



Monitoring-Quality Assurance Processor-API

A tool to support CGIAR Quality Assurance process for peer-reviewed publications

May 19th, 2021



Monitoring, Evaluation and Learning (MEL)

The Monitoring, Evaluation and Learning (MEL) Team aims to improve the impact of research organizations through four areas of expertise: Monitoring & Evaluation, Knowledge Management, Data Curation and Research Software Development.

The team's main tool, the MEL Platform, centralizes the collection, visualization, and use of data for more informed decision-making and research impact of organizations.

AUTHORS

Valentina De Col¹

CO-AUTHORS

Sara Jani¹, Enrico Bonaiuti¹

SUGGESTED CITATION

Valentina De Col, Sara Jani, Enrico Bonaiuti (2021). Monitoring-Quality Assurance Processor-API - A tool to support CGIAR Quality Assurance process for peer-reviewed publications. International Center for Agriculture Research in the Dry Areas (ICARDA).

HANDLE

<https://hdl.handle.net/20.500.11766/13115>

DISCLAIMER



This document is licensed for use under the Creative Commons Attribution-ShareAlike (CC-BY-SA) 4.0 International License. To view this license, visit <http://creativecommons.org/licenses/by-sa/4.0/>.

Unless otherwise noted, you are free to copy, duplicate, or reproduce and distribute, display, or transmit any part of this publication or portions thereof without permission, and to make translations, adaptations, or other derivative works under the following conditions:

ATTRIBUTION. The work must be attributed, but not in any way that suggests endorsement by the publisher or the author(s).

SHARE ALIKE. If this work is altered, transformed, or built upon, the resulting work must be distributed only under the same or similar license to this one.

¹ The International Center for Agricultural Research in the Dry Areas (ICARDA).

Version	Date	Originator(s)	Reviewer(s)	Description
1	04/05/2021	Valentina De Col	Sara Jani, Enrico Bonaiuti	Structure and content
2	09/05/2021	Valentina De Col	Enrico Bonaiuti	Content
3	11/05/2021	Valentina De Col	Julien Colomer	Content, grammar and style corrections
4	19/05/2021	Valentina De Col	Enrico Bonaiuti	Content and style corrections

Index

Monitoring, Evaluation and Learning (MEL)	2
What is the Monitoring-Quality Assurance Processor-API tool?	5
Why use the M-QAP-API tool?	6
Advantages and disadvantages of the M-QAP-API tool	6
How to interpret the responses from the M-QAP-API tool	6
A. ISI	7
A. ISI troubleshooting	7
A1. Invalid DOI	7
A2. DOI/Journal article is peer-reviewed but not indexed under the WoS Core Collection™	7
A3. DOI/Journal article not yet indexed in the WoS Core Collection™	9
A4. Journal article is indexed in the WoS Core Collection™ but not linked to its DOI	9
B. OA	10
B. OA troubleshooting	11
B1. Invalid DOI	11
B2. DOI/Journal article recently uploaded in a repository	11
B3. DOI registered by an Agency that is not CrossRef or DataCite	11
Additional notes	13
Recommendations	14
Useful links	14
Acknowledgements	14

A brief introduction

What is the Monitoring-Quality Assurance Processor-API tool?

The [Monitoring, Evaluation, and Learning Quality Assurance Processor](#) (M-QAP) is a publications metadata extractor that employs Application Programming Interfaces (APIs) from [Web of Science™](#) (WoS), [Scopus®](#), [Unpaywall](#), [Altmetric](#), and F.A.I.R metric from [GARDIAN](#).

The tool, for short M-QAP-API, is designed to support CGIAR results reporting, including the [CGIAR Results Dashboard](#) and ensure that publications with a Digital Object Identifier (DOI) are validated against the above-mentioned databases.

