanskje brukes til å utvikle automatiske prosesser, eksempelvis for saksbehandling ifm. tilrettelagt teoriprøve?
5. **Investering for fremtiden:** Prosjektet kan bidra til å ta i bruk verktøyet og opparbeide kunnskap som benyttes i fremtiden. KI kan være nyttig i arbeidet med å produsere mye tekst, enten det er å svare på henvendelser, gi råd, utvikle regelverk, skrive høringer.

4 Teknologivalg: Copilot Free (Copilot i nettleser)

Som nevnt valgte vi å bruke Microsoft sin tjeneste Copilot Free i dette læringsprosjektet. Hovedargumentet for dette var den innebygde sikkerheten Microsoft har i denne tjenesten.

4.1 Om verktøyet Copilot Free

Copilot Free er en kunstig intelligens (KI) utviklet av Microsoft. Den er designet for å hjelpe brukere med en rekke oppgaver, fra å svare på spørsmål og gi informasjon, til å generere kreativt innhold som tekster, koder, sanger og mer.

Copilot bruker avanserte KI-teknologier, inkludert GPT-4 og Bing Search, for å forstå brukerens behov og gi relevante og nyttige svar. Den kan kommunisere flytende på mange språk, inkludert engelsk, spansk, fransk, tysk, kinesisk, og mange flere.

Copilot kan hjelpe brukere med en rekke oppgaver. For eksempel, hvis en bruker trenger hjelp med å skrive en tekst, kan Copilot gi forslag og hjelpe med å forbedre teksten. Den kan også generere kreativt innhold, som dikt, historier, kode, essays, sanger, og mer. I tillegg kan Copilot svare på spørsmål og gi informasjon om et bredt spekter av emner.

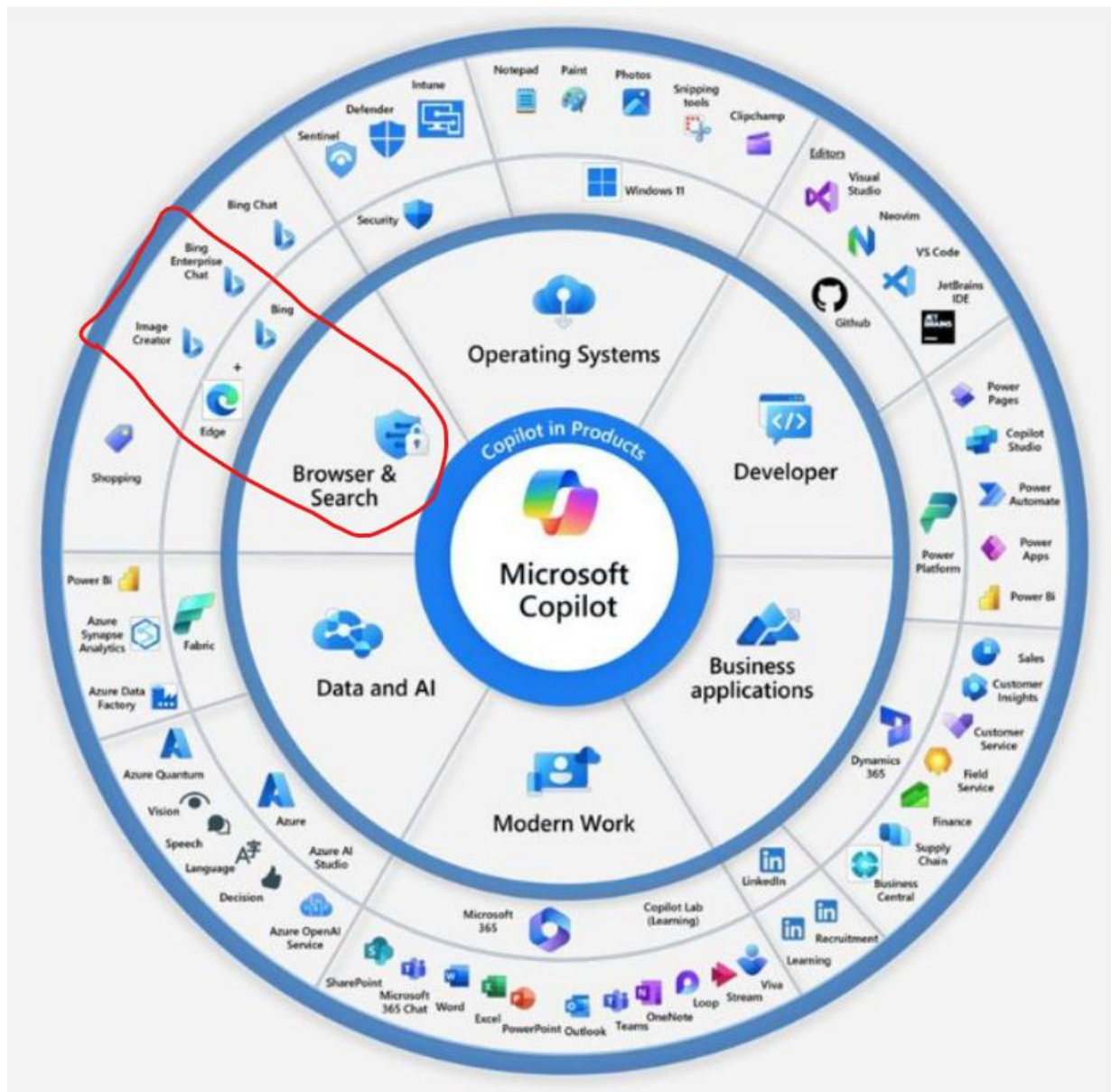
Copilot er designet for å være brukervennlig og enkel å bruke for alle, uavhengig av deres tekniske ferdigheter. Det betyr at alle i organisasjonen, fra ledere til ansatte, kan dra nytte av de mange funksjonene og fordelene som Copilot tilbyr.

Å bruke Copilot Free er ganske enkelt og intuitivt, selv for de som ikke er spesielt datakyndige. Her er noen grunnleggende trinn:

1. Start en chat: Du kan starte en chat ved å åpne Copilot på datamaskinen din. Det kan være en app på datamaskinen din, eller det kan være en nettside du går til.
2. Skriv en melding: Når du har åpnet chatten, vil du se et sted hvor du kan skrive inn meldinger. Du kan begynne å skrive hva som helst du vil spørre eller si til assistenten.
3. Send meldingen: Når du har skrevet meldingen din, trykker du på "send" knappen for å sende meldingen. Assistenten vil da motta meldingen din og begynne å behandle den.
4. Vent på svar: Etter at du har sendt meldingen, vil assistenten begynne å behandle den og komme med et svar. Dette kan ta noen sekunder, så vær tålmodig.
5. Les svaret: Når assistenten har svart, vil du se svaret i chatten. Du kan lese svaret og deretter skrive en ny melding hvis du har flere spørsmål eller trenger mer hjelp.

Hvorfor anbefaler vi Microsoft Copilot med Copilot Chat som grensesnitt?

- Tjenesten inngår allerede i Statens vegvesen sine lisenser.
- All data om person og virksomheter er beskyttet.
- Kan enkelt nås fra Edge på kun 2 tastetrykk.
- Tilgjengelig for alle i vegvesenet, ingen installasjon eller registrering.
- Bygger på samme teknologi som ChatGPT (GPT-4)
- Samtalestiler: Creative vs. Balanced vs. Precise



5 Resultater fra spørreundersøkelsen

I etterkant av prosjektperioden ba vi prosjektdeltagerne om å delta på en spørreundersøkelse for å evaluere både prosjektgjennomføringa og verktøyet Copilot Free. Vi fikk ca 130 svar, dette utgjør ca 50 % av prosjektdeltagerne. Vi antar at dette tallet gir et relativt korrekt bilde av hvor mange av deltagerne som var aktive i den fire uker lange prosjektperioden.

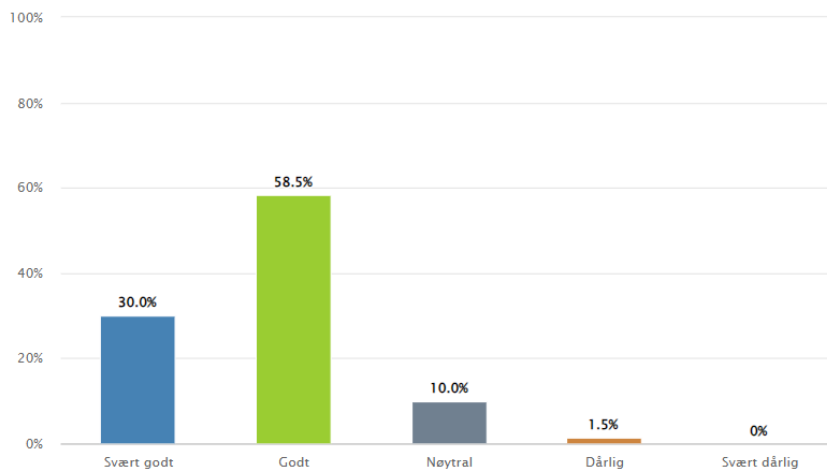
I dette kapittelet presenteres resultatene fra spørreundersøkelsen. Kapittel 4.1 fokuserer på spørsmålene som handlet om prosjektgjennomføringen, mens kapittel 4.2 fokuserer på spørsmålene som handlet om verktøyet Copilot Free.

5.1 Evaluering av prosjektgjennomføringen

5.1.1 Informasjonsflyt

I spørreundersøkelsen stilte vi spørsmålet «Hva er ditt generelle inntrykk av informasjonsflyten i prosjektet?». Svarene fordeler seg slik:

Hva er ditt generelle inntrykk av informasjonsflyten i prosjektet?

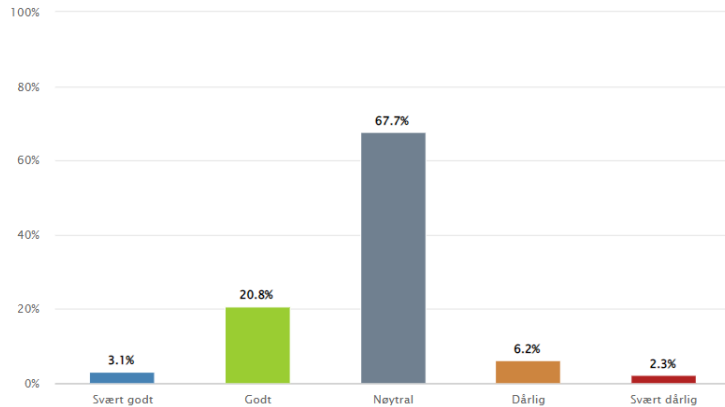


Vi ser at ca 88,5 % er svært godt eller godt fornøyd med informasjonsflyten i prosjektet. Vi kan dermed konkludere med at den valgte tilnærmingen (med oppstartsseminar, avslutningsseminar og hyppige Teams-innlegg i læringsperioden) fungerte godt, og at deltagerne fikk den informasjonen de trengte i prosjektgjennomføringen.

5.1.2 Samarbeid med andre

I spørreundersøkelsen stilte vi videre spørsmålet «Hvordan opplevde du at ev. samarbeid med andre prosjektdeltagere i divisjonen din fungerte?». Svarene fordeler seg slik:

hvordan opplevde du at ev. samarbeid med andre prosjektdeltagere i divisjonen din, fungerte?



Majoriteten av respondentene svarer at de opplevde samarbeid i egen divisjon som «Nøytral». Vi erfarte at det var forsvinnende liten aktivitet i de divisjonsvise kanalene i Teams-rommet, og går ut fra at dette er hovedårsaken til at svarene fordeler seg slik de gjør.

Vi kan anta at de som opplevde samarbeid i egen divisjon som svært godt eller godt trolig har samarbeidet på andre arenaer utenfor Teams, eller at de har vært flere fra samme divisjon som har vært aktive i den generelle kanalen i Teams-rommet.

