

# Kommentarer till *Agreement on Reforming Research Assessment*

## Bakgrund

*Agreement on reforming research assessment (ARRA)*<sup>1</sup> är en överenskommelse om att utveckla riktlinjer för att bedöma forskning med ett uttalat mål att maximera forskningens kvalitet och genomslag. Det har utarbetats av ett konsortium lett av Science Europe, European University Association (EUA) med stöd av Europeiska kommissionen. Överenskommelsen har tagits fram under överinseende av över 350 forskningsorganisationer från över 40 länder. Organisationer som skriver under ingår i en koalition, Coalition for Advancing Research Assessment (CoARA), som ska arbeta för att implementera riktlinjerna som är beskrivna i överenskommelsen.

Överenskommelsen innehåller riktlinjer och principer för bedömning utav såväl forskning som forskare och forskningsutförande organisationer. Huvudbudskapet är att utvärdering av forskning i första hand ska baseras på kvalitativ bedömning, med huvudfokus på peer review, men med visst stöd av “ansvarsfull användning” av kvantitativa indikatorer. ARRA efterfrågar också bredare bedömningar som inkluderar andra forskningsaktiviteter, bortom vetenskapliga publikationer, som tar hänsyn till skilda discipliner och forskningsinriktningar. Vi, SUHF:s arbetsgrupp för bibliometri, försöker i följande text belysa vad som åsyftas med ansvarsfull användning av kvantitativa indikatorer, med fokus på bibliometri, samt ge vår bild av dokumentets förtjänster och potentiella problem.

För att ge en nyanserad bild vill vi poängtera att bibliometri (ibland benämnt scientometri) är ett etablerat vetenskapligt fält med egna tidskrifter, konferenser och vetenskapliga organisationer som alltsedan 1970-talet utvecklats i relation till närliggande ämnen som vetenskaps- och teknikstudier, forskningspolitiska studier och biblioteks- och informationsvetenskap i vad som kan benämnas *Forskning om forskning*. Teoretisk och metodologisk utveckling under de senaste decennierna har gjort att bibliometri numera på ett relativt träffsäkert sätt kan fånga olika faktorer av forskningen. Det är i ljuset av ett ständigt utvecklande bibliometriskt forskningsfält som vi gör nedanstående reflektioner.

## Innehållet i ARRA

ARRA riktar sig till bedömning av forskning på tre nivåer; forskningsutförande organisationer; forskningsprojekt och individuella forskare och forskargrupper. Tanken är att koalitionens medlemmar ska använda riktlinjerna för såväl allokering av forskningsmedel som vid meriteringsbedömning och rekrytering av forskare, och således riktar sig ARRA både till forskningsfinansiärer och forskningsutförande organisationer.

ARRA består av ett antal principer om övergripande villkor för forskningsutvärdering och åtaganden. Principerna förefaller välgrundade och lyfter fundamentala värderingar i vetenskapssamhället, som exempelvis vikten av etik och integritet, akademisk frihet,

---

<sup>1</sup> CoARA (2022): Agreement on Reforming Research Assessment, <https://coara.eu/agreement/the-agreement-full-text/>

organisationernas autonomi, mångfald, jämställdhet samt öppna och transparenta data i utvärderingsprocessen.

Därefter följer de tio åtagandena där de fyra första anges som ”core commitments” och det är dessa vi kommer att fokusera på i detta dokument då de tydligast relaterar till frågor som vi i Arbetsgruppen för bibliometri arbetar med. De sex sista anges som ”supporting commitments” och handlar mest om hur ARRA skall implementeras av överenskommelsens medlemmar.

#### ARRA:s core commitments:

1. Erkänna mångfalden av bidrag till och karriärer inom forskning i enlighet med forskningens behov och natur.
2. Basera forskningsbedömningen främst på kvalitativ utvärdering där peer review är centralt, med stöd av ansvarsfull användning av kvantitativa indikatorer.
3. Överge olämplig användning vid forskningsbedömning av tidskrifts- och publikationsbaserade mätvärden, särskilt olämplig användning av Journal Impact Factor (JIF) och h-index.
4. Undvik att använda rankningar av forskningsorganisationer vid forskningsbedömning.

Trots att vi ser positivt på ARRA:s generella ansats och inriktning så har vi en del synpunkter på dess innehåll. I det följande diskuterar vi några specifika punkter som vi ser som problematiska med ARRA: innebörden av ansvarsfull användning av kvantitativa indikatorer, relationen peer review och bibliometri, samt förutsättningar för utvärdering på olika nivåer.