The M-QAP-API has been designed by the [Monitoring, Evaluation and Learning](#) (MEL) team at the [International Center for Agricultural Research in the Dry Areas](#) (ICARDA) with the financial support of the CGIAR System Organisation (SO) and it has been piloted with the collaboration of the [Alliance of Bioversity International and CIAT](#) to support existing processes in [CLARISA](#) and other Management Information System (MIS) platforms such as MEL and [MARLO](#).

In more detail (Fig. 1), for the quality assessment of peer-reviewed publications, the M-QAP-API retrieves data and metadata from:

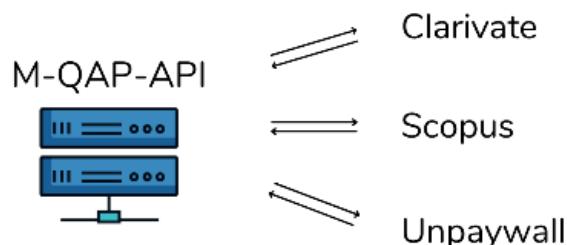


Figure 1. Simplified representation of M-QAP-API: the tool queries three different databases, namely WoS, Scopus, and Unpaywall and extracts specific information. Image designed in Canva.

- 1. WoS API**, powered by Clarivate, returns the ISI (formerly known as the Institute for Scientific Information™) status of each publication entered through the M-QAP-API to confirm if the publication is indexed or not in the [WoS Core Collection™](#). It can also return other metadata (e.g., the title of the publication, authors, year, journal name, volume, issue, pages).
- 2. Scopus API**, powered by Elsevier Developers®, returns if the publication is in the Scopus citation database and it can retrieve a set of metadata as well (e.g., the title of the publication, authors, year, journal name, volume, issue, pages).

3. **Unpaywall API** returns the response about the publication access status. If Open Access (OA), it can also provide the link where the publication is available.

Why use the M-QAP-API tool?

The M-QAP-API was developed to validate in an accurate, reliable, and automated way the assessment of peer-reviewed publications submitted by CGIAR entities as part of their annual reporting. The use of this tool allows saving time and resources. Previously the assessment was performed via a manual check by different individuals using non-standard ways of verifying the ISI and OA status. Moreover, the use of a benchmark academic database of reference such as WoS allows a consistent, reliable, replicable retrieval of data which bring values to the overall process and reduce errors and misinterpretations.

Advantages and disadvantages of the M-QAP-API tool

1. It automatizes the validation process and analyses thousands of publications at once;
2. It only requires the Digital Object Identifier (DOI) as an input;
3. It supports the consistent and reproducible collection of data;
4. It allows integration in MIS platforms and responses in real-time;
5. It relies on a paid account for WoS APIs, whereas SCOPUS, Unpaywall, Altmetric, GARDIAN are free;
6. It is not completely free from errors: a part of the results ("DOIs not found") requires validation via a manual check and/or contact the WoS/Unpaywall technical support for further investigation.

How to interpret the responses from the M-QAP-API tool

The M-QAP-API is an automated process that requires only the DOI input and supports the collection of data of crucial indicators for reporting on scientific publications, such as peer-review, ISI, and OA. As introduced in the previous section, the tool is not completely free from error and a subset of results requires validation via a manual check. Below a list of possible scenarios, grouped into the two main indicators: ISI and OA.

A. ISI

Response	Explanation	It means that that publication
Yes	The publication is included in the WoS Core Collection™ and all the metadata are retrieved from the WoS database.	- is peer-reviewed; and - is ISI, meaning indexed in the WoS Core Collection™.
No	The publication is not included in the WoS Core Collection™.	- has an invalid DOI; - is peer-reviewed, but not ISI; or - is not peer-reviewed. For any of these cases, a manual check to validate the result is needed (see troubleshooting paragraph).

A. ISI troubleshooting

An ISI status result of 'No' is caused by the following:

- A1. Invalid DOI;
- A2. DOI/Journal article is peer-reviewed but not indexed in the WoS Core Collection™;
- A3. DOI/Journal article not yet indexed in the WoS Core Collection™;
- A4. Journal article is indexed in the WoS Core Collection™ but not linked to its DOI.

