

# Avtale om bruk av opphavsrettslig beskyttet materiale i trening av norske språkmodeller

mellom

## Nasjonalbiblioteket

på den ene siden,

og

## Kopinor

på vegne av

på vegne av medlemmene i Den norske Forfatterforening, Den norske Forleggerforening, Dramatikerforbundet, Fagpressen, Forbundet Frie Fotografer, Forfatterforbundet, Grafill, Mediebedriftenes Landsforening, Musikkforleggerne, NOPA – Norsk forening for komponister og tekstforfattere, Norges Fotografforbund, Norsk faglitterær forfatter- og oversetterforening, Norsk Journalistlag, Norsk Komponistforening, Norsk kritikerlag, Norsk Oversetterforening, Norsk Redaktørforening, Norsk Revyforfatterforening, Norsk tidsskriftforening, Norske Barne- og Ungdomsbokforfattere, Norske Billedkunstnere, Norske Kunsthåndverkere og Ny Musikk's Komponistgruppe

på den andre siden.

## Bakgrunn

I møte med utviklingen av generativ kunstig intelligens dominert av globale tekngiganter, i stor grad basert på ulovlig bruk av åndsverk, har norske myndigheter ønsket å utvikle en nasjonal infrastruktur for kunstig intelligens, forankret i norsk lov.

Det er bred politisk enighet om at et sentralt ledd i en slik infrastruktur vil være å utvikle norske språkmodeller basert på norsk og samisk språk, historie og kultur.

Språkmodellene skal speile norsk erfaring og kunnskapsproduksjon, samtidig som de skal ivareta norske rettighetshaveres økonomiske og ideelle interesser. Slike modeller vil kunne ha betydelig samfunnsverdi som basis for offentlig og privat tjenesteutvikling.

Gjennom sitt forskningsprosjekt Mímir har Nasjonalbiblioteket fastslått den store verdien ved å benytte rettighetsbeskyttet materiale som treningsdata for maskinlæring. På denne bakgrunn har Nasjonalbiblioteket og Kopinor blitt enige om vilkårene for å benytte Nasjonalbibliotekets samlinger av beskyttede publikasjoner til slik maskinlæring.

## 1. Formål

Formålet med avtalen er å tillate at det etableres et sett av treningsdata bestående av opphavsrettslig beskyttet innhold og at dette benyttes av Nasjonalbiblioteket til å trene, utvikle, vedlikeholde og dele norske språkmodeller med offentligheten. På den måten ønsker det offentlige å sikre et åpent tilgjengelig grunnlag for offentlig og privat tjenesteutvikling basert på språkmodeller som ivaretar norsk språk og kultur. Avtalen skal samtidig sikre at rettighetshavernes ideelle og økonomiske interesser ivaretas og ikke forringes, samt at rettighetshaverne sikres kompensasjon for bruk av deres materiale til trening.

## 2. Avtalens område

Avtalen gjelder eksemplarfremsending og bruk av tekst til Nasjonalbibliotekets trening av språkmodeller fra opphavsrettslig beskyttede pressepublikasjoner utgitt i Norge opp til ett år fra publiseringstidspunktet. Unntatt fra tillatelsen er:

- Publikasjoner og enkelte artikler som medlemmene i Mediebedriftenes Landsforening har vurdert at det er grunnlag for å unnta i henhold til Avtalens § 16.
- Illustrasjoner, foto, lyd, levende bilder og annet visuelt materiale
- Tekst fra fagpressepublikasjoner, magasin- og ukepresse

Avtalen gir også Nasjonalbiblioteket, og i særskilte tilfeller andre, adgang til å tilgjengeliggjøre modeller som er trent for bruk i tråd med denne avtalen underlagt vedlagte lisensbetingelser.

Unntatt fra tillatelsen er trening og tilgjengeliggjøring av spesialiserte språkmodeller som kan levere media- og nyhetstjenester.

### 3. Nærmere om de relevante rettighetshavere

Tillatelsen gjelder norske publikasjoner hvis rettighetshavere Kopinor representerer.

Forutsatt at Kopinor godkjennes av Patentstyret som representativ organisasjon på avtalens område etter bestemmelsen i åndsverkloven om generell avtalelisens (§ 63 annet ledd), vil Nasjonalbiblioteket også ha tillatelse til bruk av utgitt, beskyttet materiale skapt av rettighetshavere som Kopinor ikke representerer.

### 4. Definisjoner

Med grunnmodeller menes i denne avtalen generelle språkmodeller for generativ kunstig intelligens som ikke er posttrent med et særskilt tjenesteformål for øye.

Med trening menes prosessen med å lære opp programvare basert på data (treningsdata), også omtalt som maskinlæring. Treningsprosessen inkluderer alle operasjoner som er nødvendig, herunder f eks selve treningen, evaluering og tilpassing.