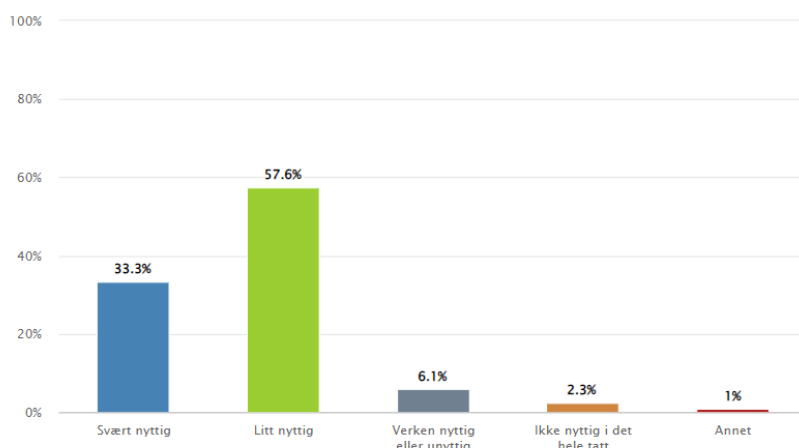
Vi kan videre anta at de som opplevde samarbeid i egen divisjon som dårlig eller svært dårlig trolig har savnet aktivitet hos nære kolleger, og at de trolig også har savnet prosjekttiltak som kunne ha økt aktiviteten i de ulike fagmiljøene.

Vi konkluderer med at vi kunne ha lagt bedre til rette for felles aktiviteter i fagområdene gjennom læringsperioden. Dette vil vi også komme tilbake til under punktet Generelle tilbakemeldinger.

5.1.3 Nytte av å delta

I spørreundersøkelsen stilte vi videre spørsmålet «Synes du deltagelse i prosjektet har vært nyttig for deg?». Svarene fordeler seg slik:

Synes du deltagelsen din i prosjektet har vært nyttig for deg?



Over 90 % av respondentene svarer at de opplevde deltagelsen i prosjektet som nyttig eller svært nyttig. Dette blir også bekreftet av de generelle tilbakemeldingene i spørreundersøkelsen.

5.1.4 Generelle tilbakemeldinger på prosjektgjennomføringen

Avslutningsvis i den delen av spørreundersøkelsen som handlet om metode stilte vi et åpent spørsmål som lød «Hvilke tilbakemeldinger vil du gi til prosjektgjennomføringen/denne måten å drive læring på?». Vi fikk totalt 76 svar, og i det følgende presenterer vi noen av disse. De er sortert på «positive tilbakemeldinger» og «Tilbakemeldinger som inneholdt forbedringspunkter».

5.1.4.1 Positive tilbakemeldinger

«Jeg kan ikke se at det kan gjøres bedre enn det ble gjort her.»

«Veldig bra tiltak!»

«Bra måte å drive læringsprosjekt på!»

«Veldig god opplevelse og informasjonsflyt!»

«Jeg likte veldig godt at det var så lavterskel.»

«Relevant måte å drive læring på. Mer av dette!»

5.1.4.2 Tilbakemeldinger som inneholdt forbedringspunkter

«Jeg har testet alene. Personlig synes jeg det er "skremmende" å jobbe på et stort felles forum. Underveis hadde jeg hatt mye å dele og å spørre om. I en mindre prosjektgruppe med andre som aktivt testet, si 5 personer, tror jeg informasjonsflyten hadde blitt en helt annen. Mer struktur rundt utprøvingen hadde også vært bra. Jeg er ganske sikker på at det er enkel funksjonalitet jeg fortsatt ikke behersker.»

«Det kunne kanskje vært litt mer opplæring på hvordan man skal skrive gode prompter, men det er vel en del av læringen.»

«Litt for lite pauser både på opplæring + avslutningsseminaret. Det er tungt å stirre på en skjerm i flere timer i strekk uten pause.»

«Mer fokus på å dele erfaringer underveis i prosjektet.»

«Det ville vært bedre med litt "undervisning" underveis.»

«Som det ble nevnt av en av de som presenterte erfaringen sin, så ville det vært nyttig å kunne samles en gang i uken/annenhver uke (f.eks. digitalt, frivillig, kun 30 minutter) basert på kunnskapsnivå/digital kompetanse for å dele erfaringer og tips til hvordan copilot kan brukes i hverdagen. Opplever at det er mer å lære ved å møte de på eget kunnskapsnivå, heller enn de som tilhører egen divisjon, da jeg tidvis har sittet igjen med opplevelsen av å bruke tiden på å lære andre heller enn bli utfordret selv.»

«Tror man kanskje har undervurdert hvor ivrige prosjektdeltakere kan være. Det er stort engasjement, og mange ønsket å komme i gang med en gang man fikk litt mer info. Det var også noe tid før vi fikk sidepanelet. Kanskje man heller burde vurdert å flytte prosjektet? For min del fikk jeg ikke så stort utbytte av funksjonene før jeg fikk sidepanelet på plass.»

«Jeg savnet informasjon om konkrete ting Copilot kunne ha vært brukt til i disse ukene prosjektet pågikk. Det var så nyttig å se hva mine kollegaer hadde funnet ut og prøvd seg på, som f.eks. det med bruk av skriv funksjonen som kunne knyttes til e-post. Tips om sånne funksjoner underveis hadde vært supert.»

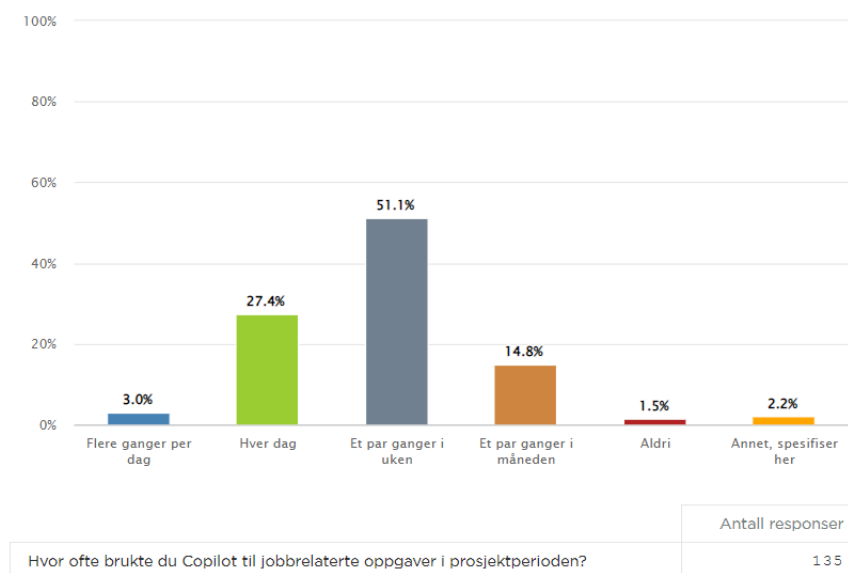
«Kick-off var veldig informativt, med gode praktiske eksempler på bruk, presentert av eksterne foredragsholdere/selgere :-). Prosessen mellom kick-off og avslutningswebinar, med egen forskning på bruk av verktøy ble veldig sporadisk i en hektisk hverdag på kjøretøy avdelingen. Syns tiden ble litt knapp også. Avslutningswebinaret var igjen bra gjennomført med erfaringer fra flere og god informasjon om løsninger. Kort oppsummert: Noe amputert testperiode med bruk av den enkle Copiloten versjonen med begrensa funksjoner, med ser frem til pro-versjonen kan testes.

5.2 Evaluering av valgt teknologi

5.2.1 Hvor ofte brukte du Copilot til jobbrelevante oppgaver i prosjektperioden?

I spørreundersøkelsen stilte vi spørsmålet «Hvor ofte brukte du Copilot til jobbrelevante oppgaver i prosjektperioden?». Svarene fordeler seg slik:

Hvor ofte brukte du Copilot til jobbrelevante oppgaver i prosjektperioden?



Her ser vi at av respondentene svarer om lag 80 % at de brukte Copilot i nettleser minst et par ganger i uken. 30 % av respondentene brukte Copilot hver dag eller flere ganger hver dag. Selv om vi har fått tilbakemeldinger på at det burde vært flere «arrangerte aktiviteter» i prosjektet ser vi altså at et stort flertall av deltagerne drev med aktiv bruk av Copilot gjennom prosjektperioden.

5.2.2 Hvilke arbeidsoppgaver så du for deg at KI kunne hjelpe deg med?

Vi har både gjennom påmeldingene og via spørreundersøkelsen fått mange innspill til hva deltagerne ønsket å bruke Copilot Free til. Vi foret Copilot med alle disse innspillene, og ba om å få en oppsummering i 5 punkter. De følgende 5 punktene er Copilot sin oppsummering av deltagerens tanker om hvilke arbeidsoppgaver Copilot kunne gjøre for dem:

- 1. Tekstproduksjon og oversettelse:** KI kan spille en rolle i produksjon og oppsummering av rapporter og tekster, samt språkoversetting. Det kan også være et verktøy i arbeidet med produksjon av egne tekster/rapporter og kvalitetssikring av rapporter fra konsulenter.
- 2. Effektivisering og forenkling av oppgaver:** KI kan potensielt forenkle oppgaver som utføres i dag og være et nyttig verktøy for å effektivisere en del arbeidsoppgaver,

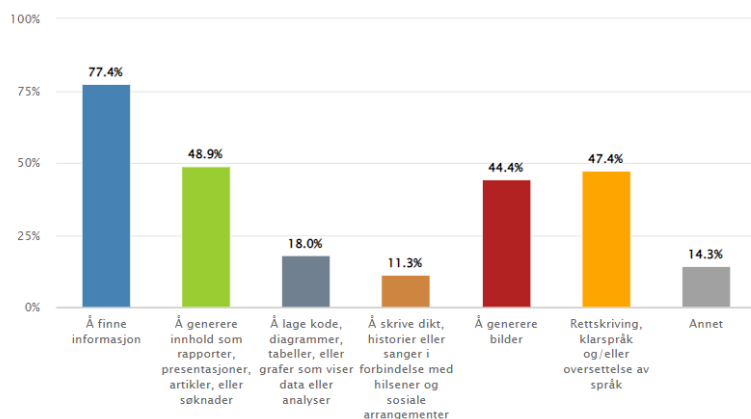
eksempelvis produksjon av nyhetsmeldinger, oppsummering av tekst, møtereferat, planlegging av agenda og oppsett for arbeidsmøter, innlegg og presentasjoner.

3. Interesse for KI i daglig arbeid: Det er interesse for å se hvordan KI kan brukes i daglig arbeid, og for samspillet mellom data/teknologi og mennesket. Dette inkluderer også en interesse for å øke den "pedagogiske merverdien" i arbeidet.
4. KI i saksbehandling: KI kan kanskje brukes til å utvikle en automatisk prosess for saksbehandling for tilrettelagt teoriprøve?
5. Fremtidig bruk av KI: Prosjektet kan bidra til å ta i bruk verktøyet og opparbeide kunnskap som benyttes i fremtiden. KI kan være nyttig i arbeidet med å produsere mye tekst, enten det er å svare på henvendelser, gi råd, utvikle regelverk, skrive høringer.

5.2.3 Hvilke jobbrelevante oppgaver brukte du Copilot til?

I spørreundersøkelsen stilte vi spørsmålet «Hvilke jobbrelevante oppgaver brukte du Copilot til?». Svarene fordeler seg slik:

Hvilke jobbrelevante oppgaver brukte du Copilot til?



	Antall responser
Hvilke jobbrelevante oppgaver brukte du Copilot til?	133

Den arbeidsoppgaven som flest benyttet Copilot til var «Å finne informasjon». Dette er samtidig en oppgave mange har opplevd utfordringer med, siden Copilot i nettleser presenterer til dels mye faktafeil i sine svar. Dette må trolig tilskrives det faktum at mange kilder på internett har feil eller utdatert informasjon. Se mer om dette i kapittel 5.2.6.

Videre ble Copilot benyttet av et relativt likt antall respondenter til oppgavene «Å generere innhold som rapporter, presentasjoner, artikler eller søknader», «Å generere bilder» og «Rettskriving, klarspråk og/eller oversettelse av språk». Mellom 45 % og 50 % av respondentene svarte at de brukte Copilot til disse oppgavene.

Et mindre antall respondenter brukte Copilot Free til «Å lage kode, diagrammer, tabeller, etc» og «Å skrive dikt, historier eller sanger i forbindelse med hilsener og sosiale arrangementer».