The suggested steps to investigate and hopefully resolve a response appearing as 'No' are the following:

A1. Invalid DOI

Check if the DOI is found online. If not, visit the DOI® System website (doi.org) and submit a request via the text interface to notify the issue. You may also contact the publisher directly to inform about the issue.

A2. DOI/Journal article is peer-reviewed but not indexed under the WoS Core Collection™

If the DOI is valid:

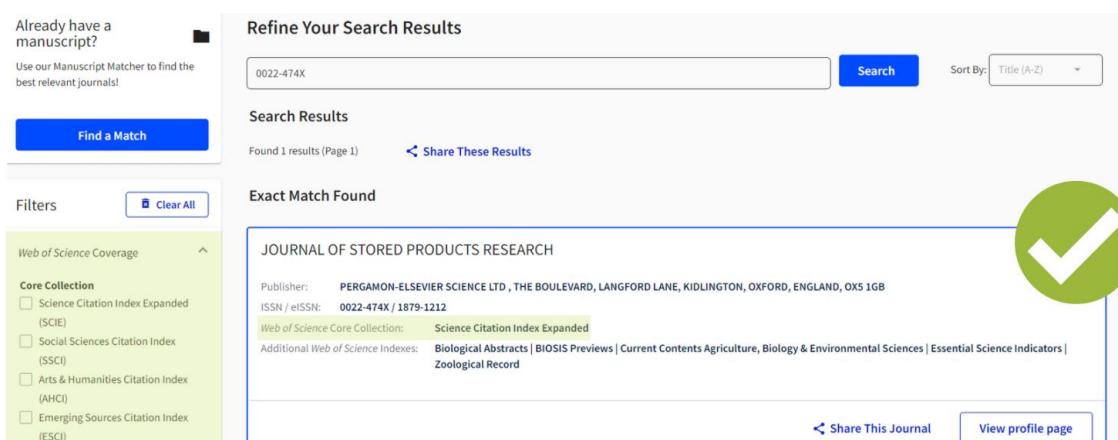
A2.1 Copy the Journal ISSN into the WoS Master Journal List (MJL) (mjl.clarivate.com) and check if the Journal is included in the WoS Core Collection™, in at least one of the indexes reported in figure 2.

Core Collection

- Science Citation Index Expanded (SCIE)
- Social Sciences Citation Index (SSCI)
- Arts & Humanities Citation Index (AHCI)
- Emerging Sources Citation Index (ESCI)

Figure 2. Indexes composing the WoS Core Collection™.

To confirm this, look at the result, as in the examples below (Fig. 3 and 4).



Already have a manuscript? Use our Manuscript Matcher to find the best relevant journals!

Find a Match

Refine Your Search Results

0022-474X **Search** Sort By: Title (A-Z)

Search Results
Found 1 results (Page 1) [Share These Results](#)

