

Synpunkter inför kommande forsknings- och innovationspolitiska proposition

Allmänt

Wikimedia Sverige har inte blivit formellt inbjudna att inkomma med synpunkter på den kommande propositionen, men hoppas att dessa synpunkter ändå tas i beaktande.¹

För oss är det centralt att den kommande forskningspropositionen lägger stor vikt vid öppen vetenskap och nyttiggörande av forskning. Föreningen står bakom några av de viktigaste plattformarna för att sprida kunskap till allmänheten, bland annat Wikipedia – en av de tio mest besökta webbplatserna i Sverige och globalt. För att plattformarnas volontärer ska kunna bidra med tillförlitlig kunskap krävs dock god tillgång till källor, och därför är öppen vetenskap en nödvändighet för volontärerna.

Forskning och forskare är dessutom viktigt för att fylla de kunskapsluckor som fortfarande finns på Wikipedia och de andra Wikimediaplattformarna, som mediedatabasen Wikimedia Commons och den strukturerade länkade databasen Wikidata. Plattformarnas oerhörda räckvidd tillsammans med forskares djupa sakkunskaper gör att nära samarbete mellan akademi och Wikimediarörelse blir en omistlig del i att uppfylla högskolelagens tredje uppgift.

Sammanfattande ståndpunkter (utan inbördes ordning)

- Kulturarvets digitalisering behöver tillräckliga resurser, bland annat för att säkerställa god tillgång till forskningsmaterial.
- Upphovsrätten behöver anpassas så att forskare utan problem kan dela upphovsrättsligt skyddat material inom forskargrupper, och att material kan tillgås för text- och datautvinning digitalt och på distans.
- Initiativ bör tas till nationella regler/riktlinjer för att forskare ska kunna bevara sina rättigheter.
- Sverige bör säkerställa att integritets- och etikprövningsregler ligger på motsvarande nivå som övriga EU-länder.
- Upphovsrättslagstiftningen behöver revideras för att omfatta parallellpublicering.
- Regeringen bör se över hur man kan underlätta för forskare att få tillgång till material, och på EU-nivå driva frågan om en rätt för tillgång till material för forskare.
- Regeringen bör analysera hur ny internetlagstiftning påverkar lärosätenas möjlighet att sprida vetenskaplig forskning via sina fillagringstjänster.

¹ Wikimedia Sverige arbetar med dessa frågor inom ramen för nätverket [KnowledgeRights21](#), som finansieras med stöd av Arcadiastiftelsen. Nätverket ger oss djupa insikter i hur dessa diskussioner förs i Europa, och hur man har löst liknande utmaningar i andra EU-länder.

Samhällsutmaningar

Öppen vetenskap och tillgång till kunskap

Kunskapssamhället förutsätter god tillgång till kunskap. Det behövs för att medborgare ska kunna fatta välgrundade beslut, för att företag ska kunna innovera och för att journalister, forskare och andra för vilka informationsinhämtning är helt nödvändigt ska kunna bedriva sitt arbete. I detta samhälle spelar Wikipedia stor roll som en av de mest använda förmedlarna av kunskap. För att säkerställa att informationen som sprids via Wikipedia är korrekt, det vill säga att den utgör kunskap, är det centralt att akademien och Wikimediarörelsen samverkar, och därtill är ett centralt första steg att vetenskaplig forskning som standard görs öppet tillgänglig.

Betydelsen av detta blev tydlig under coronapandemin. Genom volontärernas arbete på Wikipedia och de andra Wikimediaplattformarna med att sammanställa och översätta forskning fick den fick stort genomslag. Volontärernas arbete bidrog till att viktig forskning kom allmänheten till del, men det byggde i sin tur på att de kunde få tillgång till den. Därför är öppen vetenskap fundamentalt för Wikipedia och de andra Wikimediaplattformarna. Det är en förutsättning för att Wikipedia ska kunna främja tillgång till kunskap utanför akademien, och i sin tur så att offentlig finansiering av forskning även kommer allmänhet till del.

Öppen vetenskap ger också rättvis och jämlik tillgång till forskning i civilsamhället. Medan företag kan ha råd att betala för forskning de behöver är det svårt för ideella krafter, som därmed får sämre tillgång till kunskap så länge vetenskapen inte görs öppen.