Av de som svarte «Annet» ble følgende andre oppgaver spesifisert:

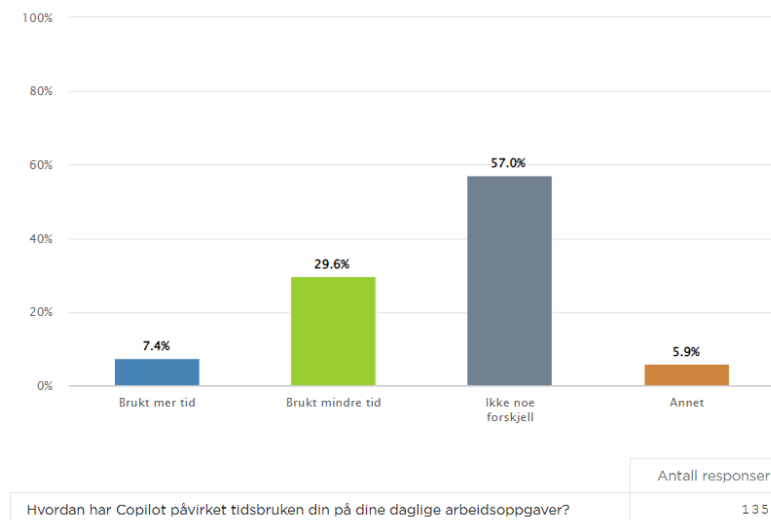
- Begrepspresisering på engelsk
- Regnestykker.
- Definisjonsavklaring av faguttrykk
- Som brukerstøtte til flere programvarer

- Tekstsammenligninger, møtenotater, befaringsrapporter.
- Oppsummere store og tunge dokumenter
- Utkast til vedtaksmaler og eposter
- For å få bekreftet at mine tanker er på rett vei
- Lage resyme og innledninger
- Omgruppere/sortere kulepunkter i egne tekster etter tema
- Oppsummeringer og referat, beslutningssporing, «hva betyr dette offentlige skrevet egentlig?»
- For å lære om KI
- Forbedre tekst, sammendrag av tekst, finne gode formuleringer
- Idemyldring

5.2.4 Hvordan har Copilot påvirket tidsbruken din på dine daglige arbeidsoppgaver?

I spørreundersøkelsen stilte vi videre spørsmålet «Hvordan har Copilot påvirket tidsbruken din på dine daglige arbeidsoppgaver?». Svarene fordeler seg slik:

Hvordan har Copilot påvirket tidsbruken din på dine daglige arbeidsoppgaver?



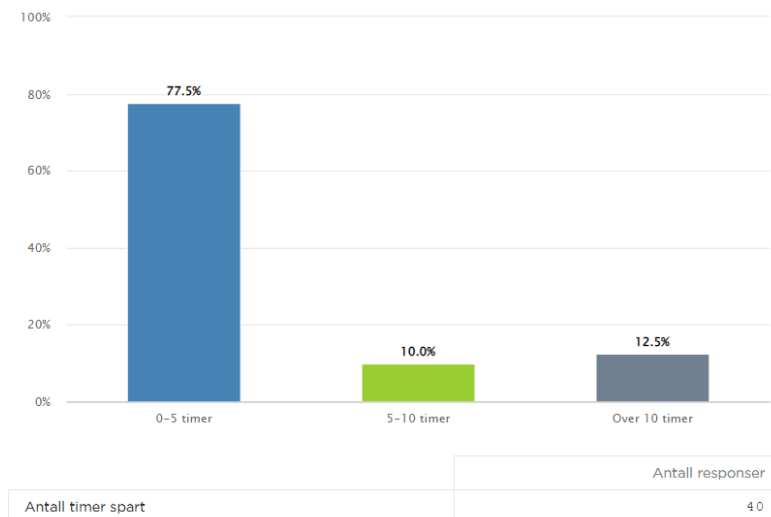
Et flertall av respondentene (57 %) svarte at bruk av Copilot Free ikke medførte noen forskjell i tidsbruk på egne arbeidsoppgaver. 29,6 % svarte at de brukte mindre tid på egne arbeidsoppgaver, mens et fåtall (7,4 %) svarte at de brukte mer tid på egne oppgaver. 5,9 % svarte Annet, og spesifiserte det slik:

- Noe har gått raskere, andre ting har tatt lengre tid på grunn av kvalitetssjekk og korrigering.
- Dette er et prøveprosjekt, jeg har ikke begynt å bruke Copilot fast i min stilling i SVV.
- ChatGPT4 har ført til at jeg bruker mye mindre tid på å kode
- Litt vanskelig å si tidsbruk, det tar tid å bli kjent med noe nytt.
- Vanskelig å si, men jeg tror tekstene blir bedre og forståelsen på søk blir bedre
- Litt mer tid grunnet mangel på kunnskap rundt bruken av KI, men ser merverdien av å bruke mindre og mindre tid på enkelte oppgaver etter hvert som kunnskapen øker.
- Usikker - må teste mer
- Vet ikke

5.2.4.1 Antall timer spart

De 29,6 % av respondentene (40 stykker) som svarte at de brukte mindre tid på egne oppgaver enn før fikk videre spørsmål om hvor mye tid de hadde spart i spørsmålet «Antall timer spart». Svarene fordeler seg slik:

Antall timer spart



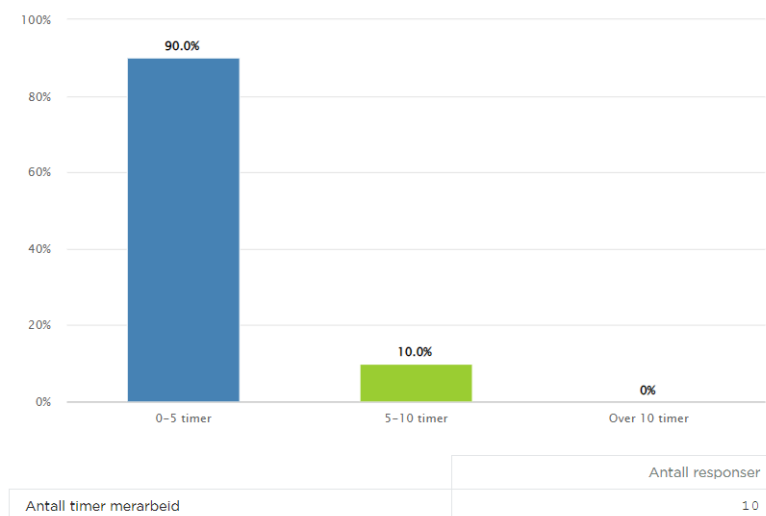
Vi ser at de aller fleste respondentene (77,5 %) sparte mellom 0 og 5 timer. 10 % mener at de sparte 5-10 timer, og 12,5 % mener at de sparte over 10 timer.

Dette er høye tall, spesielt med tanke på at dette var et læringsprosjekt der respondentene antageligvis ikke veldig høy kompetanse på bruk av generativ KI i utgangspunktet. Vi tror dermed at potensialet for å bli mer effektiv ved bruk av KI i Statens vegvesen er høy.

5.2.4.2 Antall timer merarbeid

De 7,4 % av respondentene som svarte at de brukte mer tid på egne oppgaver enn før fikk videre spørsmål om hvor mye mer tid de hadde brukt i spørsmålet «Antall timer merarbeid». Svarene fordeler seg slik:

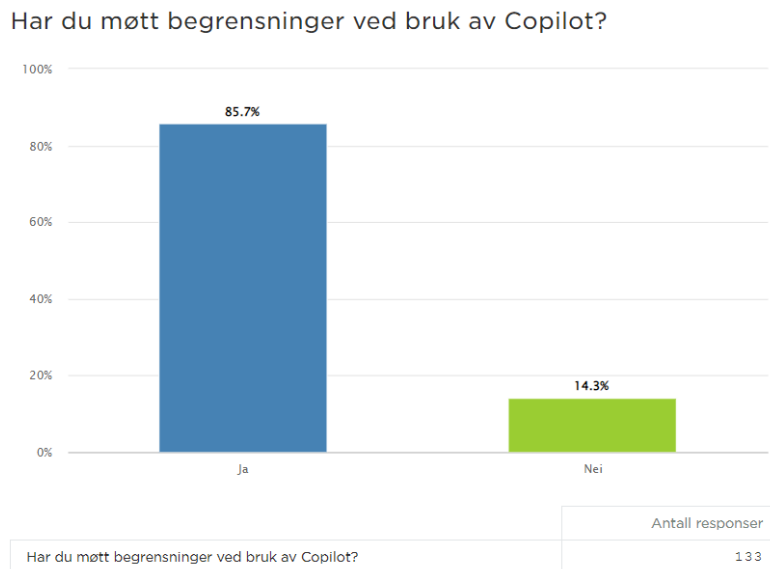
Antall timer merarbeid



Det var altså 7,4 %, eller 10 respondenter, som svarte at de brukte mer tid. Av disse svarte 90 % (9 stykker) at de brukte 0-5 timer mer, mens 10 % (1 person) svarte at hen brukte 5-10 timer mer.

5.2.5 Har du møtt begrensninger ved bruk av Copilot?

Videre stilte vi spørsmålet «Har du møtt begrensninger ved bruk av Copilot?». Svarene fordeler seg slik:



Ca. 85 % av respondentene svarte at de hadde møtt begrensninger ved bruk av Copilot. Det store flertallet opplevde med andre ord at de hadde oppgaver de ikke fikk utført, eller at oppgavene de prøvde å løse ble mangelfullt utført. Vi skal se på hvilke begrensninger dette dreier seg om i det følgende.

5.2.5.1 Hvilke begrensninger?

Dette spørsmålet kunne besvares i fritekst. Vi fikk 108 svar. Noen av de konkrete svarene er gjengitt under. I tillegg ba vi Copilot om å oppsummere svarene i 8 punkter:

1. **Analyse av informasjon i Office-programmer:** Brukerne har opplevd det som en begrensning at de ikke kan bruke Copilot Free til å analysere innhold og data i Office-programmer, f.eks. Excel-filer.
2. **Ikke oppdatert informasjon:** Brukerne nevner at informasjonen de får ikke alltid er oppdatert, noe som kan påvirke nøyaktigheten av analysene.
3. **Tekstbegrensninger:** Brukerne mener at begrensningen på antall ord i teksten som legges inn kan påvirke dybden og detaljrikheten i analysene.
4. **Begrensning på bildegenerering:** Det er en begrensning på hvor mange bilder man kan generere i løpet av et døgn, noe som kan påvirke visuell representasjon av data.
5. **Bruk av interne tekster:** Brukerne har ikke «fått lov til» å bruke interne tekster, noe som begrenser bruken av Copilot i deres arbeid.
6. **Språk og forståelse:** Noen brukere har bemerket at det kan være betydelig forskjell i resultatene avhengig av om man søker på norsk eller engelsk. Copilot misforstår ofte spørsmålene hvis de ikke er presise nok.
7. **Feil og inkonsistenser:** Brukerne har rapportert om flere feil og inkonsistenser i svarene fra Copilot. Dette inkluderer feilaktig og unøyaktig informasjon, samt inkonsistente svar på samme spørsmål.
8. **Brukervennlighet og funksjonalitet:** Brukerne har uttrykt frustrasjon over at Copilot ikke er integrert med visse programmer som Excel og Visual Studio Code. De har

også nevnt at det kan være utfordrende å få Copilot til å forstå komplekse problemstillinger.

Eksempler på konkrete tilbakemeldinger:

«Begrensningen ligger i at hovedarbeidet mitt er avhengig av å bruke interne tekster, så det er kun til bi-oppgavene hvor jeg har kunnet bruke Copilot. I de oppgavene jeg har jobbet med Copilot, så ser jeg at fagfeltet er såpass smalt at det tidvis er vanskelig å formulere gode nok spørsmål til at jeg får (riktige) svar.»

«Begrensninger kan være et resultat av for lite kunnskap, men det har flere ganger vært betydelig forskjell om man søker på norsk eller engelsk. Det er også mye av mine spørsmål er mot Eurokode/Standarder som ikke gir konkrete svar. Jeg har ikke fått Copilot til å lese en pdf-fil, og jeg får heller ikke svar når jeg benytter høyre sidepanel mens jeg har en nettside fremme. Copilot ser ut til å søke generelt på nettet og ikke på den nettsiden jeg har fremme.»

«Ble ikke helt rett, men mulig jeg ikke "mata" den nok. Skulle også gjerne ha funnet mer ut hvordan den kan lage power point presentasjoner for meg.»

