

# Bona Fide Researcher Verification

Peter Bolha

Jovan Simonoski

Alexandr Petrunin

Mihály Héder



*“Between 2014 and 2018, the researcher pool grew three times faster (13.7%) than the global population (4.6%). This translates into 8.854 million full-time equivalent (FTE) researchers.”*

UNESCO Science Report 2021

## Science is universal and favors collaboration, independent verification

- **Use case:** The user, when presented with some profile information about a researcher unknown to them, should get some indication whether said person is a “researcher in good faith” - or “bona fide researcher”
  - Reason: avoid friction - missed opportunities, unnecessary extra work - disambiguation
  - Reason: avoid fraud, made-up accounts, fake reviewers, fake authors
- **Use case:** conflict of interest detection
  - Funding Agencies, PhD boards, Open Positions

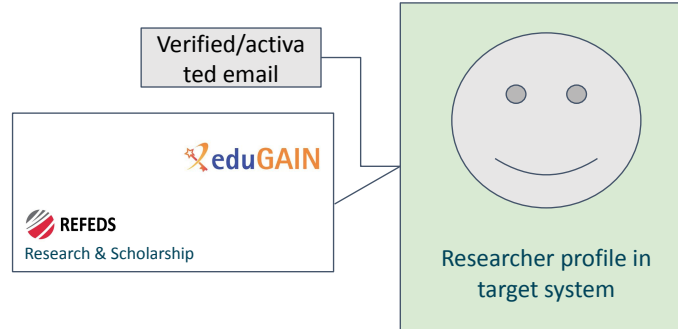
*“Oransky (2020) reports that over 600 papers were redacted due to the manipulation of the reviewers by the authors.”*

*Upcoming paper by the Incubator team*

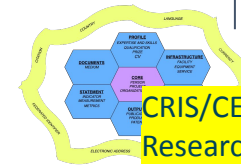
## Proof-of-Concept use case

- In the journal system the user receives - author information, pointers to reviewer, get some information

- Is the person real?
- Is the the person who they claim?
- Expertise in the field?
- No conflict of interest?



UNICEF ~ RoR

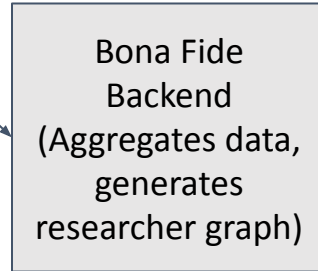


CRIS/CERIF - Common European  
Research Information Format

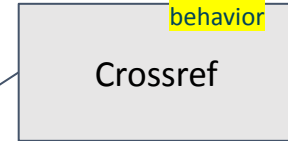
# Architecture



API



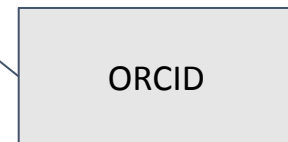
Overlap, but Different search behavior



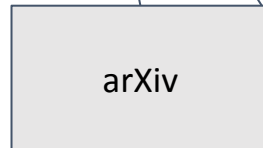
publications-> co-authorship, titles (field)