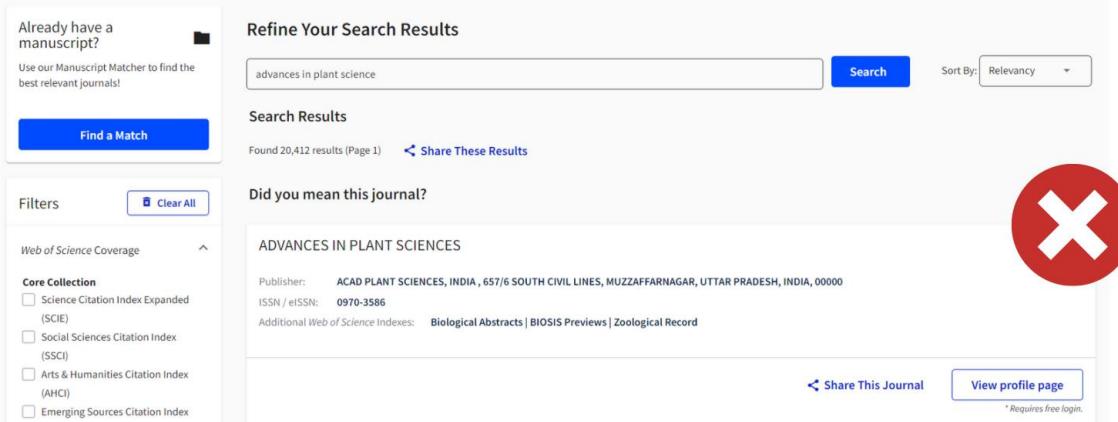
Exact Match Found

JOURNAL OF STORED PRODUCTS RESEARCH

Publisher: PERGAMON-ELSEVIER SCIENCE LTD, THE BOULEVARD, LANGFORD LANE, KIDLINGTON, OXFORD, ENGLAND, OX5 1GB
ISSN / eISSN: 0022-474X / 1879-1212
Web of Science Core Collection: Science Citation Index Expanded
Additional Web of Science Indexes: Biological Abstracts | BIOSIS Previews | Current Contents Agriculture, Biology & Environmental Sciences | Essential Science Indicators | Zoological Record

[Share This Journal](#) [View profile page](#)

Figure 3. The Journal is indexed in the WoS Science Citation Index Expanded, therefore it is ISI.



Already have a manuscript? Use our Manuscript Matcher to find the best relevant journals!

Find a Match

Refine Your Search Results

advances in plant science **Search** Sort By: Relevancy

Search Results
Found 20,412 results (Page 1) [Share These Results](#)

Did you mean this journal?

ADVANCES IN PLANT SCIENCES

Publisher: ACAD PLANT SCIENCES, INDIA, 657/6 SOUTH CIVIL LINES, MUZZAFFARNAGAR, UTTAR PRADESH, INDIA, 00000
ISSN / eISSN: 0970-3586
Additional Web of Science Indexes: Biological Abstracts | BIOSIS Previews | Zoological Record

[Share This Journal](#) [View profile page](#)

* Requires free login.

Figure 4. Journal not included in any indexes belonging to the WoS Core Collection™.

A3. DOI/Journal article not yet indexed in the WoS Core Collection™

Although ISI, there is a subset of publications returning a ‘No’ response because they are in the process of being indexed by the WoS. Falling into this category are, for example, publications that do not have the final publication date and the final assignment of a volume, issue, and page number ([Early Access](#)). The time for indexing depends on the Journal itself and on the WoS indexing service, and it could take up to 6 weeks². When the indexing is complete and the DOI is run again through the M-QAP-API, the response will change from ‘No’ into ‘Yes’.

In this scenario, after checking steps A1 and A2, it also better to:

A3.2 Visit the Journal website and check if the item is listed as ‘Article in Press’ or similar (Fig. 5). Conversely, in the case this metadata is available from the Journal side, the articles might be still under review and indexing by the WoS team.



Figure 5. An example of ‘Article in press’ with no mention to the final volume, issue, pages (as of May 2021).

Currently, there is no way for the user to know if the publication undergoing indexing. Therefore, while waiting for the article to be indexed, if the Journal itself is ISI (A2.1), the article can be reasonably considered as ISI as well.

A4. Journal article is indexed in the WoS Core Collection™ but not linked to its DOI

There is another subset of publications returning a ‘No’ response because the publications do not have a DOI appearing in the article/PDF (Fig. 6). In this case, it is recommended to:

A4.1 Ask the user of the WoS subscription to search the article by title to confirm it is ISI in the WoS Core Collection™. This will require that manual quality assurance is performed.

If not in the WoS Core Collection™, it is suggested to go back to step A1 and A2 and:

A4.2 Visit the Journal website and check if the DOI is not appearing within the publication (Fig.6).