Med **pretrening** menes treningsprosess for å bygge grunnmodeller.

Med **posttrening** menes en videre treningsprosess basert på mer spesialiserte treningsdata, enten for å sikre kvaliteten i grunnmodellene eller som en utvikling av spesialiserte modeller.

Med **spesialiserte modeller** menes modeller som bygger på grunnmodeller, men som i tillegg er trent på spesialiserte eller domenespesifikke data med et særskilt tjenesteformål for øye.

Med **tjenester** menes grensesnitt der brukere eller andre maskiner samhandler med en eller flere grunnmodeller og/eller spesialiserte modeller.

Med **tilpassing** menes å sikre at modellene har ytelser som er innenfor ønskede rammer.

Med **treningskorpus** menes et eller flere datasett fremstilt av Nasjonalbiblioteket bestående av rettighetshavernes pressepublikasjoner til bruk for trening av språkmodeller.

Med **basemodeller** menes arkitektur eller pretrente modeller Nasjonalbiblioteket trener grunnmodeller på.

## 5. Lisens for trening

Nasjonalbiblioteket gis, på de vilkår og betingelser som fremgår av denne avtale, en ikke-eksklusiv lisens til å benytte den del av publikasjonene, slik det fremgår av Avtalens § 2, til å fremstille et eller flere treningskorpus. Treningskorpusene kan brukes for Nasjonalbibliotekets egen trening av ulike språkmodeller, herunder posttrening som er nødvendig for å sikre kvalitet og unngå skjevheter i de grunnleggende egenskapene til generelle språkmodeller.

Eierskap og bruksrett til språkmodellene, slik de er modifisert etter treningen, er regulert i §§ 7 til 10.

Nasjonalbiblioteket skal ikke trene modeller i strid med rettighetshavernes interesser som omtalt i avtalens formål. Begrensningen gjelder også forskningsinstitusjoner som distribuerer modeller under denne avtalen, jf. § 8.

## 6. Nasjonalbibliotekets råderett ved trening

Nasjonalbiblioteket bestemmer selv hvilke publikasjoner som skal benyttes i trening av de ulike språkmodellene.

## 7. Deling av språkmodeller

Nasjonalbiblioteket kan dele de trente språkmodellene under vederlagsfrie, ikke-eksklusive bruksvilkår (lisens) for tredjepartsbruk. Lisensen som modellene skal tilgjengeliggjøres under fremgår av vedlegg 1 til avtalen.

Punkt 4d i lisensvilkårene i vedlegg 1 gjelder ikke for utgiver/utgivergruppe omfattet av denne avtale, så fremt disse ikke har reservert seg, jf. Avtalens § 16. Språkmodellene skal være dokumenterte i tråd med offentlige krav.

De trente språkmodellene kan deles for intern forskningsaktivitet i forskningsinstitusjoner og i forskningsprosjekt i samarbeid mellom forskningsinstitusjoner uten avtalte bruksbetingelser (lisensvilkår).

Forskningsinstitusjonene kan kun dele modellene, i opprinnelig eller modifisert form, på samme vilkår og under samme lisens som Nasjonalbibliotekets modeller.

## 8. Deling av språkmodeller trent i eller for forskning

Der Nasjonalbiblioteket lar norske universiteter trene språkmodeller for forskningsformål på materialet omfattet av denne avtalen gis disse institusjonene rett til å dele språkmodellene på samme vilkår som Nasjonalbiblioteket. Modellene må deles under samme lisens som Nasjonalbibliotekets modeller, jf. § 7.

Institusjonenes rett til å dele språkmodeller i tråd med denne paragrafen gjelder kun der Nasjonalbiblioteket har lovlig hjemmel til å dele korpus for trening i lov, forskrift eller avtale.

## 9. Deling av treningskorpus til øvrige tredjeparter

Denne avtalen gir ikke Nasjonalbiblioteket rett til å dele treningskorpuset med tredjeparter. Deling med tredjeparter kan hjemles i avtale med rettighetshaverne.

Dersom tredjepart har innhentet nødvendig tillatelse fra rettighetshaverne, vil Nasjonalbiblioteket etter nærmere avtale legge til rette for utlevering av treningsdata/treningskorpus til tredjepart for tjenesteutvikling. Nasjonalbiblioteket vil tilsvarende kunne bistå utgiver eller utgivergruppe med tilgang til eget innhold for deres egen tjenesteutvikling.

## 10. Rettigheter til treningskorpus og modeller

Ingenting i denne Avtalen skal begrense rettighetshavernes eierskap til, eller immaterielle rettigheter i og til, treningskorpuset og de pressepublikasjonene som inngår i treningskorpuset.

Rettighetshaverne oppnår ikke eierskap til noen språkmodeller som er utviklet ved bruk av treningskorpuset under denne Avtalen.