Desinformations spridning

Internet är en gigantisk källa till information, men dessvärre också till desinformation. Illvilliga krafter utnyttjar internetns öppenhet för att så split och medvetet sprida felaktigheter. Det är en central del av det öppna samhällets motståndskraft att hitta sätt att motverka detta.

Wikimediarörelsen har i tjugio år framgångsrikt utvecklat robusta strategier och metoder för att motverka att desinformation sprids via Wikimediaplattformarna. Det faktum att kommersiella, stora plattformar som Facebook och Youtube länkar vidare till Wikipedia för verifiering av kontroversiella uppgifter är ett tecken på hur robusta dessa strategier och metoder är. En förutsättning för att metoderna ska kunna fungera är dock att det finns god och bred tillgång till tillförlitliga och lättillgängliga källor, och där är vetenskaplig forskning central.

Bred tillgång innebär att vetenskap är öppet tillgänglig som standard. När Wikipediavolontärer källbelägger information utgår de i regel från vad som finns lättillgängligt, ofta källor som inte är bakom betalvägg. Råder det brist på öppen vetenskap har man sällan möjlighet att göra en värdering vilken forskning som är mest tillförlitlig eller håller högst kvalitet: Man använder det som finns tillgängligt, eller utgår från en kort sammanfattning som inte alltid ger hela bilden. I arbetet med en av de viktigaste samhällsresurserna för fri kunskap är detta ett stort problem. Det riskerar också att snedvrider kunskapen, om den kunskap som väl är öppen har öppnats upp av specifika intressen eller motiv. Bred tillgång till öppen vetenskap är ett effektivt verktyg för att motverka desinformations spridning. När vetenskap är lätt att tillgå och hitta kommer andelen innehåll på Wikipedia som stöds av vetenskaplig forskning att öka. På så vis blir öppen vetenskap även ett viktigt verktyg för att uppfylla högskolelagens tredje uppgift.

Artificiell intelligens

- Sverige bör säkerställa att integritets- och etikprövningsregler ligger på motsvarande nivå som övriga EU-länder.
- Kulturarvets digitalisering behöver tillräckliga resurser, bland annat för att säkerställa god tillgång till forskningsmaterial.
- Upphovsrätten behöver anpassas så att forskare utan problem kan dela upphovsrättsligt skyddat material inom forskargrupper, och att material kan tillgås för text- och datautvinning digitalt.

AI erbjuder enorma möjligheter och utmaningar. Utmaningarna handlar om hur den används, vem som använder den och vad den tränas på. Utan lösningar på dessa utmaningar riskerar EU att halka efter länder med flexibla lagar, som USA och Japan, eller repressiva, som Kina. Kinesisk AI riskerar att sprida användning av AI tränad på biometrik, ansiktsigenkänning och censurerad data, vilket återspeglar den auktoritära regimen världsbild. Stelbenta regler hotar så inte bara EU:s konkurrenskraft, utan även mänskliga rättigheter och EU:s grundprinciper. Att stödja europeisk utveckling av AI blir en viktig byggsten för främjandet av etisk AI.

AI-utvecklare har i stor utsträckning använt material från Wikipedia för att träna AI. Wikimedia Sverige har inga större invändningar mot detta, utan vill främja återanvändning och spridning av materialet. Däremot är vi medvetna om den stora bias som finns på våra plattformar. Det finns mer kunskap om män än om kvinnor, och den geografiska fördelningen är skev. Dessutom finns bias i hur olika fenomen beskrivs, och AI tenderar att förstärka bias i datan. För att utveckla en AI som i sig inte innehåller ett stort bias måste sådana skevheter tas i beaktande.

Wikimedia Sverige samarbetar med bland annat KTH för att utveckla en talsyntes för Wikipedia. Projektet innehåller AI-komponenter, där allmänheten involveras för att crowdsourca taldata till en talsyntes (exempelvis genom utveckling av lexikon med uttalsregler och inspelning av tal). Skärmläsare har ofta bara manliga röster på standarddialekt och taligenkänning fungerar olika bra beroende på kön, dialekt, brytning, talfel, osv. Omfattande och fritt licensierade datamängder gör produkterna bättre för alla, vilket innebär att initiativ där AI utvecklas med medborgarforskning är viktiga för att främja bra och etisk AI. Crowdsourcing gör att man kan arbeta riktat för att få in taldata från en mångfald av personer, speciellt de grupper som inte prioriteras av kommersiella aktörer, och därmed få en AI som fungerar likvärdigt för alla.