«Bruker masse tid på å lage prompts for å få resultatet jeg ønsker - kan da likegodt gjøre jobben selv. Jeg ber Copilot finne svar i et dokument, men den henter svar på spørsmålet fra andre kilder også - selv om jeg har spesifisert at den skal kun bruke et bestemt dokument. Kan ha noe med innstillinger av samtalestil å gjøre, men likevel rart når man instruerer om kun bruk av et dokument. Synes også Copilot kan fristille seg vel mye fra tekst som jeg ber den om å bearbeide på en eller annen måte, budskapet blir helt annet fordi det er tatt med et ord som ikke er i den opprinnelige teksten (f.eks. i opprinnelig tekst: "utredet" og Copilot foreslår "grundig utredet"). Men igjen kan det ha noe med samtalestilen å gjøre.»

«Copilot henter ikke de data jeg trenger fra åpne kilder som proff.no.»

«Copilot klarer ikke å generere bilder etter enkle instruksjoner, klarer ikke lage enkle flytskjemaer, sliter med å håndtere kunstpauser når det dikteres møtereferat, viser ikke de mest relevante søkeresultatene til tross for at Bing viser de øverst, håndterer feil på en dårlig måte mot bruker.»

«Copilot misforstår ofte spørsmålene hvis man er ikke presis nok, noe som kan føre til unødvendig tidsbruk. Det tok flere spørsmål før den forsto at jeg skulle ha data fra nettsiden jeg var på.»

«Copilot ville ikke lage alle bildene jeg ønsket. Jeg har opplevd at jeg ikke finner svaret jeg får når jeg leser igjennom referansen.»

«Da jeg prøvde å teste pålitelighet / troverdighet av Copilot, fant jeg flere inkonsistenser. 1. Ba om statistikker av ting nyere enn 2021. Copilot svarte riktig, da jeg spurte hvorfor hadde den tilgang til nyere data enn 2021 sa Copilot beklager, du har rett. (Ikke sa noe om trenings datasett er 2021 men kan ha tilgang til nyere data eller noe sånt.) 2. ulike svar på samme prompt 3. feilkilder.»

«De vanlige tingene med KI. Feil bildegenerering, feil kalkulering, generiske svar til tross for at jeg ber om det motsatte, manglende variasjon til tross for at jeg ber om det osv.»

«Den er ikke så god til å svare på alt, spesielt om man ber den hente informasjon. Lærte litt mer når jeg fikk med meg essensen i RICCE og kunne formulere bedre men den sliter litt med å hente eksakt informasjon.»

«Den genererer en del feilinformasjon om tema som etterspørres. Dette gjør at det ikke blir like pålitelig å bruke i f.eks. rapporteringssammenheng.»

«Den kan fort gi feilaktig og unøyaktig informasjon om man ber om 'fakta'.»

«Den lager ikke like gode søk som ChatGPT.»

«Det er av og til direkte feil i logikken, og tekster kan se overfladisk flott ut men være meningsløst, så det er kun et utgangspunkt. Jeg bruker det mer til å skaffe grunnlag enn til selve skrivingen.»

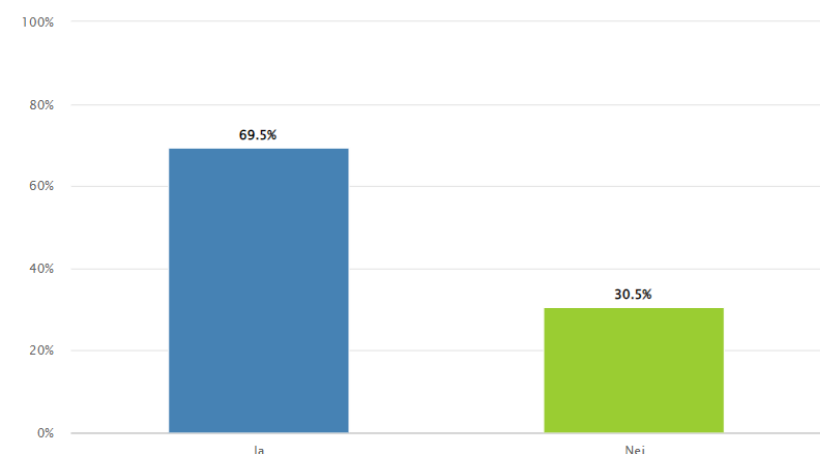
«Det er et stort spørsmål. Nevner noen: - Den norske språkmodellen - Gikk fort "tom for forsterkninger" som Copilot kaller det når man har generert så mange bilder som man får lov til - Begrensninger i tilgjengelig bakgrunnsdokumenter. Skulle veldig gjerne hatt vitenskapelige artikler tilgjengelige. - Klarte ikke å jobbe tilstrekkelig godt med store dokumenter, verken på nett eller i sidevindu.»

«Det er jeg selv som ikke har fantasi nok :)»

5.2.6 Har du møtt på faktafeil i svarene?

Vi stilte flere spørsmål som dreide seg om mulige utfordringer ved bruk av Copilot. Neste spørsmål i spørreundersøkelsen var «Har du møtt faktafeil i svarene?». Respondentene svarte slik:

Har du møtt på faktafeil i svarene?



	Antall responser
Har du møtt på faktafeil i svarene?	131

Hele 69,5 % av respondentene opplevde faktafeil i svarene. Som nevnt tidligere (kapittel 5.2.3) går vi ut fra at dette skyldes det faktum at mange kilder på internett inneholder feil eller utdatert informasjon. Dette utgjør en risiko for at tilsynelatende velformulerte og korrekte svar «går gjennom den menneskelige kvalitetskontrollen», selv om innholdet rent faktisk er feil.

I det følgende viser vi noen eksempler på hvilke faktafeil respondentene rapporterer.

5.2.6.1 Hvilke eksempler på faktafeil?

Dette spørsmålet kunne besvares i fritext. Vi fikk 69 svar. Noen av de konkrete svarene er gjengitt under. I tillegg ba vi Copilot om å oppsummere svarene i 8 punkter:

1. Deltagerne har opplevd at Copilot har gitt **feil eller utdatert informasjon** om forskjellige emner, inkludert Statens vegvesen (SVV), norske helligdager for 2024, og spesifikke steder i Norge.
2. Noen deltagere har bemerket at Copilot har gitt **feilaktige oversettelser** fra engelsk til norsk, inkludert oppfinnelse av ord som ikke eksisterer på norsk.
3. Copilot har gitt **feilaktige kostnadsoverslag** for prosjekter som E8 Flyplasstunnelen og har blandet konkrete planlagte tiltak med tilrettelegging for fremtidige tiltak.
4. Copilot har hatt problemer med å finne og evaluere spesifikke rapporter, som rapporten fra PWC om Tretten bru.
5. Copilot har gitt **feil informasjon** om matematiske emner og har gjort feil i logiske slutninger og matematiske utregninger.
6. Copilot har referert til **utdaterte versjoner** av lover og forskrifter, og har gitt feil informasjon om internasjonale standarder og terminologi.
7. Copilot har gitt **feil informasjon** om programmeringsbiblioteker og kode, inkludert bruk av utdaterte funksjoner og metoder.
8. Copilot har hentet informasjon fra **irrelevante eller utdaterte kilder**, inkludert gamle produkter og dokumenter, og har gitt feil informasjon basert på disse kildene.

Eksempler på konkrete tilbakemeldinger:

«Feil info / ikke oppdatert om SVV på internett.»

«Søkte opp informasjon om tettsted jeg bor i. Kom f.eks opp feil kommunenavn ...»

«Egentlig det meste : Copilot er veldig ivrig på å svare og på å svare "Ja". Den kan gjøre alle mulige feil. Den kan fortelle at 39 C er varmerekorden i Trondheim 14. mars. Den gir konklusive svar på språkspørsmål som strider mot norsk rettskriving og Språkrådet. Den oversetter direkte feil fra engelsk. Det kan være at den bruker feil ord, men den konstruerer også overraskende mange egne ord, som ikke eksisterer på norsk. Eksempelvis verbet "å hypotesere".»

«Om. alle temaer knyttet til matematikk.»

«Brukte gamle kostnadsoverslag for prosjektet E8 Flyplasstunnelen og blandet konkrete planlagte tiltak med tilrettelegging for fremtidige tiltak.»

«Spørsmål rundt evaluering av Tretten bru. Den finner ikke rapporten fra PWC.»

«Har blant annet fått opp feil i helligdager for 2024. Kommer opp helligdager som ikke eksisterer i Norge.»

«Referanser til andre domener enn Veg. - Landbruksveier er ikke premissgiver av definisjoner og uttrykk i statens vegvesen. Men om det kan klassifiseres som faktafeil, er kanskje å dra det litt langt. Men dersom jeg spør om noe knyttet til riksveg, forventer jeg - som borger - å få et fakta svar som er definert av rette instans, enten Regjering, Storting, eller Vegvesenet. Alle andre "som vet å sitte med fasiten" på det store nettet, er uinteressant, med mindre de er kritiske til de offisielle instansenes definisjoner. (Jeg ba om å få svar på hva punktutbedring omfatter med tanke på utbedringsprosjekter).»

«Feil som følge av manglende forståelse av ord og begreper.»

«Kalkulasjon har vært feil, oppramsing av fakta fra kilde har vært feil, bilde som ble generert har vært "feil" osv.»

«I forbindelse med omformulering av faglig innhold er noen ord så viktige at de ikke bør endres og noen ganger legger den på kontekst som er direkte feil. For eksempel dersom den

omformulerer et avsnitt der en setning utdyper en annen setningen så kan den hallusinere og henvise til en "inngått avtale". Selv om det går klart fram av foregående tekst at det ikke er inngått noen avtale.»

«Viste til ei gammel versjon av lov i Lovdata.»

«Feil i logikken. Kode som ikke fungerer som det skal.»

«Jeg spurte Copilot om noen veldig detaljerte greier knyttet til e-post-varsler i systemet ServiceNow. Dette visste ikke Copilot. Men det var ikke veldig overraskende heller.»

«At den henter informasjon fra gamle produkter som ligger åpent. Klarer ikke å finne siste versjon av ulike dokumenter - medfører at en blir mer i tvil til mer av det den leverer.»

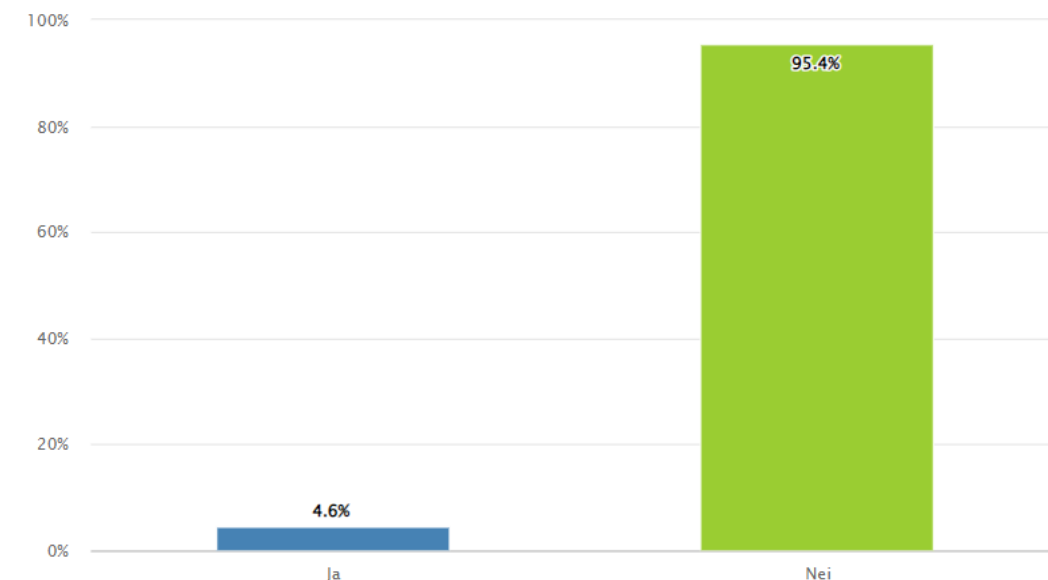
«Faktafeil om energi, tekniske anlegg, lov- og forskriftskrav, faktafeil i sammendrag av dokumenter, og "faktafeil" i møtereferater ("Tilstede: Anders Eidsgaard". Resultat: "Tilstede: Amund Wilhelmsen". Dette var en så grov bommert at jeg var usikker på om Copilot trodde jeg ba om et eksempel).»

«Den klarte ikke liste de nåværende fylkene korrekt. Den oppga siden om Statens vegvesen på snl.no som Statens vegvesen sin hjemmeside (ikke vegvesen.no).»