## 11. Ansvar for forhåndstrening av basemodeller

Denne avtalen er begrenset til å regulere Nasjonalbibliotekets trening av språkmodeller på de data som avtalen regulerer. Den berører dermed ikke eventuell pretrening som de basemodellene Nasjonalbiblioteket bruker har gjennomgått.

## 12. Oppfølging av tredjeparter-eventuelle krenkelser

Nasjonalbiblioteket forplikter seg til aktivt å følge opp eventuelle brudd på bruksbetingelsene, enten på eget initiativ eller etter varsel fra Kopinor. Dersom det avdekkes brudd på bruksbetingelsene, skal Nasjonalbiblioteket iverksette nødvendige tiltak for å stanse den ulovlige bruken. Nasjonalbiblioteket vil i slike tilfeller sende skriftlige advarsler, og hvis dette ikke etterkommes, si opp lisensen med aktøren det gjelder. Nasjonalbiblioteket skal uten ugrunnet opphold informere Kopinor om brudd og hvilke tiltak som er iverksatt. Partene er enige om at rettighetshaverne har rett og anledning til å forfølge eventuelle brudd på bruksbetingelsene. Nasjonalbiblioteket forplikter seg til i den forbindelse å bistå rettighetshaverne, blant annet gjennom dialog og med relevant dokumentasjon som Nasjonalbiblioteket har tilgjengelig, kan tilgjengeliggjøre med enkle midler og har anledning til å dele.

## 13. Personvern

Nasjonalbiblioteket skal opptre som selvstendig behandlingsansvarlig med hensyn til behandlingen av personopplysninger som skjer som en del av samarbeidet regulert i denne Avtalen. Begge parter erkjenner at de skal overholde alle gjeldende lover og krav og, om nødvendig, inngå en separat juridisk avtale vedrørende overføring og/eller behandling av personopplysninger, i henhold til EUs personvernforordning (2016/679) (GDPR).

## 14. Datasikkerhet

Når treningskorpuset er i besittelse av eller under kontroll av Nasjonalbiblioteket skal Nasjonalbiblioteket implementere og opprettholde rimelige fysiske, administrative og tekniske sikkerhetstiltak for å beskytte treningskorpuset mot utilsiktet eller uautorisert tilgang, offentliggjøring, bruk eller endring.

Nasjonalbiblioteket skal uten ugrunnet opphold varsle rettighetshaverne ved eventuell uautorisert tilgang, offentliggjøring, bruk eller endring av treningskorpuset og samarbeide med rettighetshaverne for å avhjelpe og løse sikkerhetsbruddet på en måte som er tilfredsstillende for rettighetshaverne.

Nasjonalbiblioteket skal ikke særskilt legge til rette for muligheten til å identifisere noen enkeltperson ut fra anonymiserte eller aidentifiserte personopplysninger som inngår i treningskorpuset.

## 15. Innhold og kvalitet på treningskorpuset

Nasjonalbiblioteket vil få tilgang til rettighetshavernes pressepublikasjoner for å fremstille et eller flere treningskorpuser på en «som den er»-basis og kan bruke dataene deretter.

## 16. Reservasjonsrett

Berørte rettighetshavere skal ha rett til å unnta sine verk fra avtalen, i forkant av avtalens ikrafttredelse eller senere, da med virkning for fremtidige treninger.

Partene avtaler nærmere rutiner for håndtering av reservasjoner. Rettighetshaverne kan fremstille en liste over artikler eller annet materiale som inngår i og som må fjernes fra treningskorpuset grunnet personvern hensyn, injurier, brudd på opphavsrett eller andre grunner som rettighetshaverne anser relevant. Materiale på listen skal ha tydelige identifikasjonsmarkører. Listen vil bli oppdatert hver tredje måned og Nasjonalbiblioteket plikter å gjennomgå listen på dette tidspunktet for å fjerne alle elementer på listen fra treningskorpuset. Nasjonalbiblioteket skal føre en oversikt over artikler/elementer som er slettet fra treningskorpuset og sikre at denne samsvarer med

rettighetshavernes liste. Rettighetshaverne skal ha tilgang til Nasjonalbibliotekets liste på forespørsel.

## 17. Vederlag

Det betales et årlig samlet vederlag på 45 000 000 kr (i 2026) for bruk av beskyttet materiale etter denne avtalen.

Vederlaget indeksreguleres, første gang med virkning for 2027-vederlaget, deretter årlig per 31. desember i henhold til endringer i Statistisk Sentralbyrås konsumprisindeks eller annen konsumprisindeks som måtte erstatte denne.

Justeringen foretas på grunnlag av endring i konsumprisindeksen fra 15. oktober forrige år til 15. oktober inneværende år og tilsvarende for de kommende år.

Vederlaget forfaller til betaling årlig tidligst per 30. juni. Kopinor sender faktura med 30 dagers betalingsfrist.