I dag ställer lagstiftningen till det för dylika projekt. Ett stort problem är GDPR, vars relation till AI är komplicerad. Fortfarande återstår många oklarheter, i synnerhet när det gäller taldata. Europeiska kommissionen är medveten om frågan men saknar också svar, samtidigt som de kraftiga sanktionerna i GDPR gör alla involverade ovilliga att pröva lagen. Liknande utmaningar medför Sveriges etikprövningslag. Integritet och etikprövning är mycket viktigt, men den rådande stelbentheten hämmar viktig forskning, inte minst inom humaniora. När Sverige har högre krav än övriga EU riskerar dessutom forskningsprojekt att flytta till dessa länder, eller utesluta Sverige i kommande ansökningar, med negativa följder för Sverige som forskningsnation. Därför är det viktigt att Sverige ligger på motsvarande nivå som resten av EU.

Samarbetet kring talsyntesen visar också på vikten av att tillåta och främja samarbeten mellan forskningsinstitutioner, företag och civilsamhälle, och att minimera mängden lagstiftning som ställer stelbenta krav på icke-kommersiella syften vad gäller forskning.

Legala hinder för TDM och AI-relaterade forskningsprojekt

När det är svårt för forskare att få laglig tillgång till material att träna AI på tar de det material som finns tillgängligt. En AI kan bara "veta" utifrån dess träningsdata. Bristen på digitaliserat material, äldre material, kunskap på andra språk än engelska och så vidare förstärker befintliga bias. Det kan få farliga konsekvenser inom områden som [medicin](#) och [rättsväsende](#). En viktig dellösning är storskaliga satsningar på kulturarvets digitalisering, för att ge bredare tillgång till material. Det behöver finnas tillräcklig finansiering för detta.

Ansatser till att förbättra den legala grunden gjordes genom artiklarna om text- och datautvinning (TDM) i Upphovsrättsdirektivet (URL § 15 a och b i svensk lag). Sedan dess har utvecklingen gått snabbt och lagstiftningen ställer nya problem. En översyn av EU:s regelverk görs inom [ERA Policy Agenda 2022–2024](#), men även svensk upphovsrätt behöver ses över med forskningsfokus, bland annat för att främja AI-relaterade forskningsprojekt och minimera gråzoner. Det kan göras proaktivt, i fall där EU-rätten redan möjliggör nationella lösningar.

En sådan utmaning är forskares möjlighet att dela material för eller resultat av TDM inom forskargrupper. Eftersom URL §15b bara omfattar exemplarframställan (kopiering) finns ingen rättslig grund för att överföra kopior till andra forskare, i synnerhet om andra forskare anses utgöra en del av allmänheten. Det är inte klarlagt om forskare utgör en allmänhet, men många EU-forskningsprojekt består av tiotals forskare fördelade över institutioner i många länder. Bristen på rättslig grund för överföring för forskningsändamål skapar en juridisk gråzon och stor oro. En lösning som redan är kompatibel med EU-rätt är att införliva Infosoc-direktivets undantag för kommunikation av forskning (art. 5.3.a), utifrån förslagsvis italienskt ([Art. 70 ter \(1\) LdA](#)) eller slovenskt ([57.b.1](#)) exempel. Störst effekt skulle detta få om 5.3.a gjordes obligatorisk i hela EU, vilket skulle underlätta för gränsöverskridande forskning inom EU.

En annan utmaning för forskare är att få tillgång till material digitalt för TDM-syften. I dag behöver forskare ta sig till institutionen som har laglig tillgång till materialet. Det kan vara svårt i ett stort land som Sverige, men blir mycket dyrt och tidskrävande om material är utspritt över flera länder, i synnerhet som TDM ofta kräver att man bearbetar material löpande och finjusterar utifrån behov. En lösning som den slovenska, ovan, löser även detta.