² Personal communication with the WoS Technical Support.



Figure 6. An article with the DOI mentioned within the PDF (left side, highlighted in green) versus one without. Only the one on the left can be found via the API, whereas the one on the right can only be searched manually.

Note: Issues within the API and/or within the WoS database might occur. In these cases, contact the support service of the platform that is using M-QAP-API (for CLARISA: planningandreporting@cgiar.org) to direct your request to the relevant technical officer.

B. OA

Assessing the OA can be slightly more complex than the ISI due to different publishing policies from Journals and Publishers, and due to the internal rules set by the Unpaywall API. OA might also change over time, for instance, in the case of Green OA which is often subject to an embargo period. For example, an article with an embargo period of 6 months can be limited access if checked immediately after the publication, but open after 6 months. This requires performing periodic queries through the M-QAP- API tool.

To know more about each Journal publishing policies, please visit the Sherpa Romeo website (v2.sherpa.ac.uk/romeo). The ‘Help’ section on the website guides on understanding and reading Publisher policies (v2.sherpa.ac.uk/romeo/help.html).

Similarly to the ISI, the suggested steps to investigate and hopefully resolve a response appearing as ‘No’ are the following:

Response	Explanation	It means that that publication
Yes	The API finds the DOI in the publisher website or another repository. This is also the case of Green OA articles (an author version is available in a repository but not in the Journal webpage).	- is OA ³ .
No	The publications appear to be of limited access or not found.	- is not OA and will stay limited access; - is not OA at the moment, but it could become OA in the future (e.g., Green OA). For any of these cases, a manual check to validate the result is needed (see troubleshooting paragraph).

B. OA troubleshooting

When the result of the OA is ‘No’ there could be three main reasons:

- B1. Invalid DOI;
- B2. DOI/Journal article recently uploaded in a repository;
- B3. DOI registered by an Agency that is not CrossRef or DataCite.

B1. Invalid DOI

Check if the DOI is found online. If not, visit the DOI® System website (doi.org) and submit a request via the text interface to notify the issue. You may also contact the publisher directly to inform about the issue.

B2. DOI/Journal article recently uploaded in a repository

If the DOI is valid (step B1), check if:

B2.1 A downloadable version of the publication has been recently uploaded in a repository or about to be available on the Journal website (e.g., in the case of Gold OA). In this case, the API finds results 3 or 4 days from the time the repository record has been updated.

If confident that the result is OA after visiting the repository of the Journal website, the Quality Assessor responsible should overwrite the response of the API.

B3. DOI registered by an Agency that is not CrossRef or DataCite

To date, Unpaywall only considers DOI emitted from a specific Registration Agency called CrossRef⁴. Unpaywall does not include publications DOIs emitted by another Agency called [DataCite](#) since almost all publications from DataCite are OA. In the M-QAP-API we have

³ We have encountered a few cases where only the PDF of the cover page of the publication was available and the API returned that the publication was OA. It is the responsibility of repository managers to ensure that uploaded documents comply with Journal policies and are the correct ones.

⁴ <https://support.unpaywall.org/support/solutions/articles/44001900286-which-dois-does-unpaywall-cover->.

added DataCite to increase coverage. Users can manually check the DOI Registration Agency with the CrossRef REST API (github.com/CrossRef/rest-api-doc) by copying the following link into their browser:

- <https://api.crossref.org/works/{DOI}/agency>

If a DOI is registered by another Agency it does not mean it is not OA, therefore a manual check is needed.

