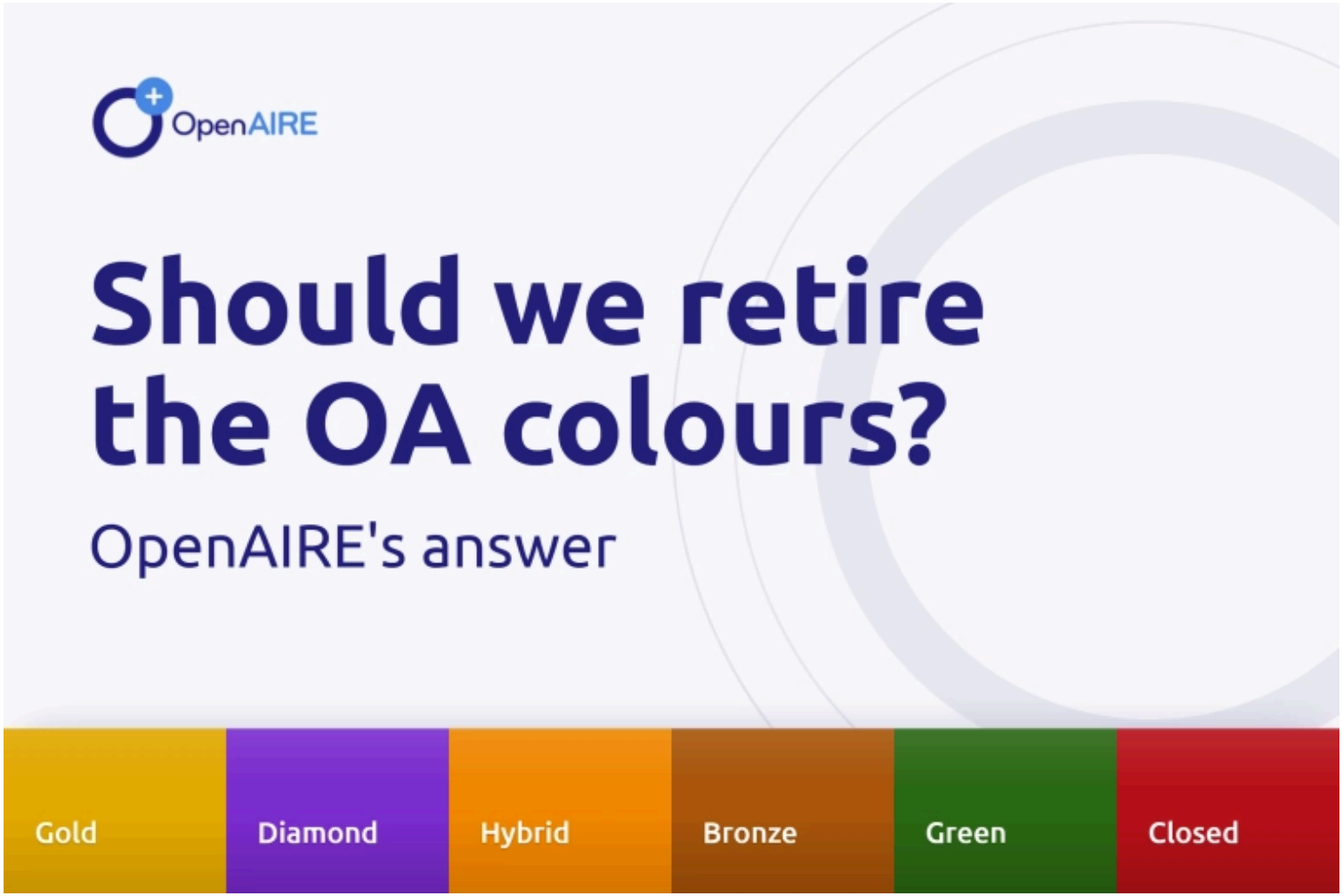
Written on Nov 12, 2026.



OpenAlex has asked a good question in public: is it time to retire the open access colours (Gold, Green, Hybrid, Bronze and Closed)? It is a fair question, and one the whole community should answer.

The question comes from the report of the OpenAlex Colors Working Group (Alperin et al., 2026). The report names four problems with the colours and proposes a fix. Here is our answer to each.