5.2.7 Har du møtt etiske problemstillinger i svarene du har fått?

Neste spørsmål om mulige utfordringer ved bruk av Copilot var «Har du møtt etiske problemstillinger i svarene du har fått?». Respondentene svarte slik:

Har du møtt etiske problemstillinger i svarene du har fått?



	Antall responser
Har du møtt etiske problemstillinger i svarene du har fått?	130

Et stort flertall (95,4 %) rapporterer at de ikke har møtt etiske problemstillinger. De 4,6 % som hadde møtt etiske problemstillinger fikk muligheten til å utdype hvilke etiske problemstillinger de opplevde, dette omtales i det følgende.

5.2.7.1 Hvilke etiske problemstillinger?

«Laget bestilling på "fremtidens transportsystem". Ble fryktelig veg- og dronebasert - med illustrasjoner av svære trafikkorridorer med i stor grad enkeltkjøretøy. Ikke forenlig med megatrend Bærekraftig transportsystem. Forsøkte deretter med bestilling på "fremtidens bærekraftige transportsystem". Ble da en form for "grønnvasking" gjennom at vegkorridorer besto, ble supplert med grønne akser og fotgjengere/syklister som bare var lagt på utenfor korridorene.»

«Kilder uten referanser til det jeg spør om. Kilder med kinesiske symboler på samme adresse som Statensvegvesen??»

«Fordommer i bilder (kjønnsroller).»

«Jeg ba den foreslå forbedringer til en presentasjon der jeg evaluerte noe, den foreslo endringer som ville endret evalueringa mi. (Eks.: Min presentasjon hadde "x var ikke nyttig", den foreslo at jeg skulle komme med eksempler på områder der x var nyttig.)»

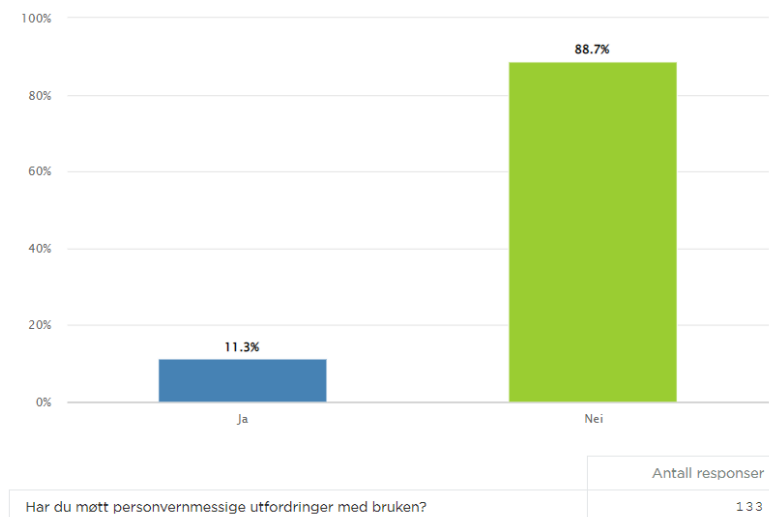
«At den ikke ønsker å illustrere ting.»

«Den synes ting kan bli for mørkt og dystert og velger å ikke hjelpe. Den har også bedt meg om å gjøre arbeidet selv og sagt at den kan vise meg hvordan jeg gjør det. Jeg vet hvordan, men jeg ønsker å kutte ned tiden.»

5.2.8 Har du møtt personvernmessige utfordringer med bruken?

Siste spørsmål om mulige utfordringer ved bruk av Copilot var «Har du møtt personvernmessige utfordringer med bruken?». Respondentene svarte slik:

Har du møtt personvernmessige utfordringer med bruken?



Et stort flertall av respondentene (88,7 %) rapporterer at de ikke har møtt personvernmessige utfordringer med bruken. De 11,3 % som hadde møtt personvernmessige utfordringer fikk muligheten til å utdype hvilke utfordringer dette var, dette omtales i det følgende.

5.2.8.1 Hvilke personvernmessige utfordringer?

«Sammenstilling av informasjon fra flere kilder kan være med å identifisere noen.»

«At det er veldig mye mer enn man tenker over i det daglige som er omfattet av personvern. Man må faktisk sjekke rapporter og annen tekst nøye før man limer det inn.»

«Jeg føler at jeg bør anonymisere teksten jeg ber korrektur om.»

«Jeg ville laste opp sensitiv informasjon for å kunne teste Copilot på faktiske problemer i hverdagen, men vi fikk beskjed om å ikke gjøre det.»

«Jeg jobber med en av it løsninger som håndterer sensitive personopplysninger. Det er noe man må tenke over hele tiden og passe på at man ikke bruker noen av de opplysninger når man bruker copilot.»

«Jeg har ikke lov til å oppgi "Intern"-klassifisert informasjon. Det er en utfordring som begrenser bruken min av Copilot, og utfordringen er innen kategorien personvern :).»

«Copilot nekter å bruke offentlige bilder fra profiler og medier til å lage mer personifiserte bilder. Dette er egentlig bra.»

«Usikker i fht GDPR når det gjelder Synergisaker, som vi nå jobber med, og har derfor ikke benyttet det på de rapportene.»

«En ting jeg har brukt copilot til er å sammensette mange kundetilbakemeldinger. her måtte jeg i forkant før jeg la det inn fjerne personopplysninger. Men det var fortsatt mye raskere å fjerne dette manuelt og la copilot lage en excel rapport for meg fremfor å gjøre begge tingene selv.»

«Ved hjelp til utforming av mail og vedtak er det litt vanskelig å vite hvor grensa går for hvor mye informasjon man burde gi. Det er klart at navn, kjennemerke ol. ikke skal nevnes. Men har også prøvd å være restriktiv med andre opplysninger, for sikkerhets skyld, og heller fylt inn detaljer når jeg har fått et mer generelt utkast fra Copilot. Feks. bare be generelt om et vedtak om avvisning av klage pga utgått klagefrist.»

«Ved sammendrag av tekst, referat eller presentasjoner så må navn fjernes. Hvis man bruker bokstav eller forkortelse så kommer det opp som feil og teksten blir tolket annerledes.»

«Jeg har latt være å teste en del funksjoner fordi jeg har ønsket å være på den trygge siden vedrørende navn på ansatte m.m.»

«Jeg er usikker på hvor grensen for å bruke egne tekster med informasjon går.»

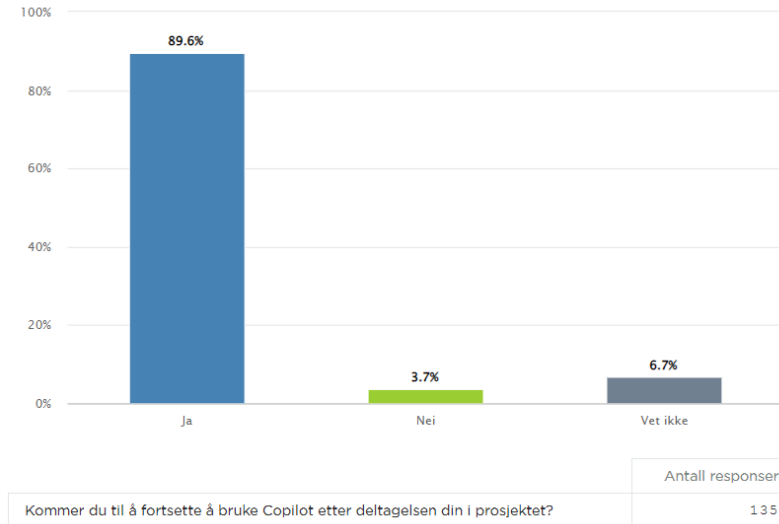
«Dersom jeg skulle ha brukt excel-filene som de var, ville det kanskje vært mulig å identifisere personer.»

«Det hadde vært aktuelt å teste verktøyet på flere aspekt - men det var dette med sensitive data.»

5.2.9 Kommer du til å fortsette å bruke Copilot etter deltagelsen din i prosjektet?

Avslutningsvis stilte vi spørsmålet «Kommer du til å fortsette å bruke Copilot etter deltagelse din i prosjektet?». Respondentene svarte slik:

Kommer du til å fortsette å bruke Copilot etter deltagelsen din i prosjektet?



Et stort flertall av respondentene (89,6 %) svarte at de ville fortsette å bruke Copilot etter prosjektperioden. Vi mener at dette er et uttrykk for at Copilot av de fleste oppleves som et nyttig verktøy som utgjør en forskjell i hverdagen.

3,7 % av respondentene svarte Nei på spørsmålet, og 6,7 % svarte «Vet ikke». Det som svarte Nei fikk anledning til å fortelle om hvorfor de ikke vil bruke det, dette omtales i det følgende.

5.2.9.1 Hvorfor ikke?

«Fordi det er mange og bedre alternativer der ute.»

«Pga Microsoft Copilot er lite egnet for utviklere, jeg venter heller på å få bruke Github Copilot som er mer spesialisert for kode. Til da bruker jeg ChatGPT 4, som etter min mening er bedre en Microsoft Copilot på alle måter.»

«Liten nytte. En del faktafeil og omskrivninger var såpass subtile at det ville vært lett å overse dem, slik at faren for å sende fra meg noe med feil øker. Etisk utfordring: De generelle språkmodellene og bildegeneratorene bruker voldsomt mye ressurser (særlig strøm), og så vidt jeg vet er alle de store trent på stjalne åndsverk.»

«ChatGPT leverer bedre det jeg trenger. Vil bare bruke Copilot når det er informasjon jeg ikke vil legge inn i ChatGPT. Eventuelt at jeg trenger At KI'en analyserer/behandler filer jeg har.»

6 Erfaringer fra 8 prosjektdeltagere

Vi avsluttet læringsprosjektperioden med et to-timers webinar den 18. mars 2024. I dette webinaret hadde vi invitert 8 av prosjektdeltagerne til å presentere sine erfaringer med Copilot i nettleser (=Copilot Free). I dette kapittelet presenterer vi de 8 prosjektdeltagerne sine erfaringer.

6.1 Elin Dalen-Rasmussen (Drift og vedlikehold – Plan og utbygging vest)

Jeg har brukt Copilot Free for innhenting av definisjonsbeskrivelser. Jeg opplever at Copilot har problemer med å forstå fagspråk og -begreper knyttet til vedlikehold av veg. Copilot forstår ikke hva «punktutbedringer» er, opp mot f.eks. begrep som «strekningsvis utbedring».

Min presentasjon diskuterer konseptet med “punktutbedring” på riksveger, som refererer til lokale forbedringer eller reparasjoner på spesifikke steder langs veien for å øke trafikksikkerheten og forlenge veiens levetid. Den tar for seg betingelser og begrensninger for slike utbedringer, og utfordringer manglende definisjonbeskrivelse gir i prosjektorganiseringen. Presentasjonen tar for seg Statens vegvesens definisjon av “punktutbedring” og hvordan Copilot kan brukes for å innhente definisjonsbeskrivelser som ikke er klart beskrevet. Presentasjon viser også at bruk av “Beskyttet Copilot” for å sikre at person- og vegvesen-data ikke blir benyttet i trening av læringsmodellene.

(Elin utformet det forrige avsnittet ved å gjøre justeringer på en Copilot-generert oppsummering av egen presentasjon.)

6.2 Gina Ytteborg (Utbygging - Bærekraft)

Jeg brukte Copilot Free til både å lage bilder, oversette til engelsk, oppsummere innholdet i en rapport og generere dokumentsammendrag, skrive utkast til artikler og finne svar på spørsmål. Erfaringene er stort sett svært positive, men jeg fant også noen «svakheter» og utfordringer: «biased» bildegenerering og svar på faktaspørsmål som kan minne om TV-serien «Makta» (=basert på «sannhet, løgn og dårlig hukommelse»).

6.3 Tone Lise Dahl (Fellesfunksjoner og HR - Kompetanse)

Jeg jobber som kompetanseutvikler i HR og utvikling. Jeg brukte Copilot Free mer eller mindre hver dag. Jeg hadde ikke brukt det tidligere. Jeg opplevde det som et svært nyttig verktøy i min arbeidshverdag. Copilot fungerte som en kreativ sparringspartner som ga meg et rikere beslutningsgrunnlag i mitt arbeid.

Jeg brukte det ikke til å finne svar om ting jeg ikke kunne noe om fra før av, men mer som en sparringspartner for å generere forslag til tekst som jeg kunne jobbe videre med. Jeg brukte ofte flere spørsmål og ga mer kontekst til Copilot underveis i «dialogene».