Lagändringarna säger heller ingenting om tekniska skyddsåtgärder (TPM) som utestänger forskare från möjligheten att ta del av material. Det blir ett problem när de inte kan ta del av material från exempelvis sociala medie-plattformar som Facebook – en central del av den digitala infrastrukturen och offentliga miljön – eller stora strömningstjänster som Netflix. Det blir dock även ett problem när material som egentligen är fritt och inte omfattas av upphovsrätt, eller där användningen skulle vara tillåten med stöd av en inskränkning, ändå låses in med hjälp av TPM. TPM för upphovsrättsligt fritt material såsom material i *public domain*, bör inte vara tillåtet, vare sig för forskare eller allmänhet. Fritt material ska vara fritt.

Behov för ökat nyttiggörande

Långsiktig infrastruktur för öppen vetenskap

- Initiativ bör tas till nationella regler/riktlinjer för att forskare ska kunna bevara sina rättigheter.

Öppen vetenskap är viktig som princip eftersom det innebär att forskning som finansieras av allmänheten också är tillgänglig för allmänheten, vilket även konstaterades i den förra forskningspropositionen. Ofta tvingas dock lärosäten spendera stora summor för att göra vetenskap öppet tillgänglig, då de mest meriterande vetenskapliga tidskrifterna genom de gällande transformativa avtalen kräver ersättning för öppen publicering av vetenskap. Det är bra att de stora forskningsfinansiärerna [tar på sig publiceringskostnaden](#) för att främja ett öppet tillgängligt publiceringssystem, men det innebär i sig inte nödvändigtvis minskade publiceringskostnader. De stora kommersiella tidskriftsförlagens priser saknar skälig grund, ibland tiotusentals kronor för en artikel. Summor som skulle kunna ha gått till forskningen i sig. Det är därför viktigt att bygga en långsiktig infrastruktur som minskar beroendet av de marknadsdominerande kommersiella förlagen, exempelvis genom så kallad [diamant-OA](#).

Ett problem i sammanhanget är att det är svårt för lärosäten, i synnerhet mindre, att förhandla med de stora kommersiella tidskriftsförlagen. På nationell nivå bör stöd ges till lärosäten för att göra förhandlingsläget mer gynnsamt. Det handlar inte minst om möjligheten för forskare att bevara sina rättigheter till vetenskapliga artiklar och forskningsresultat, en förutsättning för att forskaren ska kunna sätta fria licenser (såsom [CC-licenser](#)) på sina verk. Nationella riktlinjer, ännu hellre regler, bör tas fram för hur forskare ska kunna bevara rättigheter. Ett problem som uppstår när forskare förlorar rättigheter gäller [sammanläggningsavhandlingar](#), som blir vanligare inom många discipliner. Många doktorander har svårt att få tillstånd att använda sina egna publicerade artiklar i sina avhandlingar, eftersom de inte har rättigheter. Förhandlingen för att få använda sin egen artikel i avhandlingen kan vara tidskrävande, och flera avhandlingar har försenats bara av den anledning att tidskriftsförlagen inte vill ge sitt tillstånd.

Rapporten ”[Retaining Rights and Open Licensing in Europe 2023](#)” gör en översyn av hur olika länder i Europa hanterar frågan om forskares rätt till sina verk. Storbritannien har kommit längst med att garantera forskares rätt, men fyra svenska lärosäten berör detta i sina policyer för öppen tillgång: [Stockholms universitet](#), [Chalmers](#), [Lunds universitet](#) och [Högskolan i Skövde](#). Utvecklingen sker dock per lärosäte, snarare än med nationell koordinering. Det finns ett stort behov av nationell koordinering, att insikter sprids från de som har gått före till övriga lärosäten. I samtal vi haft med forskningsbibliotek är det tydligt att en utmaning för många är just förhandlingsläget. Nationella regler för att forskare ska kunna behålla sina rättigheter vore ett viktigt steg framåt och skulle göra förhandlingsläget bättre. Här skulle Bibsam kunna få ett tydligt mandat. [Deras handlingsplan](#) 2022–2024 innehåller viktiga punkter, men ofta lämnas de enskilda lärosätena ändå i en problematisk förhandlingssituation gentemot tidskriftsförlagen.