THE REPORT SAYS	OUR ANSWER
The colours were built for journal articles only.	Agreed. We describe datasets and software the same way, copy by copy, with access and licence. They get no colour, because the colours assume a journal.
The colours follow journal policy, not the work.	Not in our case. We look at the article's own copies, and the journal's class is only one input. If the copy on the journal is closed, the article is not Gold, even if the journal is listed as fully open access.
The colours mix access with business model.	In the names, yes. In our data, no: the two are kept separate. The colour combines them on purpose, because funders want both answers in one word.
The colours ignore time.	Agreed. We have the same gap. Fixing it needs versions and dates from sources, repositories and publishers alike.
Proposal: describe each work by access, location and licence.	We already record these three facts for every copy. We can publish them next to the colours.

Our position in one line: the information is already there; keep the colour names people know, show the data underneath them clearly, and close the remaining gaps together.

In this post we make three points.

We agree there is a problem. The report is right, and we see the same issues in our own work.

We already have the information, and we already build the colours from it. For every work, the OpenAIRE Graph keeps each copy separately: where it is, whether it is open, and what licence it has. Journals are classified on their own. So the three facts the report asks for are already in our data.

The main task is to make it easy to understand. Funders, ministries and research offices know the colour names and use them, so we would keep the names. What needs to change is how clearly we explain what each colour means and how it is built, and showing the three facts next to the colours. Where real gaps remain, such as licences and article versions, we propose working on them together with OpenAlex and others.

How OpenAIRE builds the colours today

The definitions are the usual ones:

- **Gold**: open in a fully open access journal.
- **Diamond**: open in a fully open access journal that charges no fees. We report it separately, so our Gold figures exclude diamond publications.
- **Hybrid**: open, with a licence, in a subscription journal that also publishes some open articles.
- **Bronze**: free to read on the publisher's site, but with no licence.
- **Green**: an open copy in a repository, with or without a licence.

What matters is how we decide. In the OpenAIRE Graph, a work is one merged record, and every copy of it that any source reports is kept as a separate instance, with the access right and licence that source declared and the data source that hosts it. So we do not look only at the journal. We look at the article's own copies. As we run it today, this works in three steps.

Step 1: classify the journals. We use public lists: a reconstructed Gold OA ISSN list for fully open access journals, DOAJ's APC field and the Diamond Discovery Hub for Diamond, and observed mixed content for Hybrid. We are now adding further sources, including the Plan S Journal Checker Tool and publisher journal model lists. The class belongs to the journal, not to the article.

Step 2: look at the copy at the journal. We test it in a fixed order: Diamond, then Gold, then Hybrid, then Bronze. A work gets the first colour that fits. If the copy at the journal is closed, it gets none of these, even if the journal is fully open access.

Step 3: look at repository copies separately. If any copy in a repository or CRIS is open, the work is Green. It is "Green with licence" if at least one of those copies has open access rights and a licence.

This gives three useful results:

- A work can be Gold and Green at the same time. One label per work would lose that.
- The journal's business model is one input, not the answer. It cannot make a closed copy open.
- Datasets and software are described the same way, copy by copy, but get no colour, because the colours assume a journal.

Our “Green without licence” is the report’s Platform Public. For “Green with licence” we also record which licence the copy carries, so it can be split into the report’s Platform Open and Platform Reserved. In the Irish national monitor, about two in three Green works now carry a licence on at least one open repository copy; the remaining third do not. Repository licence metadata is improving, but a third of deposits still say nothing about reuse.

The definitions are published on the Irish monitor's terminology page. New sources will be documented there as we add them. The vocabulary itself does not change.

One article, three copies

Mereghetti, S., Rigoselli, M., Salvaterra, R. et al. A magnetar giant flare in the nearby starburst galaxy M82. Nature 629, 58–61 (2024). https://doi.org/10.1038/s41586-024-07285-4

COPY	WHERE	TYPE	ACCESS RIGHTS & LICENCE
1	Nature	Journal	Closed Access
2	HAL	Repository	Open with CC-BY licence
3	arXiv	Repository	Open, no licence recorded

Our colour: Green with licence, because copy 2 is in a repository with a licence. Nature’s model is Hybrid, but the copy at Nature is closed, so the article does not take the journal colour.

The report's labels, from the same rows: Platform Open (copy 2), Platform Public (copy 3). The closed copy at Nature gets no label of its own.

The same data gives both answers. The colour summarises; the labels list each copy.

(NB: This is a simplified example. There are more copies of this article in the OpenAIRE Graph, but they belong to the same 3 categories, and do not change its colours.)

Where we have the same problems

Licences. Today we check whether a licence is recorded, not what it allows. Any licence counts, and that is too generous. We have the licence strings and about 98% are normalised. What is missing is an agreed grouping into open, restrictive and unknown that several infrastructures accept. This is the most important gap: without it, the three facts are not usable. It is shared work that needs to be maintained: OpenAlex, OpenAIRE, DOAJ and others each hold part of the list.