Tip: users can check the Unpaywall API response for DOIs of their choice via:

- the **Unpaywall Simple Query Tool** at unpaywall.org/products/simple-query-tool. The user can submit a list of DOIs and will receive an email with the response of the API in a CSV, JSONL, or Excel (.xlsx) format including more metadata, such as the OA type (e.g., Green, Gold, Hybrid, Bronze), to assist with validation. For recently uploaded articles into a repository, the Simple Query Tool needs more than 4 days to return the updated value.
- the **Unpaywall REST API** at unpaywall.org/products/api#get-doi: adding the DOI in the URL bar of your browser and your email as following:
<https://api.unpaywall.org/v2/{DOI}?email={YOUR EMAIL}>

Example:

<https://api.unpaywall.org/v2/10.1093/femsec/fiaa069?email=user@cgmel.org>

By using the REST API, the user will receive the same response as the API. If only interested in validating if the Journal Article is OA, see directly the line “is_oa” (Fig. 7).

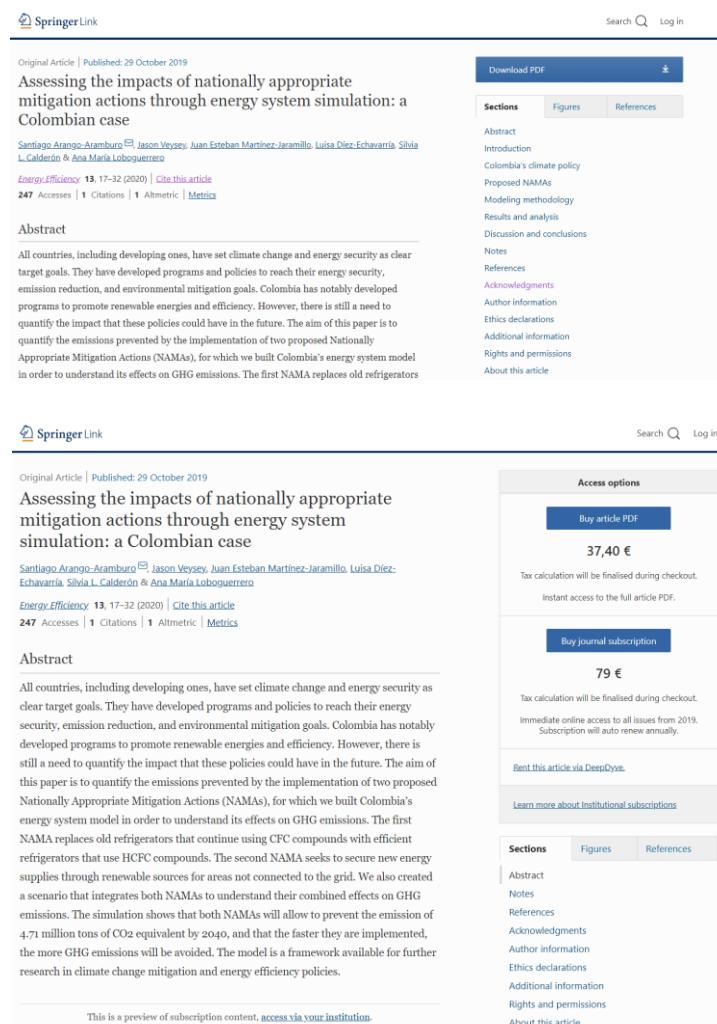
```
{
  "doi": "10.1016/j.gfs.2020.100466",
  "doi_url": "https://doi.org/10.1016/j.gfs.2020.100466",
  "title": "The processed food revolution in African food systems and the double burden of malnutrition",
  "genre": "journal-article",
  "is_paratext": false,
  "published_date": "2021-03-01",
  "year": 2021,
  "journal_name": "Global Food Security",
  "journal_issns": "2211-9124",
  "journal_issn_l": "2211-9124",
  "journal_is_oa": false,
  "journal_is_in_doa": false,
  "publisher": "Elsevier BV",
  "is_oa": true,
  "oa_status": "hybrid",
  "has_repository_copy": true,
  "best_oa_location": {
    "repository_institution": null,
    "host_type": "publisher",
    "url_for_pdf": null,
    "evidence": "open (via crossref license)",
    "url": "https://doi.org/10.1016/j.gfs.2020.100466",
    "pmh_id": null,
    "is_best": true,
    "url_for_landing_page": "https://doi.org/10.1016/j.gfs.2020.100466",
    "version": "publishedVersion",
    "updated": "2021-05-05T15:29:13.687866",
    "license": "cc-by-nc-nd",
    "oa_date": "2020-11-25",
    "endpoint_id": null
  }
}
```