Upphovsrätten för forskningsresultat

→ Upphovsrättslagstiftningen behöver revideras för att omfatta parallellpublicering.

Öppen vetenskap prioriterades under svenska EU-ordförandeskapet, vilket ledde till rådsslutsatser som lyfter förändringar som måste till även på nationell nivå, bland annat lagstiftning som ”säkerställer upphovspersoners rätt att omedelbart öppet tillgängliggöra sina forskningspublikationer oavsett eventuella överenskommelser med förlag.” Detta återges även i [delredovisningen](#) av KB:s arbete med nationella riktlinjer.

Flera EU-länder har redan infört parallellpublicering, eller sekundära publiceringsrättigheter som det heter i rådsslutsatserna, i lag. Det innebär att offentligt finansierad forskning görs fritt tillgänglig omedelbart eller inom en kort tidsperiod, och att forskare *har rätt* att publicera sina

verk som öppen vetenskap. I viss mån finns redan laglig grund för detta i svensk lag. SUHF konstaterar i sin [färdplan för öppen vetenskap](#) att den rådande uppfattningen är att forskningsdata och resultat utgör allmän handling, och att det även gäller artiklar, avhandlingar och rapporter: ”Det innebär att vi redan idag har en passiv form av öppen tillgång till forskningsdata och forskningsresultat.” Mycket skulle dock vara vunnet om denna passiva form av öppen tillgång gjordes aktiv, och parallellpublicering infördes som en direkt bestämmelse.

För att främja öppen vetenskap i Sverige bör liknande bestämmelser införas även i svensk lag. En bra lösning som verkligen skulle främja öppen vetenskap, och därmed spridningen av forskning i samhället, vore att *ställa krav* på universitet och högskolor att offentligt finansierad forskning parallellpubliceras som öppen vetenskap, vilket bland annat är fallet i Spanien.

Forskares tillgång till material

- Regeringen bör se över hur man kan underlätta för forskare att få tillgång till material, och på EU-nivå driva frågan om en rätt för tillgång till material för forskare.

Utan tillgång till material att forska på kan forskning och innovation inte ske. I dag är det en utmaning för många forskare att få tillgång till material de vill beforska, särskilt då tillgången i allt högre grad styrs av avtal snarare än upphovsrättigheter. Forskare bör ha lagstadgade möjligheter att få tillgång till material och data för forskningssyften, material som annars är bakom avtalsväggar, brandväggar (vad gäller exempelvis algoritmer) eller föremål för tekniska skyddsåtgärder. I dag är det mycket svårt att beforska exempelvis ljud- och filmmaterial via strömningstjänster som Spotify och Netflix, eller data och algoritmer från stora kommersiella plattformar som Facebook. Avtalsberoendet bidrar dessutom till att datan fragmenteras, vilket i sig sänker validiteten i resultatet och gör det svårare att jämföra studier. Det bidrar också till att forskare forskar på de plattformar de kan få avtal med, vilket inte nödvändigtvis är detsamma som de plattformar som är av störst relevans för just det forskningsprojektet. Tillgången på avtal styr således forskningen snarare än plattformens relevans i sammanhanget.

Både forskningsinstitutioner och forskare befinner sig i ett förhandlingsunderläge när det gäller avtal. Godkänner man inte villkoren får man inte tillgång till undervisnings- eller forskningsmaterial. Här finns en stor maktobalans. Obalansen är för stor för att kunna prata om avtalsfrihet. För många forskare råder också en stor osäkerhet kring hur de på ett lagenligt sätt kan få tillgång till material, liksom hur de får använda materialet. Utifrån upphovsrättsligt perspektiv är den ekonomiska skadan för uphovspersoner liten vid användning av deras verk för forskningssyften, medan samhällsintresset kan vara mycket stort. Forskare bör därför ha en särskild rätt till tillgång till material för sin vetenskapliga forskning.

Ansvarsskyldigheter för fillagringstjänster

- Regeringen bör analysera hur ny internetlagstiftning påverkar lärosätenas möjlighet att sprida vetenskaplig forskning via sina fillagringstjänster.