Journal history. We update journal classes once a year and apply them to all years. So if a journal changes model, its old articles take the new status; our terminology page says so. The report (section 4.5) suggests classifying journals by what they actually publish rather than by their stated policy. We do a bit of both: lists for fully open access and Diamond journals, observed content for Hybrid. Lists are easier to check; observed behaviour would catch changes sooner. Nobody has yet compared the two.

Versions and time. The three facts say nothing about which version a copy is (preprint, accepted manuscript, published) or when it became open. Policy monitoring needs both. In our 2021 Horizon 2020 study for the European Commission, 11.2% of deposited publications had no usable version, and on embargoes we could only report that "it is only possible to know the access rights of the last updated version". Repositories could fix most of this by exposing three things: the version (using the COAR vocabulary already in the OpenAIRE Guidelines), the date of deposit and the date of last update.

Access values. The report allows only "free to read" or "closed". Sources also say "embargoed" or "restricted", and often say nothing: 16.9% of peer-reviewed publications in the Irish monitor have no access status (unspecified access rights). We need "embargoed" and "unknown" as values. Embargoed is the one fact about time that sources sometimes provide. The proposal already allows "unknown" for licence, and should allow it for access too.

What we can do now

- **Publish both side by side**. Because our data is per copy, we can compute the report's labels from the same records as the colours, and publish both during a transition, once the licence grouping above is agreed. This avoids reusers splitting

into two camps, a risk the report raises in section 7.2.
- **Add versions and dates** when repositories expose them.
- **Track change over time** for groups of works, by comparing our archived monthly releases. Tracking a single work over time needs dates that sources do not yet provide.

What we propose

Keep the names, agree on how they are built. The report is right that Gold and Hybrid differ only by the type of journal (section 4.3), and that Gold has come to mean "APC-funded" (section 5.2). But colour users (funder policies, national dashboards, reimbursement rules) want two answers in one word: can it be read, and in what kind of venue. The report's own section 5.2 says so. The report's two-part system makes users join these answers themselves. A colour built from the copies, with the journal class kept separate underneath, does that join for them; that is how we build ours, and we would rather agree on one methodology of construction than retire the names. We accept that "Bronze" is misleading; we already report it as unrealised open access, and "public" is a good description. The report's Appendix B allows both vocabularies to exist together, and we would run both.

Improve the proposal in four ways. Add "embargoed" and "unknown" to Access. Record the version of each copy, with "unknown" allowed. Record the source that reported each copy next to where it is hosted, so errors can be sent back to the right place. And, on section 4.4, show the per-copy information first, because whatever is shown first becomes the indicator people reuse.

Our commitments

1. **Licences**. We will help build a shared, maintained grouping of licences into open, restrictive and unknown. Our normalised licence strings are available for it, and we will adopt the result.
2. **Journals**. Our journal classification and its sources are available to the business model initiative the report recommends. One lesson from running it: it needs clearly identified venues underneath, and it must include book and proceedings publishers.
3. **Evidence**. OpenAlex and OpenAIRE both hold per-copy data, but nobody has compared them on the same works. We propose a joint comparison on a shared sample, copy by copy, starting with journals that have changed model.

4. **Running both vocabularies**. We will publish the colours and the report's labels side by side for as long as the community needs.

In short

The colours were always three inputs and a rule. We have the inputs. We have changed the rule before and can change it again. We would keep the names. What everyone needs is agreement on licences and better metadata from the sources. That is the work the report points to, and we would like to do it together.

References

Alperin, J. P., Kramer, B., van Eck, N. J., Farley, A., Church, A., Mitchell, D., Marcum, C., Takenaka, R., Harrison, M., & Gatti, R. (2026). Recommendations for revising the open access classification system in OpenAlex: Report of the OpenAlex Colors Working Group (Draft for community consultation). Zenodo. https://doi.org/10.5281/zenodo.21934115

Demes, K. (2026, September 17). Is it time to retire the OA colours? Tell us what you think. OpenAlex Blog. https://blog.openalex.org/is-it-time-to-retire-the-oa-colours-tell-us-what-you-think/

European Commission, Directorate-General for Research and Innovation. (2021). Monitoring the open access policy of Horizon 2020. Publications Office of the European Union. https://data.europa.eu/doi/10.2777/268348

Grypari, I., Pispiringas, L., & Manola, N. (2024). National Open Access Monitor, Ireland: Report. Zenodo. https://doi.org/10.5281/zenodo.10837317

OpenAIRE. (n.d.). OpenAIRE Graph documentation. https://graph.openaire.eu/docs/

infrastructure, OpenAIRE Graph, Research Monitoring