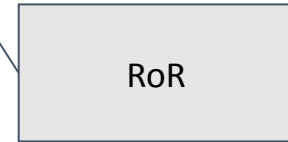
publications-> affiliations



ORCID itself email



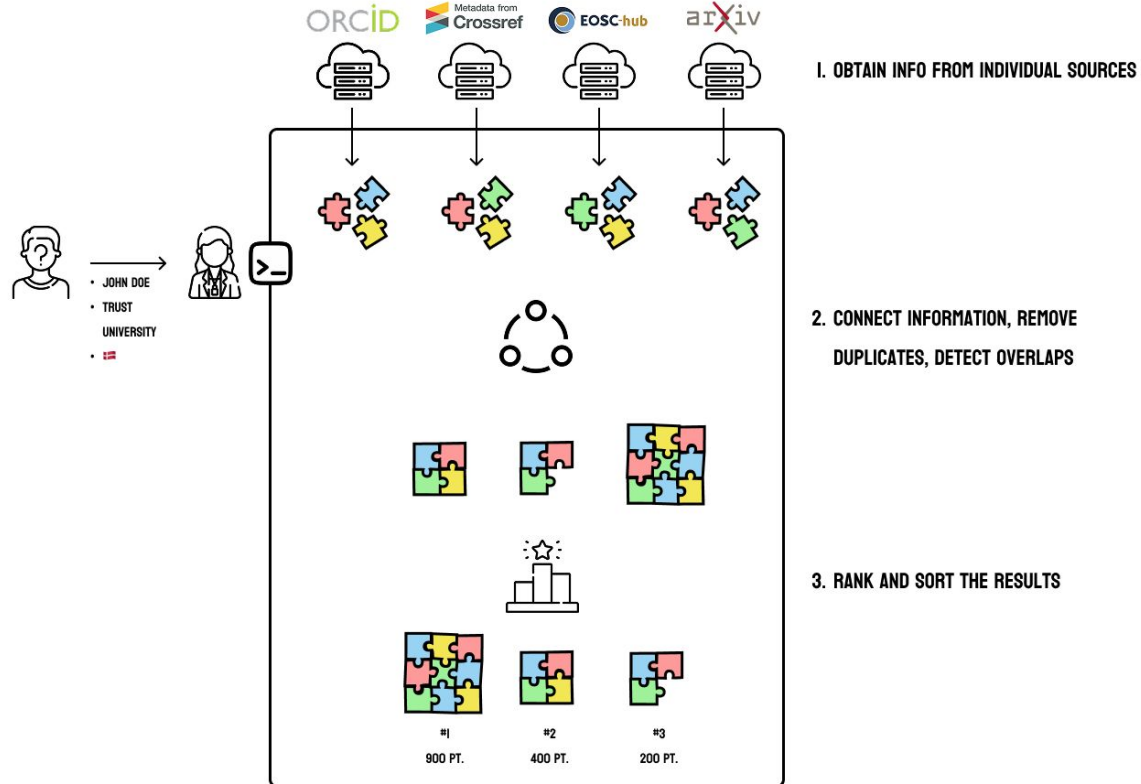
Field-specific, graph generation



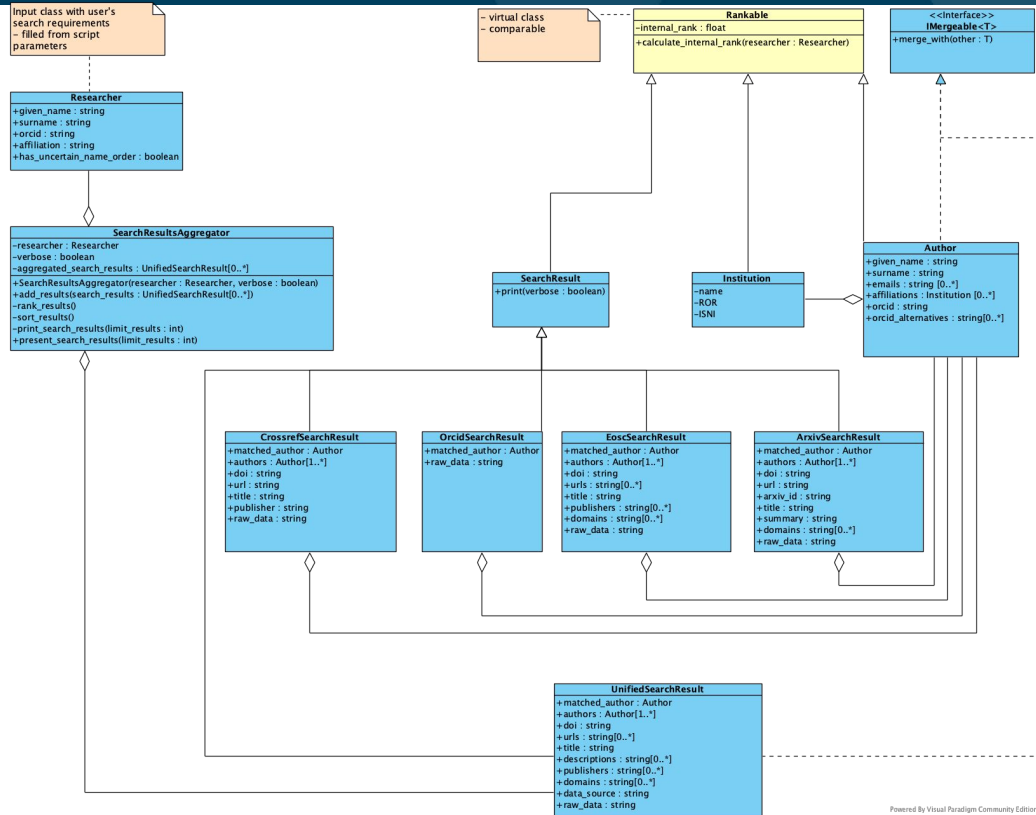
Institution (affiliation, email domain)

**Other candidates investigated:**  
SSRN  
DBLP  
IEEE Xplore (throttled/limited API)  
Sherlock, maigret

# Basic principles



# Data representation



## Names are hard

- transliterations, initials, name order, cultural differences
- **Solution:** fuzzy matching, flexible name order

## Merging

- We get the similar data in various paths authors, institutions, articles
- **Solution:** data aggregation with alternative values f