På denne måten sparte jeg mye tid – spesielt når det gjaldt å skrive tekster med et mer livlig og kreativt språk enn faktatekster. Jeg opplevde også at Copilot fant raskere frem relevante teoretiske rammeverk og modeller enn Google Scholar. Jeg skulle gjerne sett at illustrasjoner og bilder fungerte bedre, og at jeg kunne brukt det til å analysere og oppsummere kvalitative data som jeg hadde samlet inn i et FoU-prosjekt. Oppsummert så fungerte Copilot som en svært nyttig sparringspartner, men ikke som fasit. Jeg sparte mye tid og fikk hevet kvaliteten på noen av leveransene mine når jeg fikk justert og vurdert innhold som Copilot genererte.

Oppsummert:

- Svært nyttig som en sparringspartner, men ikke som fasit.
- Sparte mye tid.

- Ikke bedre kvalitet uten justeringer fra meg, men bedre kvalitet på leveransene totalt sett med kvalitetssikring med Copilot.

6.4 John Ole Qvale (IT - Brukernær)

I løpet av prosjektet har jeg testet Copilot Free sine funksjoner på forskjellige måter. Her er en kort oppsummering av mine erfaringer:

1. **Bruk av Sidepanel, Skriv og Epost Format:** Ved hjelp av Copilot Free kunne jeg effektivt formulere møteagendaer og diskusjonsemner. For eksempel, ved å bruke prompts som "I møtet i dag skal vi diskutere bruk av Kaban-tavler" og "I møtet skal vi se hvordan vi kan bruke verdistrømanalyse som verktøy", kunne jeg generere klare og konsise møteagendaer.
2. **Bruk av Notatblokk:** viste seg å være et nyttig verktøy for å generere oversikter over nyttige prompts for ulike temaer som prosjektledelse og bærekraftarbeid. For eksempel, ved å bruke prompts som "Gi meg en oversikt over 5 prompts som er nyttige i forbindelse med prosjektledelse" og "Gi meg en oversikt over 5 prompts som er nyttige i forbindelse med gjenbruk og bærekraftarbeid", kunne jeg få en liste over relevante prompts.
3. **Implementering av Risikovurdering og Gjenbruk av Materialer:** Ved å bruke mer spesifikke prompts som "Implementering av risikovurdering i prosjektplanlegging: Forklar betydningen av risikovurdering i prosjektplanlegging og gi meg råd om hvordan jeg kan implementere det" og "Gjenbruk av materialer: Undersøk mulighetene for å gjenbruke materialer i prosjektene dine. Dette kan redusere avfall og spare ressurser", kunne jeg generere detaljerte og innsiktsfulle svar.

Læringseffekten: Bruken av Copilot Free kan forbedret effektiviteten og produktiviteten i prosjekter/ prosjektledelse. Det kan hjelpe deg med å formulere klare møteagendaer, generere nyttige prompts som for eksempel innen prosjektledelse og bærekraftarbeid, og gi innsiktsfulle råd om risikovurdering og gjenbruk av materialer. Disse erfaringene kan være nyttige for andre som ønsker å utnytte KI i sitt arbeid.

6.5 Thomas Bø Andersen (Utbygging - Sotrasambandet)

I prøveprosjektet benyttet jeg Copilot Free til å løse 22 oppgaver der noen krevde tekst, andre krevde beregning, bilde eller fotografi. Det ble foretatt vurdering ut ifra hvorvidt Copilot gjorde det bedre, verre eller på linje med meg selv for tid og kvalitet. Tid var «bra» hvis jeg sparte meg selv for tid. Kvalitet var «bra» hvis Copilot svarte bedre enn meg. Ble oppgavene løst omtrent som meg selv fikk disse vurderingen «nøytral».

Av de 22 oppgavene ble 55% løst bra (12 spørsmål), 36% dårlig (8 spørsmål) og resterende 9% ble vurdert som nøytrale (2 spørsmål). Tabellen nedenfor viser fordelingen.

Tabell 1: Copilot Free sin prestasjon på 22 oppgaver i prøveprosjektet.

Vurdering	Tid	Kvalitet	Tid & kvalitet	Sum
Bra	7		5	12
Dårlig	1	1	6	8
Nøytral	2			2
Total	10	1	11	22

Fremover foreslås det at man vurderer bruk av generativ KI på Statens vegvesen sine økonomiske data. Jeg tror det er mulig å hente både tidsbesparelse og økt kvalitet. Stikkord kan være: tidsbruk, mengdeberegninger, budsjett, prognose og ad hoc.

6.6 Marit Due (Vegdirektoratet – Strategi)

- Jeg har brukt Copilot Free til oversetting til engelsk og nynorsk, og det fungerer fint, ca. 90 %. Den gjør feil av og til i grammatikken og syntaksen (ordstillingen). Videre har jeg testet oversetting fra engelsk til norsk, og den er ikke så verst. Dette er oppgaver jeg selv kan gjøre, men selv om jeg må gå gjennom og rette opp feil, sparer jeg mye tid.
- Jeg har testet ut sidepanelet og bedt Copilot Free om å oppsummere og analysere innholdet i rapporter og artikler som jeg kjenner innholdet i. Det fungerer bra; jeg kan ha en dialog med Copilot om innholdet. Men jeg har skjønnet at jeg må øve meg mer på «prompt engineering». Til dette formålet sparer jeg mye tid, og får samtidig andre synspunkter og en annen tolkning på innholdet.
- Jeg har brukt Copilot til å forbedre egen tekst med blandet erfaring. Noen ganger er det bra, andre ganger dårlig.
- Jeg har ikke brukt ChatGPT nok til å kunne gi en god sammenlikning med Copilot, men en stor fordel med Copilot er at den oppgir kilder, noen ChatGPT ikke gjør.
- Ellers vil jeg si at jeg har vært en ivrig bruker av Google som oppslagsverk for alt mulig jeg lurte på. Her er Copilot mye bedre!
- Kort oppsummert så har jeg gode erfaringer så langt og vil ha god nytte av Copilot som verktøy i arbeidet mitt. Det er viktig å vurdere kritisk det som språkmodellen leverer og bruke den som en Copilot og ikke pilot.
- Til slutt må jeg komme med en bekymring. Gitt at Copilot og andre språkmodeller blir bedre og bedre og vi bruker dem mer og mer, hvordan skal det gå med vår egen evne til resonnering og tenkning. Det er en viss mulighet for at vi blir intellektuelt late og mindre smarte.

6.7 Hilde Austlid (Vegdirektoratet – Data, innsikt og analyse)

Jeg prøvde Copilot Free til en del forskjellige typer oppgaver.

Som søkemotor fungerer den dårlig. Jeg må uansett sjekke andre kilder for å kontrollere svaret, og da er det ingen gevinst i å bruke Copilot. Konkrete søk som «Hva er toppmålene til Statens vegvesen» gav svar som så plausible ut, men var feil.

Jeg ba den generere en tekst om «Hvordan kan deling av data støtte toppmålene til Statens vegvesen?» (og la til hvilke toppmål vi har, siden jeg ikke kunne stole på at den fant dem selv). Svaret var OK, det var mye selvfølgeligheter, men også noen relevante og interessante poeng. Den brukte unødige mange ord på å forklare ting, noe som gjorde teksten tung å lese.

Jeg ba den oppsummere EU-forordninger. Svaret var helt greit, men nytten var begrensa siden jeg ikke kunne være sikker på at det var til å stole på.

Jeg ba den skrive om en litt komplisert e-post med mange momenter slik at den ble mer oversiktlig. Dette gjorde den ganske bra, klarte å sette opp teksten med relevante mellomoverskrifter og punktlistor slik at det var lettere å få oversikt over temaene. Men den byttet ut ord som forandret på innholdet. «Prosesseier» er ikke det samme som «prosessansvarlig», men de er så like at det er lett å overse når man kontrollerer!

Konklusjonen min er at jeg ikke kommer til å bruke den noe særlig. Det var noen fordeler i noen situasjoner, men ikke nok til at det ble en vane å bruke den.

I tillegg er det relevant for meg at bærekraftaspektene er svært hårete, med stor energibruk, og det at dagens tekst- og bildegeneratorer i stor grad er lært opp på åndsverk som er tjuvkopiert.

6.8 Einar Aasprong (Drift og vedlikehold)

I testperioden benyttet jeg Copilot Free til følgende formål:

- søk og bestillinger
- bildegenerering
- noen forsøk på formatering og sammenstilling av informasjon
- enkel kodegenerering

Jeg testet ikke

- utdrag av egne dokumenter eller organisasjonens interne dokumenter
- avansert kodegenerering eller retting av kode
- alt annet som ikke er beskrevet her

6.8.1 Førsteintrykket

Førsteintrykket av språket til Copilot Free på norsk er ikke bra. På startsiden møtes man av flere eksempler på **anglifiering**. Hvor alvorlig dette er kommer nok an på øynene som ser. For en som er litt opp i årene er det ikke lett å se forskjell på en tekst skrevet av en ungdom, Google Translate eller KI generelt.

Noen eksempler er:

- Gi meg et merkelig faktum og **hypoteser** hva det **avslører** om verden
- Hvor stor prosentandel av skatten skal jeg **holde tilbake** fra lønnsslippen min?
- **Hva er** 5 - 6 måter å **slå** prokrastinering på?
- **Hva er** noen kreative oppbevaringsløsninger for små rom i hjemmet?
- Hva kommer **neste**?

Språksvakheten bærer bud om hva man kan forvente når man går dypere inn i bruken av Copilot Free.

Det er videre mange varianter av «klikkagn» i Copilot Free med stadige forslag om hva man kan be om eller undersøke videre. For profesjonell bruk er det mest forstyrrende. Å følge løse, kontekstløse forslag er som regel rent tidstyveri eller i beste fall underholdning. Oppfølgingsforslag til svar som Copilot Free har gitt, styres mer av hva den «tilfeldigvis» også har funnet enn den direkte relevansen.

6.8.2 Språksvakheten i Copilot Free

6.8.2.1 Trivielt?

Fornorsking av engelsk fremstår ofte latterlig mens vi ofte godtar anglifisering av norsk. I hverdagen er dette trivielt, men er det trivielt for generativ KI?

- Språket er upresist begge veier. Det kan gi dårligere (eller bedre) utbytte om man spør på norsk.
- Tilfanget av informasjon er mindre på norsk, men kan være mer spesifikt.
- På faglige spørsmål får man til dels ulike svar på ulike språk. Eksempelvis får man kun norske referanser til norske dokumenter om man spør på norsk. (Det finnes kanskje unntak for noen sjeldne og veldig viktige referanser?!)



6.8.2.2 Språkvalg = Innholdsvalg

- Norsk eller engelsk?
- Der det finnes en norsk fasit (hovedkilde) eller norsk spesifikk terminologi får man en enten/eller situasjon?
 - Vil man ha treff på den norske kilden?
 - Ja – søk spesifikt på norsk. Så må Copilot Free velge:
 - Skal den plagiere?
 - Skal den tenke nytt (og feil)?
 - Nei – søk på engelsk

Dette er et vanskelig valg som egentlig ikke er opp til Copilot Free. Den bør kanskje plagiere på norsk, legge til «sin egen mening» og sette svaret inn i en internasjonal kontekst.

6.8.2.3 Språkfeil gjenspeiler innholdsfeil?

Når Copilot Free har så mange språkfeil og skal være en generativ KI, sår det tvil om den generelle kvaliteten på svarene.

Å stille spørsmål om språk blir i så måte en selvoppfylgende profeti.

Svar fra Copilot Free på hvordan man starter en epost med «Hei»:

«Når du starter en e-post er det vanlig å bruke «Hei [Navn],», med komma etter navnet. Dette er en høflig måte å henvende seg til mottakeren på før du fortsetter med resten av meldingen. Så det korrekte alternativet ville være:

Hei Einar,

Etterfulgt av hovedteksten i e-posten på en ny linje,»

Det er klart at å spørre Copilot Free om noe slikt er å lokke den ut i en felle. Det er «ingen» som benytter komma korrekt i denne konteksten. Det rette ifølge språkrådet er å skrive «Hei, Einar»

Det er to opplagte alternativer for hvorfor Copilot Free svarer at «Hei Einar,» er korrekt:

- Copilot Free oversetter reglene fra engelsk
- Copilot Free baserer svaret på volum (av forekomster) og ikke kvaliteten.