## Oklarheter kring innebörden av ansvarsfull användning av kvantitativa indikatorer

Enligt ARRA ska bedömning av forskning baseras på kvalitativ bedömning där peer review är centralt och att den endast kan stödjas av ansvarsfullt användande av kvantitativa indikatorer. Användande av h-index och Journal Impact Factor ges som exempel på icke ansvarfull användning av bibliometri. När det gäller vad som anses vara ansvarfull användning av bibliometri saknas däremot konkreta exempel, vilket vi ser som en svaghet med dokumentet. Det hänvisas till Leidenmanifestet<sup>2</sup>, DORA<sup>3</sup>, INORMS-projektet<sup>4</sup> och den brittiska rapporten *The Metric Tide*<sup>5</sup>, vilka samtliga har framtagits eller utvecklats inom det bibliometriska och forskningspolitiska professionsområdet. Det är anmärkningsvärt att ARRA-överenskommelsen till så stor del bygger sin argumentation på resonemang från dessa källor, gällande behovet av ett reformerat forskningsutvärderingsförfarande, utan att notera att de är framtagna i en bibliometrisk kontext. Samtidigt avviker ARRA på viktiga punkter från exempelvis Leidenmanifestet som inte alls menar att man ska undvika kvantitativa metoder. I Leidenmanifestets första punkt anges exempelvis att kvantitativa metoder behövs för att utmana den bias/subjektivitet som kan förekomma under kollegial granskning och att sådan information kan underlätta beslutsprocesser. Även Leidenmanifestet anger att JIF och H-index är olämpliga mått för specifika utvärderingspraktiker men erbjuder samtidigt ett större tillämpningsområde för kvantitativa indikatorer än vad som går att finna i ARRA.

---

<sup>2</sup> Hicks, Diana et al. The Leiden Manifesto, *Nature*, 520, 429-431, <https://doi.org/10.1038/520429a>

<sup>3</sup> The Declaration on Research Assessment (DORA) (2012), <https://sfdora.org>

<sup>4</sup> INORMS Research Evaluation Working Group (2022): Fair and responsible university assessment.

<https://inorms.net/wp-content/uploads/2022/07/principles-for-fair-and-responsible-university-assessment-v5.pdf>

<sup>5</sup> Wilsdon, James, et al. (2015) *The Metric Tide*, HEFCE. <https://doi.org/10.13140/RG.2.1.4929.1363>

Vi menar att ARRA:s starka skrivningar riskerar att leda till oklar implementering om de läses allt för bokstavligt. Men genom dess hänvisningar till ovanstående beskrivningar av ansvarsfull användning av bibliometri uppfattar vi att det inte är avsikten att utestänga kvantitativa indikatorer som underlag för forskningsutvärdering helt och hållet.

## Relationen peer review och bibliometri

ARRA anger peer review som den mest robusta utvärderingsmetoden (s. 5) och menar att den ska användas vid utvärdering på olika nivåer. Trots att det tas upp många nackdelar med kvantitativa metoder saknas helt problematisering kring peer review. Potentiella problem med peer review, som inte nämns, är till exempel att det är enormt resurs- och tidskrävande, det är svårt att finna kompetenta sakkunniga att utföra det och att det är svårt att tillämpa peer review på hög övergripande nivå och på tvärs av ämnesområden. Det finns även en ”reviewer’s fatigue” och en viss misstro till peer review i forskarsamhället, vilket har framhållits av bland annat Liv Langfeldt (NIFU)<sup>6</sup>, som inte alls diskuteras i ARRA.

Som nämndes i punkten ovan finns det enligt Leidenmanifestet fördelar i användandet av bibliometri, som objektiv tillämpning, effektivitet och möjligheterna till årlig eller löpande uppföljning. Ingenting av detta nämns i ARRA. ARRA ger därmed en onyanserad och ensidig bild i relationen mellan kvalitativ och kvantitativ bedömning. Man nämner heller inte hur kvalitativa och kvantitativa former av utvärdering kan komplettera och utmana varandra, på ett sätt som kan styrka de slutgiltiga slutsatserna. Även på denna punkt har bibliometrifältet kommit med relevanta inspel, exempelvis Leydesdorff et al.<sup>7</sup>.

Vidare har olika forskningsapproacher och ämnesinriktningar skilda förutsättningar. Medan utvärdering av viss forskning huvudsakligen bygger på publiceringsdata, kan annan, mer tillämpad eller samhälleligt förankrad forskning värderas enligt delvis annorlunda principer. Likaså kan publiceringsformer, språk och tilltänkt publik skilja sig åt. Alla dessa variationer kräver genomtänkta utvärderingsinstrument. Detta innebär att balansen mellan peer review och kvantitativa indikatorer bör se olika ut vid bedömning av olika ämnesområden.