## Ranking

- more data = more trust ?, more precise matches = more trust ?
- some authors with many articles, affiliations etc. skew the results
- **Solution:** primary and secondary ranking - perfect matches & supporting info

# CLI Tool Demo - successful case

## Target: Mihály Héder

```
User attributes given:
=====
Given name: Mihály
Surname: Héder
Email: mihaly.heder@sztaki.hu
Uncertain name order
=====
```

```
#1 (INDEX)
Result score breakdown:
  Author total score: 340.0
  Perfectly matched attributes: 1
  Articles total score: 5596.0
  Number of articles: 11
  Average score of article: 508.73
Author: Héder (given name), Mihály (surname) [NAME MATCH - 200/200]
Affiliations:
  - Institution: Hungarian Academy of Sciences - identifiers: https://ror.org/02ks8qq67 (ROR), 0000000121494407 (ISNI)
  - Institution: Budapest University of Technology and Economics - identifiers: https://ror.org/02w42ss30 (ROR), 0000000121800451 (ISNI)
  - Institution: Library and Information Centre of the Hungarian Academy of Sciences - identifiers: https://ror.org/04ws47v52 (ROR), 0000000121789790 (ISNI)
ORCID: 0000-0002-9979-9101
Email: ?
Collected works:
=====
Title: Comparing institutional strategies for engineering ethics education in regional, national and European contexts
DOI: 10.5281/zenodo.14652136
URLS:
  - http://dx.doi.org/10.5281/zenodo.14652136
  - https://research.tue.nl/en/publications/6932c138-d4cc-455d-b61a-cfda7d5889b5
Publishers: ?
Domains: ?
=====
Title: A black market for upvotes and likes
DOI: 10.22503/inftars.xix.2019.4.2
URLS:
  - https://doi.org/10.22503/inftars.xix.2019.4.2
  - http://arxiv.org/abs/1803.07029
  - https://dx.doi.org/10.48550/arxiv.1803.07029
  - http://dblp.uni-trier.de/db/journals/corr/corr1803.html#abs-1803-07029
  - https://doi.org/https://doi.org/10.22503/inftars.XIX.2019.4.2
Publishers:
  - Információs Társadalom
Domains:
  - Social Sciences
=====
Title: A criticism of AI ethics guidelines
DOI: 10.22503/inftars.xx.2020.4.5
URLS:
  - https://doi.org/10.22503/inftars.xx.2020.4.5
```

# CLI Tool Demo - successful case

```
Organization Matches:
=====

1. Computer and Automation Research Institute
Match Score: 100%
Website(s): https://www.sztaki.hu/en

Score Breakdown:
Fully Qualified Domain Match      : 100/100 points
Base Domain Match                  :   0/80 points
Email in Website Subdomain        :   0/20 points
Website in Email Subdomain        :   0/20 points
DNS Similarity Bonus              :   0/15 points
WHOIS Bonus                        :   0/20 points
Crossref Bonus                    :   0/5  points

Crossref ID: 100007626
Crossref Data:
Name      : Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézet
Location  : Hungary
Works Count : 62
Also known as: MTA SZTAKI, Computer and Automation Research Institute, Hungarian Academy of Sciences, Számítástechnikai és Automatizálási Kutatóintézet, ...

DNS Comparison
No comparison available (same domain: sztaki.hu)

WHOIS Comparison
No comparison available (same domain: sztaki.hu)

-----
RDR researcher email domain verification was successful.
```

# CLI Tool Demo - problematic case

Target: John Smith



[Sign in / Register](#) English 



## Search

[ADVANCED SEARCH](#)

Showing 50 of 141811 results.

Items per page:

50 

Page 1 of 2837



ORCID ID	First Name	Last Name	Other Names	Affiliations
<a href="#">0000-0003-1660-3511</a>	John	Smith		Fairleigh Dickinson University, Flagburn Health Center, Nova Scotia Health Authority South Western Nova Scotia, University of Ottawa, University of York
<a href="#">0000-0002-9315-0678</a>	John	Smith		Fluent Technology, Sample invited position, Ulster Univeristy, Ulster University
<a href="#">0000-0002-4216-1107</a>	John	Smith		East Carolina University, Michigan State University, University of Central Missouri
<a href="#">0000-0001-5376-4926</a>	John	Smith		Bureau of Ocean Energy Management, Pacific OCS Region, OCS Energy Consultant, Pennsylvania State University
<a href="#">0000-0001-7793-0079</a>	John	Smith		Charles Sturt University - Wagga Wagga Campus, CottonInfo, New South Wales Department of Primary Industries Agriculture
<a href="#">0000-0002-2351-4982</a>	John	Smith		Stony Brook University, Virginia Commonwealth University
<a href="#">0000-0002-4409-653X</a>	John	Smith		American Music Therapy Association, Sam Houston State University, Sun Behavioral