För spridning av vetenskaplig forskning är digitala publiceringssystem centrala. År 2021 hade exempelvis DiVA, där studenter och forskare själva kan publicera sina alster, fler än 500 000 publicerade fulltexter, som laddades ned mer än 75 miljoner gånger per år. Det motsvarar drygt [200 000 gånger om dagen](#), eller knappt 9 000 nedladdningar i timmen. Det säger sig självt att ett publiceringssystem som omfattar 500 000 fulltexter, från fler än femtio lärosäten och övriga medlemmar, medför utmaningar. På många sätt liknar dessa utmaningar de som Wikipedia och

de andra Wikimediaplattformarna står inför, som stora, välanvända, användargenererade plattformar som drivs icke-kommersiellt i folkbildningssyfte.

Länge har icke-kommersiella plattformar haft begränsade ansvarsskyldigheter för innehåll som användare tillgängliggör. I Upphovsrättsdirektivets artikel 2.6 undantas exempelvis både onlineencyklopedier som Wikipedia och fillagringstjänster som DiVA från upphovsrättsligt ansvar vid eventuellt intrång. Lagstiftaren var ute efter stora, kommersiella plattformar där det sker omfattande intrång i upphovsrätten, inte ideella folkbildningsplattformar med välutvecklade metoder för att motverka intrång och lagbrott.

Detsamma gäller dock inte i fallet med [rättsakten för digitala tjänster](#) (DSA). Där omfattas Wikipedia – som en av 17 ”[VLOPs](#)” – liksom även fillagringstjänster som DiVA. Det är nödvändigtvis inte orimligt att ställa krav på onlineplattformar att motverka att innehåll som av något skäl inte hör hemma på plattformen plockas bort. Det är dock oerhört viktigt att motverka en utveckling där kraven som ställs blir så betungande att bara stora, kommersiella aktörer klarar av att uppfylla dem. Det vore en stor förlust för internet, men i synnerhet för spridning av kunskap och folkbildning. Sveriges regering bör därför noga undersöka hur man, i enlighet med europeisk och svensk rätt, kan arbeta tillsammans med icke-kommersiella internet-aktörer, för att säkerställa att mängden olagligt innehåll minimeras, samtidigt som bördan inte blir större än vad de icke-kommersiella kunskapsspridande aktörerna kan hantera.

Lagstiftningen såsom den är formulerad tar sikte på stora, globala kommersiella plattformar som har resurser långt bortom folkbildningsaktörer och forskningsinstitutioner. Det framstår heller inte som en optimal resursanvändning att tvinga forskningsinstitutioner att lägga omfattande offentliga medel på juridisk prövning snarare än akademisk spridning.

Medborgarforskning

Medborgarforskning är centralt för att öka tilliten till och engagemanget i högre utbildning, och bör därför bli en prioritet i den kommande forsknings- och innovationspolitiska propositionen.

Diskussionen om nyttiggörande av forskning handlar i hög grad om hur akademiska resultat kan komma återanvändare (som Wikimediarörelsen) och allmänheten till godo. Vi har dock i ökande grad sett hur akademien och forskarsamhället engagerar sig i och samarbetar med Wikimediavolontärer, bland annat i den strukturerade länkade databasen Wikidata, för att knyta samman forskningsdata med öppen data, samt strukturera, berika och därmed förbättra den. Flera stora forskningsprojekt har under bara de senaste åren i sina forskningsansökningar aktivt inkluderat arbete med Wikidata och samarbete med dess gemenskap. Detta har positiva resultat både för den akademiska forskningen och för den fria kunskapen (via Wikidata), och samarbeten mellan medborgare och forskare av liknande slag bör uppmuntras.

Medborgarforskning är också centralt i projektet Wikispeech som KTH driver tillsammans med Wikimedia Sverige och Språkbanken Tal, där medborgare bistår med taldatainsamling och transkribering. Den taldata som samlas in kan nyttjas av forskarsamhället, och andra intresserade, för att skapa och förbättra *large language models* (LLMs) som används för talsyntes.

Utöver att öka tilliten till och delaktigheten i forskningen kan medborgarforskning även bidra till en större mångfald i forskningen, när människor från olika delar av samhället och med olika bakgrunder får förutsättningar för att bidra.