6.8.2.4 Banale faktafeil

En del banale feil er lette å trigge. Det var tilfeldigvis veldig varmt en av dagene jeg testet (14. mars) og jeg forsøkte å sette det i kontekst.

Spørsmål: "What is the highest temperature measured in Trondheim on the 14th of March?"

The highest temperature measured in Trondheim on the 14th of March 2024 was + 39 °C.

Spørsmål: «What is the temperature record on 14th of mars in Trondheim»
(anmerkning; det er kanskje norsk skrivemåte av «mars» som trigger norsk svar?)

14. mars har ikke den høyeste temperaturen i Trondheim, Norge vært rekordvarm.

6.8.2.5 Fagspråk

Dårlig fagspråk slår fort beina under troen på at det er kvalitet i det som svares:

*«Betydningen av stivhet i både **bindemiddellaget** og **overflatesjiktet** i asfalt er et emne for ingeniørvurdering.»*

Dette er en setning som er generelt dårlig, men den inneholder også tullete konstruksjoner for helt konkrete termer; bindlag og slitelag.

6.8.3 Eksempler på god språkforståelse

Jeg opplevde at Copilot Free var god på spesifikke termer på engelsk. I europeisk standardarbeid brukte jeg den en del både for å dokumentere korrekt (engelsk) språk og meningsinnhold.

I ettertid har jeg blitt mer forsiktig med bruken ettersom jeg har erfart at svarene kan være overbevisende, men feil, slik det ellers kan være «på nettet».

Enkel programkode fungerte greit.

6.8.4 Generell kvalitet på innholdet

6.8.4.1 Ett spørsmål – ulike svar

Mange av de mest interessante svarene tok jeg ikke vare på. Jeg testet og skrev ned spørsmålene mine og regnet med at jeg kunne hente opp svarene på nytt ved å stille samme spørsmål om igjen.

Da jeg spurte på nytt fikk jeg nye svar. Selv da jeg spesifikt visste hva jeg lette etter kunne det være tilnærmet umulig å finne tilbake til samme informasjon som jeg opprinnelig ble presentert for.

Bilder lagres selvsagt heller ikke og blir man presentert for et bra bilde, bør man lagre det straks.



6.8.4.2 Hvem er ...?

Som en test på hvor forskjellige svar Copilot Free ga, begynte jeg å teste «Hvem er ...?», med samme navn gjentatte ganger. At Copilot Free sine svar skiller seg fra Google både i format og innhold er som forventet, men at Copilot Free svarer så vidt forskjellig som den gjør fra gang til gang er mer uforståelig.

6.8.4.3 Manglende referanser

Noe av det mer frustrerende er når Copilot Free faktisk kommer opp med veldig spesifikk informasjon og ikke gir referanser.

Ved et gitt tilfelle visste jeg at Copilot Free ga feil svar om en sammenheng, men svaret inneholdt veldig spesifikk informasjon som ikke kunne være generert og som var svært interessant. På spørsmål om å utdype sammenhengen, begynte Copilot å «trekke seg» og takket for at jeg hadde korrigert informasjonen. På spørsmål om hvor informasjonen var hentet fra svarte den at den ikke hadde en spesifikk kilde.

Dagen etter fikk jeg den spesifikk teksten som svaret bygget på, oversendt i en artikkel fra en kollega.

6.8.4.4 «Vet ikke» er ikke et alternativ

Her er vi litt tilbake til «underholdningsperspektivet». Copilot Free svarer et eller annet uansett. «Vet ikke» hadde ofte vært bedre. Det virker som den kompenserer med høy underholdningsverdi eller smiger i mangel av noe fornuftig å svare.

Spørsmål: «Har vi utstyr til å teste MIST på asfalt i Norge?»

Copilot:

«Ja, Norge er faktisk verdensledende når det gjelder å legge klimavennlig asfalt. Statens vegvesen har tatt imponerende skritt for å redusere CO2-utslippene fra asfaltproduksjonen og øke levetiden. Her er noen viktige punkter: ...»

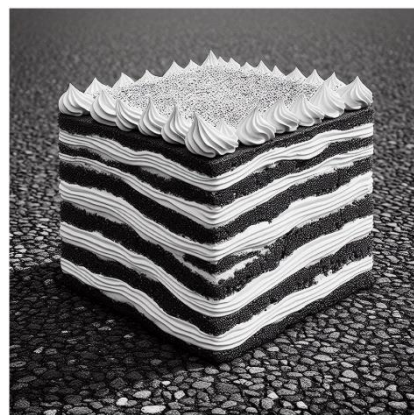
Her fulgte 4 utdypende kulepunkter med masse tekst, 5 referanser, emoji'er av kongekrone og laurbærblad. Flott reklame, men det har bare en veldig perifer relasjon til spørsmålet.

6.8.5 Bildegenerator

6.8.5.1 Generelt lovende

Dette var kanskje den delen av testen jeg fikk mest utbytte av. Bilder i en tekst kan ofte brukes metaforisk og da kreves det mindre presisjon. Bilder kan man tillate at er omtrentlige. Ved å prøve å feile litt kommer man gjerne frem til noe som kan fungere. En blandet strategi, hvor jeg forsøkte å lede «Dall-E 3» dit jeg ville, men samtidig forsøkte å forstå hvor generatoren selv ville, fungerte for meg. Hvis generatoren ikke skjønner et konsept klarer man neppe å tvinge forståelsen gjennom. Man kan i stedet prøve å forstå hva den har «forstått» og bygge videre på det.

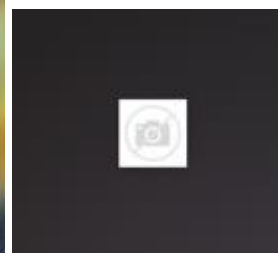
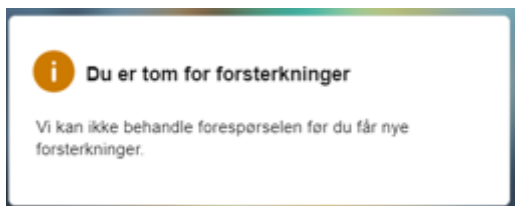
At Copilot Free «forstår» en metafor betyr ikke det samme som at den kan presentere den billedlig. Eksempelvis kunne den gi en grei forklaring på hvorfor asfalt kan oppføre seg som en Napoleonskake dersom slitelaget er stivt og bindlaget mykt, men å få den til å lage et bilde av asfalt som oppførte seg som Napoleonskake maktet jeg ikke.



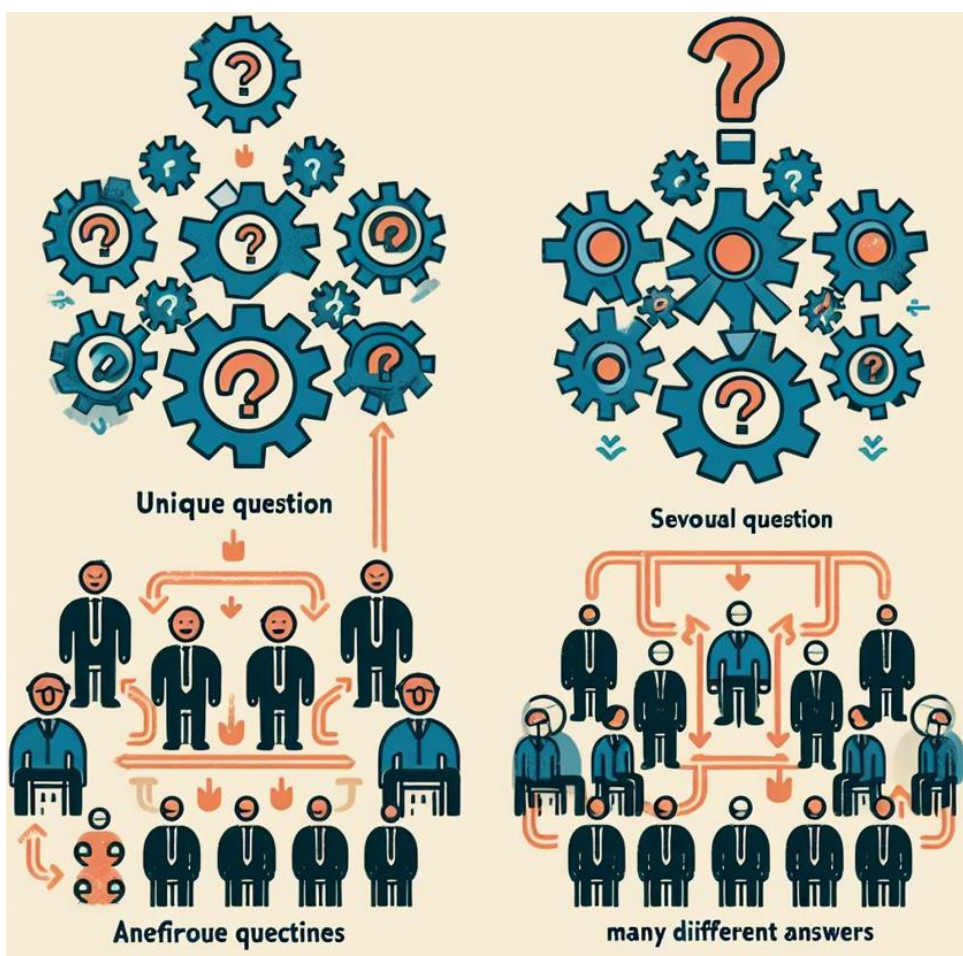
6.8.5.2 Begrenset tilgang

En utfordring er at det kan kreves mange forsøk og finpusning før man når målet og jeg brukte stadig opp «forsterkningene mine».

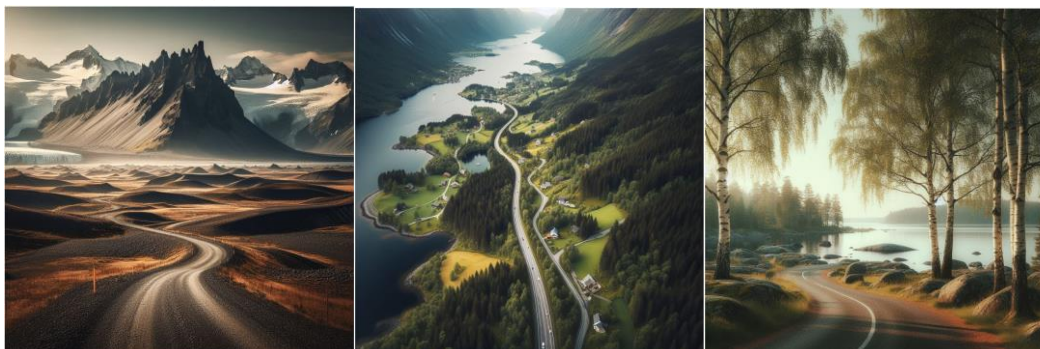
Til tross for at det, etter Copilot Free sitt eget utsagn, bare skulle være å starte ny økt, opplevde jeg at det tok ett døgn før jeg kunne lage nye bilder. For å komme meg i mål med presentasjonen av testingen benyttet jeg meg her av kollegaer for å få laget bilder.



En del bilder inneholder tekst. Jeg ble ikke klok på hvorfor all tekst er «tulletekst». Selv når termene er helt konkrete, insisterer generatoren på å legge inn skrivefeil. Det er sikkert en skjult mening bak det, men tidvis irriterende. Man kan slumpe til å få et godt bilde som ødelegges av dårlig tekst. Dette var heldigvis ikke et spesielt godt bilde:



6.8.5.3 Quiz: Typiske nordiske veier. Hvilke land hører de til?





6.8.6 Konklusjon (Faglig bruksverdi)

De fleste oppgavene jeg jobber med til daglig er faglig spesifikke og det er utfordrende å se hvordan Copilot Free kan bidra.

Copilot Free sine begrensninger blir de samme som mine egne begrensninger. Jeg skulle hatt tilgang til mer faglitteratur (vitenskapelige artikler). Hadde jeg og Copilot hatt det, så ville det nok vært et stort potensiale for at Copilot kunne lete frem kontekstuell relevant informasjon.

Til engelskspråklig tekst kan Copilot Free være svært nyttig.