Figure 7. An example of reply from the Unpaywall REST API. If only interested in validating the OA value Yes/No, see directly the result showing next to “is_oa” (highlighted in green in the image).

Additional notes



For the OA status, the country from which you are basing your search from can affect the response. An article that you see freely accessible could depend on the country you are searching from (and therefore your IP address) or the one where you set in your virtual private network (VPN)⁵ (Fig. 8). Access to articles could be, for instance, subsidized by universities, libraries, museums, foundations, societies or government agencies⁶. Run the Unpaywall Simple Query Tool to be sure about the result.



The figure consists of two vertically stacked screenshots of a journal article page from SpringerLink. Both screenshots show the same article details: "Assessing the impacts of nationally appropriate mitigation actions through energy system simulation: a Colombian case" by Santiago Arango-Aramburu, Jason Vessey, Juan Esteban Martínez-Jaramillo, Luisa Díez-Echavarría, Silvia J. Calderón, & Ana María Loboguerrero, published in Energy Efficiency 13, 17–32 (2020). The top screenshot shows a "Download PDF" button and a sidebar with tabs for Abstract, Sections, Figures, and References. The bottom screenshot shows a "Buy article PDF" button with a price of 37,40 €, and a sidebar with tabs for Access options, Buy journal subscription, and the same Abstract, Sections, Figures, and References tabs as the top version.

Figure 8. An example of the same journal article freely downloadable (upper image) or subject to a paywall or journal subscription (lower image).

⁵ A virtual private network (VPN) extends a private network across a public network and enables users to send and receive data across shared or public networks as if their computing devices were directly connected to the private network (en.wikipedia.org/wiki/Virtual_private_network).

⁶ https://en.wikipedia.org/wiki/Open_access#Subsidized_or_no-fee.

Recommendations

It is recommended to refresh the query over time to update the records, both for the WoS indexing issue and Unpaywall.

Useful links

- **A guide to Open Access policies (WoS)**
<https://clarivate.com/webofsciencegroup/article/a-researchers-complete-guide-to-open-access-papers/>
- **What is an OA license? (Unpaywall)**
<https://support.unpaywall.org/support/solutions/articles/44002063718>
- **What do the types of oa_status (green, gold, hybrid, and bronze) mean? (Unpaywall)**
<https://support.unpaywall.org/support/solutions/articles/44001777288-what-do-the-types-of-oa-status-green-gold-hybrid-and-bronze-mean->
- **Sherpa Romeo: a useful resource to the Journal's policies on OA**
<https://v2.sherpa.ac.uk/romeo/>

Acknowledgements

The development and improvements on the M-QAP-API tool were possible thanks to the financial support of the CGIAR System Organisation (SO) as well as for the WOS service subscription.

Enrico Bonaiuti from the CRP on Roots, Tubers and Bananas (RTB) provided the overall design of the process and support to the team.

Moayad Al-Najdawi from Code Land For System Development (CodeObia) developed the script while the Team at the Alliance of Bioversity International and CIAT tested the script in MISs and provided feedback.

Valentina De Col from the CGIAR Research Program (CRP) on Grain Legumes and Dryland Cereals (GLDC) provided a review of the data, improvement of API responses, and interaction with external services such as WoS and UnPaywall.

Sara Jani from The International Center for Agricultural Research in Dry Areas (ICARDA) provided an advisory role to the team in her capacity of Publications Repository Manager and metadata curator.