## 18. Mislighold og erstatning

Ved mislighold av avtalen gjelder en frist for retting på 30 dager fra skriftlig varsel er mottatt.

Dersom en part vesentlig misligholder avtalen og forholdet ikke rettes, kan den annen part heve avtalen med umiddelbar virkning. Heving skal erklæres skriftlig med angivelse av hevingsgrunn. Erstatning for mislighold kan kreves uavhengig av om avtalen heves.

Rettighetshaverne har ved denne avtalen ikke fraskrevet seg adgangen til selv å kreve erstatning og reise søksmål mot Nasjonalbiblioteket i anledning kopiering og bruk som ikke er tillatt ved denne avtalen. Har Nasjonalbiblioteket innfridd krav fra Kopinor for slik kopiering, plikter Kopinor å innfri rettighetshavernes krav vedrørende denne utnyttelsen.

## 19. Endringsadgang

Dersom vesentlige forutsetninger for avtalen skulle endre seg, har partene rett til å si opp avtalen med umiddelbar virkning eller be om reforhandling av vilkårene i avtalen, herunder størrelsen på vederlaget. Fra Nasjonalbibliotekets side er det en forutsetning at treningssettet som gjenstår etter at individuelle rettighetshavere eventuelt har unntatt sine verk fra avtalen, er volummessig, språklig, genremessig og tidsmessig representativt.

Alle tillegg og endringer i avtale og vedlegg skal være skriftlige.

## 20. Eventuelle krav fra utenforstående rettighetshavere

Partene har et likeverdig økonomisk ansvar for krav fra utenforstående rettighetshavere dersom det viser seg at avtalelisensbestemmelsene i åndsverkloven ikke omfatter den avtalte bruken. Ut over det er Kopinor forpliktet til å dekke eventuelle krav om vederlag eller erstatning som Nasjonalbiblioteket måtte motta fra norske og utenlandske rettighetshavere for så vidt kravene gjelde bruk som er betalt for og utført i overensstemmelse med denne avtalens bestemmelser, jf. for øvrig åndsverkloven § 64. Nasjonalbiblioteket er ikke forpliktet til å dekke slike krav, og Kopinor skal holde Nasjonalbiblioteket skadesløs for krav om kostnader som Nasjonalbiblioteket måtte bli påført i forbindelse med et slikt krav. Dersom krav mottas, skal Nasjonalbiblioteket avvise kravet og henvise til Kopinor. Krav skal ikke innfris uten etter samråd med Kopinor. Ved tvist om kravet har Kopinor rett til å intervenere.

## 21. Rapportering

Nasjonalbiblioteket vil, etter nærmere avtale, gi Kopinor tilgang til informasjon om innholdet i datasettet for fordelingsformål.

## 22. Overvåkingsorgan

Partene skal avholde regelmessige kontaktmøter for å overvåke og drøfte praktiseringen av avtalen og om utviklingen innen teknologi for kunstig intelligens og språkmodeller innebærer at avtalen bør justeres. Det påhviler Nasjonalbiblioteket et spesielt ansvar å overvåke hvorvidt lisensbetingelsene for videre bruk av språkmodellene overholdes, slik at rettighetshavernes økonomiske og ideelle rettigheter ikke krenkes.

## 23. Tvisteløsning

Tvist mellom avtalepartene om avtalens forståelse eller bestående, om brudd på avtalen eller om krav som på annen måte grunner seg på avtale mellom partene vedrørende kopiering og bruk, skal i avtaleperioden først søkes løst ved forhandlinger mellom partene.

Fører ikke disse forhandlinger fram, kan hver av partene kreve tvisten avgjort ved voldgift. Voldgiftsretten sammensettes av tre medlemmer. Partene oppnevner ett medlem hver og i fellesskap en leder. Blir partene ikke enige om leder, oppnevnes denne av lederen i Den Norske Advokatforening. For øvrig gjelder bestemmelsene i lov om voldgift.

Partene kan hver for seg begjære mekling etter åndsverkloven § 65.

## 24. Avtalens varighet

Avtalen trer i kraft ved signering og gjelder for 2026. Avtalen løper deretter videre for ett kalenderår av gangen, med mindre en av partene sier opp avtalen skriftlig innen 3 - tre - måneder før kalenderårets slutt (innen utløpet av 30. september), likevel slik at avtalen første gang kan sies opp innen utløpet av september 2027 med virkning fra 1. januar 2028. Avtalens opphør skal ikke ha virkning for språkmodeller som allerede er trent og delt på delingsplattformer i tråd med denne avtalen.

Oslo, den 19. desember 2025

### For Nasjonalbiblioteket

Trond Myklebust (sign.)

Fungerende Nasjonalbibliotekar

### For Kopinor

Hege Munch Gundersen (sign.)

Administrerende direktør