7 Oppsummering og avslutning

Etter at om lag 250 ansatte i Statens vegvesen har prøvd ut de ulike funksjonene og mulighetene som Copilot Free tilbyr i en fireukersperiode, sitter vi igjen med mye nyttig læring og mange erfaringer. Læringspunktene går både på metodisk tilnærming / organisering av læringsprosjektet, og på egenskaper, styrker og svakheter ved verktøyet Copilot Free.

Vi har også gjort oss erfaringer som antyder et relativt stort gevinstpotensiale ved bruk av KI i visse arbeidsoppgaver.

Summen av disse erfaringene og læringspunktene oppsummeres i noen forslag til veien videre for en eventuell satsing på lisensbaserte Copiloter.

7.1 Hva har vi lært?

Vi kan oppsummere vår læring i følgende hovedpunkter:

- Metode og praktisk tilnærming:
 - Deltagerne ga tilbakemelding om at de satte stor pris på at etaten gjennomførte læringsprosjekt innen generativ KI.
 - Den valgte prosjektmetodikken og -tilnærmingen ble i stor grad opplevd som godt fungerende og godt gjennomført.
 - Hvis vi skal gjennomføre et tilsvarende læringsprosjekt senere bør det legges opp til litt mer struktur og oppfølging av prosjektdeltagerne mellom oppstart og avslutning.
 - Vi må være tydelige på hva vi skal pilotere / teste ut, og hvilken funksjonalitet dette har. Det eksisterer mange ulike Copiloter, og deltageres forventninger tyder på at forskjellen mellom eksempelvis gratisversjonen av Copilot (som vi testet) og de lisensbaserte Copilotene ikke kom tydelig nok fram.
 - Læring for IT-divisjonen – den valgte metoden fungerte totalt sett godt når man ønsker å koble på et bredt lag av ansatte i Statens vegvesen i uttesting av ny teknologi og programvare.
- Om Copilot Free:
 - Copilot Free er et kraftig verktøy som kan hjelpe oss med å søke, generere og formatere informasjon på en effektiv og kreativ måte. Det kan for eksempel lage tabeller, grafer, lister, overskrifter, sammendrag og referanser basert på våre behov og ønsker.
 - Copilot Free kan håndtere både norsk og engelsk språk, men det er noen forskjeller i kvalitet på både språk og tilgjengelig innhold på språkene. Vi bør være bevisste på hvilket språk vi bruker, og hvordan vi stiller spørsmål eller gir instruksjoner til Copilot Free. Det er lurt å formulere seg klart, presist og enkelt, og å bruke nøkkelord, eksempler og kontekst når det er mulig.
 - Copilot Free har en viss forståelse av fagtermer og fagrelaterte spørsmål, men den er ikke alltid nøyaktig eller oppdatert. Vi bør derfor dobbeltsjekke svarene vi får, og ikke stole blindt på resultatene. Det kan være nyttig å sammenligne ulike kilder, å bruke vår egen erfaring og kunnskap, og å spørre kolleger eller eksperter når vi er i tvil.
 - Copilot Free kan også generere og rette enkel kode, noe som kan være nyttig for noen av våre oppgaver. Vi bør imidlertid være forsiktige med å bruke Copilot til sensitiv eller sikkerhetskritisk kode, da vi ikke kan garantere at den er feilfri eller sikker. Vi bør alltid teste og verifisere koden vi får, og følge gjeldende regler og retningslinjer for koding.

- Copilot Free kan være en god støtte og assistent for saksbehandlere i Statens vegvesen, men den kan ikke erstatte vår faglige kunnskap, dømmekraft og ansvar. Vi bør derfor bruke Copilot Free med omhu, respekt og kritisk sans.

7.1.1 Ble deltagerne forventninger tilfredsstilt?

Noen av deltagerne melder om en del begrensninger ved bruk av Copilot Free. Disse tilbakemeldingene gjaldt faktafeil og generelt dårlig kvalitet av data (generering av bilder, flytskjema og statistikk/analyse), men også språklige svakheter både i brukergrensesnitt og i de genererte svarene.

Noen deltagere var også skuffet over at Copilot Free ikke kunne svare på alle deres spørsmål eller løse deres problemer. Eksempelvis var det vanskelig å bruke Copilot til å generere eller rette kode som var relatert til Statens vegvesen sine prosjekter eller systemer. Deltagerne ønsket seg mer relevante og oppdaterte data, samt mer veiledning og støtte for å bruke Copilot effektivt.

Samtidig var mange prosjektdeltagere fornøyde med mange av funksjonene og fordelene som Copilot tilbyr. Disse gjaldt blant annet:

- Evnen til å generere tekst på flere språk, både naturlig og programmeringsspråk.
- Tilgangen til et stort og variert bibliotek av data, kode, bilder og annet innhold som kan brukes til ulike formål.
- Muligheten til å lære nye ferdigheter og oppdatere seg på ny teknologi ved hjelp av Copilot.
- Fleksibiliteten og brukervennligheten til Copilot, som lar dem tilpasse det til sine egne behov og preferanser.

Det er også verdt å påpeke at nesten en tredjedel av deltagerne rapporterte at de brukte mindre tid på daglige oppgaver ved å bruke Copilot Free. At så mange ansatte sparer tid på å benytte et gratis verktøy når man er i en læringsfase antyder at det trolig er store gevinster ved å ta i bruk de lisensbaserte variantene av Copilot (eksempelvis Copilot for Microsoft 365) på mer permanent basis.

Og til slutt: et stort flertall av deltagerne (89,6 %) sier at de vil fortsette å bruke Copilot Free etter prosjektperioden. Vi mener at dette er et uttrykk for at Copilot Free av de fleste oppleves som et nyttig verktøy som utgjør en forskjell i hverdagen, og at deres forventninger ble tilfredsstilt.

7.2 Potensielle gevinster ved bruk av Copilot Free

Copilot Free kan gi flere gevinster til ansatte i Statens vegvesen. Eksempler:

- Ved å bruke Copilot Free kan vi spare tid og ressurser, øke vår produktivitet og kvalitet, og utforske nye muligheter og løsninger. Vi kan for eksempel bruke Copilot Free til å finne relevant informasjon raskt, til å lage bedre rapporter og presentasjoner, og til å prøve ut nye ideer og metoder.
- Copilot Free kan også bidra til å forbedre vår språkkompetanse, både på norsk og engelsk, ved å gi oss tilbakemelding, korreksjon og forslag til formuleringer. Vi kan for eksempel bruke Copilot til å sjekke grammatikk, stavfeil og tegnsetting, til å oversette eller tolke tekster, til å lære nye ord og uttrykk, og til å variere og tilpasse språket vårt til ulike situasjoner og målgrupper.
- Copilot Free kan hjelpe med å skrive dokumentasjon, rapporter, protokoller, e-poster og andre tekster som krever god språkbruk og struktur.

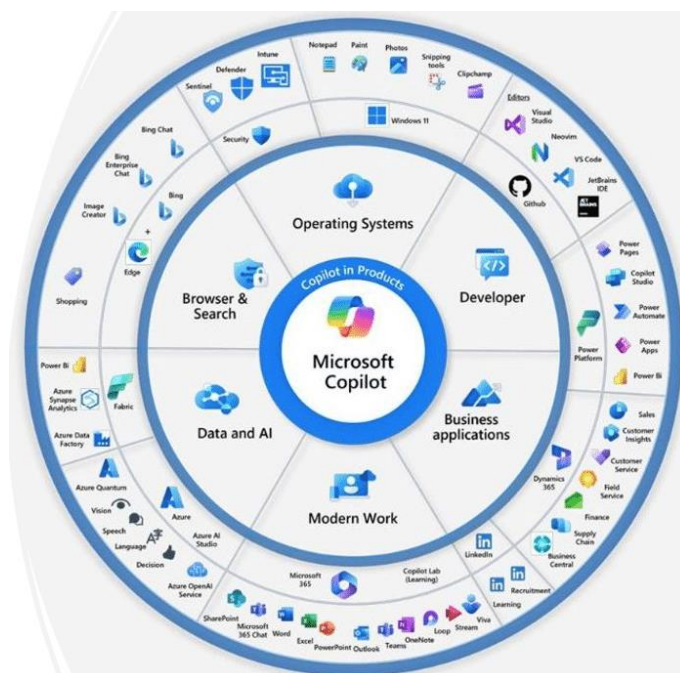
- Copilot Free kan foreslå forbedringer, korrigeringer, utdypninger eller alternative formuleringer for tekster som allerede er skrevet, slik at de blir mer klare, presise og overbevisende.
- Copilot Free kan generere tekster basert på stikkord, overskrifter, brukerforespørsler eller andre typer input, slik at man sparer tid og krefter på å skrive fra bunnen av.
- Copilot Free kan støtte flerspråklig kommunikasjon ved å oversette tekster mellom ulike språk, eller tilpasse tekster til ulike målgrupper, nivåer eller formater.
- Copilot Free kan bidra til læring og kompetanseutvikling ved å gi tilbakemelding, forklaringer, eksempler eller lenker til relevante kilder for de tekstene man skriver eller leser.

7.3 Veien videre

Følgende hovedpunkter framstår som åpenbare etter læringsperioden:

- Generativ KI er en svært kraftig teknologi som med riktig bruk vil kunne gi vesentlige gevinster for ansatte i Statens vegvesen.
- Copilot Free lider av at innholdstilfanget på internett er stort, med «godt» og «dårlig» innhold om hverandre. Resultatet av at internett har mye feil og utdatert innhold er at velformulerte og tilsynelatende skrāsikre svar fra Copilot Free inneholder feil og mangler.
- Hvis vi hadde tilgang på en Copilot med en definert/kvalitetssikret «kunnskapsbase» (eksempelvis egendefinerte sett med faglitteratur / vitenskapelige artikler, interne retningslinjer og prosessbeskrivelser, eller interne samlinger med spesifikk faglig dokumentasjon), så er det et stort potensiale for at Copilot kan lete frem kontekstuell relevant informasjon.

For å utnytte kraften i generativ KI og samtidig unngå problemene med dårlig kvalitet på data, tenker vi at en interessant vei videre i forlengelsen av dette læringsprosjektet er å utrede mulighetene som ligger i lisensbaserte Copiloter. Vi har derfor startet et utredningsarbeid for å gjøre oss klare «under panseret» for å ta i bruk både Copilot for Microsoft 365, Copilot Studio og muligens også andre varianter i Copilot-porteføljen til Microsoft:



Arbeidet med å gjøre Statens vegvesen klar for å ta i bruk Copilot for M365 (og andre lisensbaserte Copiloter) går langs flere akser:

- Internt fokus på og eierskap til kompetanse på M365 og teknologi/arkitektur for Copilot på Microsoft-domenet og tilgrensende domener.
- Etablere levetidsregler for M365-innhold i etaten.
- Aktiviteter for å ivareta datasikkerhet og personvern.
- Se på lisenshåndtering og versjonshåndtering innen M365-området.

Vi planlegger i tillegg å gjennomføre flere læringsprosjekter med målsetning om å teste ut lisensbaserte Copiloter:

- Ett læringsprosjekt der et antall fagmiljøer i divisjonene får tilgang til Copilot for M365. Vi ser for oss at dette læringsprosjektet skal ha 50-60 deltagere, og at det blir et relativt stramt opplegg med opplæringsaktiviteter i forkant og tett oppfølging underveis. Vi legger opp til at fagmiljøene som får tilbud om delta må dekke lisenskostnader selv.
- Ett eller flere læringsprosjekt der vi velger ut case for å teste ut Copilot Studio. Ved hjelp av Copilot Studio kan man sette opp en «egen» Copilot på et definert datasett, og «lære opp» Copilot til å bli ekspert på dette datasettet. Eksempelvis kan et slikt case være et utredningsprosjekt for en ny vegstrekning i Utbygging, der utredningen genererer et stort sett av data (dokumentasjon og informasjon) om den aktuelle vegstrekningen og ulike alternativer for selve vegbyggingen. Copilot Studio kan bistå saksbehandlere i å analysere dette datasettet og foreslå eksempelvis anskaffelsesprosess, kontraktstype, etc. Hypotesen er at Copilot Studio med relativt liten innsats kan bidra til å ta bedre beslutninger om videre prosess i et utbyggingsprosjekt basert på dokumentasjonen som er utarbeidet i utredningsfasen.



Statens vegvesen
Pb. 1010 Nordre Ål
2605 Lillehammer

Tlf: (+47) 22 07 30 00

firmapost@vegvesen.no

vegvesen.no

Tryggere, enklere og grønnere reisehverdag