Vi menar att ARRA lite för lättvindigt skyller den uppkomna publish-or-perish-kulturen enkom på begränsade tidskrifts- och publiceringsbaserade indikatorer. I bakgrundsbeskrivningen till varför en förändring krävs står att vi måste röra oss bort från ”publication-based metrics” för att återfå förtroendet för forskning. Här kan inte en övergång till sakkunniggranskad utvärdering vara hela lösningen.

## Utvärdering på olika nivåer

ARRA innehåller riktlinjer för bedömning på olika nivåer: från organisationer, enheter (”units”), forskningsprojekt till individuella forskare. Det saknas en diskussion om hur förutsättningarna för bedömning skiljer sig mellan dessa nivåer. Undertecknad arbetsgrupp tog 2016 fram ett dokument ”Råd för utvärderande bibliometri i Sverige”<sup>8</sup> där det poängteras att val av utvärderingsmetod bör beakta utvärderingsnivå. Peer review är särskilt användbart för utvärdering av enskilda forskare eller mindre forskargrupper. Däremot är peer review mycket

---

<sup>6</sup> Langfeldt, Liv (2021): The many dilemmas of grant peer review. *R-QUEST Policy Brief*, no. 6 <https://www.r-quest.no/wp-content/uploads/2021/11/R-QUEST-Policy-Brief-6-2021.pdf>

<sup>7</sup> Leydesdorff, Loet et al. (2020): Bridging the divide between qualitative and quantitative science studies. *Quantitative Science Studies*, 1(3): 918–926. [https://doi.org/10.1162/qss\\_e\\_00061](https://doi.org/10.1162/qss_e_00061)

<sup>8</sup> SUHF:s forum för bibliotekschefer arbetsgrupp för bibliometri (2016): Råd för utvärderande bibliometri i Sverige, [https://suhf.se/app/uploads/2019/03/Dnr-0060-16-R%C3%A5d-utv%C3%A4rderande-bibliometri\\_161202.pdf](https://suhf.se/app/uploads/2019/03/Dnr-0060-16-R%C3%A5d-utv%C3%A4rderande-bibliometri_161202.pdf)

resurskrävande på aggregerade nivåer som exempelvis organisationsnivå. Kvantitativa metoder kan på dessa nivåer vara ett lämpligare alternativ. Vi ser ARRA:s ensidiga fokus på peer review som problematisk, både från ett praktiskt genomförandeperspektiv som ett rent metodologiskt, utifrån möjligheterna till årlig eller löpande uppföljning.

ARRA uppmanar deltagande organisationer att sträva efter att “undvika motsägelser mellan utvärdering av forskning, forskare och organisationer”, samt mellan organisationer (s. 7). Vi instämmer i att motsägelser bör undvikas. Däremot ser vi risker med att försöka likforma utvärderingsmodeller allt för mycket: Det är inte en motsägelse att ta de olika förutsättningarna för dessa nivåer i beaktande. Om samma utvärderingsmodell används på olika nivåer finns det en tydlig risk att modellen fungerar sämre på endera nivån. Man bör också ha i beaktande att syftet med varje form av utvärdering bör motiveras.

### Sammanfattande resonemang

Vi ser positivt på ett perspektiv på utvärdering som tar ett sammantaget grepp om hela forskningsprocessen, tar hänsyn till olika forskningsområdets förutsättningar och betonar vikten av öppen vetenskap och transparens i utvärderingens genomförande. Dock anser vi att ARRA:s starka skrivningar mot användningen av kvantitativa indikatorer, när det samtidigt tillåter viss *ansvarsfull* användning av desamma, är otydliga och kan leda till onödigt starka principiella ställningstaganden.

ARRA antyder att en övergång från tidskrifts- och publiceringsbaserade indikatorer till att endast använda sakkunniggranskning skulle förbättra den nuvarande publish-or-perish-kulturen. Vi anser att detta är naivt, då den snarare kan ses som en effekt av mycket större styrningar mot effektivitetsdrivande och ansvarsutkrävande principer i samhället.

ARRA betonar huvudsakligen bedömning på individuell nivå och vi anser att det är enklare att se hur ARRA kan implementeras på denna nivå. Däremot har vi svårt att se hur ARRA:s principer skall kunna appliceras på exempelvis organisationsnivå där praktiska förutsättningar och resurser riskerar att bli oöverstigligen.

Slutligen skulle vi gärna se ett konkretiserande resonemang om hur kvalitativa och kvantitativa metoder för utvärdering kan kombineras på olika nivåer och tidshorisonter. Denna konkretisering kan eventuellt behövas både på EU-nivå, via ett kompletterande dokument, men skulle även kunna ske i en nationell kontext, för att bidra till en mer likformig implementation hos svenska lärosäten och forskningsfinansiärer.