 Help

# CLI Tool Demo - problematic case

## Target: John Smith

User attributes given:

```
=====
Given name: John
Surname: Smith
Email: john.smith@triangelescscientific.com
Fixed name order
=====
```

```
#1 (INDEX)
Result score breakdown:
  Author total score: 200.0
  Articles total score: 74200.0
  Number of articles: 301
  Average score of article: 246.51
Author: John (given name), Smith (surname) [NAME MATCH - 200/200]
?
ORCID: ?
Email: ?
Collected works:
=====
Title: Brian Macdonald-Milne. 2020. <i>Seeking Peace in the Pacific: The Story of Conflict and Christianity in the Central Solomon Islands</i>
DOI: 10.3366/swc.2022.0381
URLS:
  - https://doi.org/10.3366/swc.2022.0381
Publishers:
  - Edinburgh University Press
Domains: ?
=====
Title: John B. Smith Archive - Chronological Bibliography Bundle 4 (Technical Reports)
DOI: 10.7939/r33j3943d
URLS:
  - https://era.library.ualberta.ca/items/b363bb91-5e80-4b52-83c3-abbbe878dbc2
  - https://dx.doi.org/10.7939/r33j3943d
Publishers: ?
Domains: ?
=====
Title: John B. Smith Archive - Chronological Bibliography Bundle 3 (Bibliographies)
DOI: 10.7939/r39469
URLS:
  - https://era.library.ualberta.ca/items/a3d421eb-3e7b-4908-8d79-a4e8a87ac4b0
  - https://dx.doi.org/10.7939/r39469
Publishers: ?
Domains: ?
=====
Title: The SpineCor Brace in 2015: A Form of Chiropractic Child Abuse? A Report of Three Cases
DOI: 10.1055/s-0036-1583108
```

- CLI
  - running a script `verify_eduperson.py` with arguments
  - response presented in the terminal window as structured text
- API
  - querying the researcher info using JWT in HTTP request
  - hybrid response:
    - active polling of the status endpoint
    - passive awaiting at the callback url
  - response presented in JSON format
  - multiple researchers can be queried at the same time, concurrent processing

# Research Organization Registry - Email Lookup

```
https://api.ror.org/v2/organizations?query.advanced=  
locations.geonames_details.country_code:[COUNTRY_CODE] + AND +  
links.type:website + AND +  
links.value:[DOMAIN_RELATED_WORD]
```

Enter the email address: *mihaly.heder@sztaki.mta.hu*

Generated ROR Queries:

```
https://api.ror.org/v2/organizations?query.advanced=locations.geonames\_details.country\_code:HU%20AND%20links.type:website%20AND%20links.value:sztaki  
https://api.ror.org/v2/organizations?query.advanced=locations.geonames\_details.country\_code:HU%20AND%20links.type:website%20AND%20links.value:mta  
https://api.ror.org/v2/organizations?query.advanced=locations.geonames\_details.country\_code:HU%20AND%20links.type:website%20AND%20links.value:sztaki.mta
```

# Research Organization Registry - Email Lookup

## Organization Matches:

=====

### 1. Hungarian Academy of Sciences

Match Score: 100%

Website(s): <http://mta.hu/english/>

#### Score Breakdown:

Fully Qualified Domain Match	:	0/100 points
Base Domain Match	:	80/80 points
Email is Subdomain of Website	:	10/10 points
Website is Subdomain of Email	:	0/-50 points
Subdomain Mismatch Penalty	:	0/-50 points
DNS Bonus	:	8/30 points
WHOIS Bonus	:	0/30 points
Crossref Bonus	:	5/5 points

Crossref ID: 501100003825

#### Crossref Data:

Name : Magyar Tudományos Akadémia  
Location : Hungary  
Works Count : 6025  
Also known as: Hungarian Academy of Sciences, MTA

## DNS comparison

DNS Comparison: sztaki.mta.hu and mta.hu

-----

Domain Relationship Score: 25/100

#### Matching Name Servers:

- ns2.iif.hu

WHOIS Comparison: sztaki.mta.hu and mta.hu

-----

## WHOIS comparison

WHOIS Match Score: 0/100

Website of the organization

Score breakdown

Crossref Reference

# Research Organization Registry - Email Lookup

19. Institute for Computer Science and Control

Match Score: 42%

Website(s): <https://www.sztaki.hu/?en>

Score Breakdown:

Fully Qualified Domain Match	:	0/100	points
Base Domain Match	:	0/80	points
Email in Website Subdomain	:	0/20	points
Website in Email Subdomain	:	20/20	points
DNS Similarity Bonus	:	12/15	points
Crossref Bonus	:	0/5	points

Crossref ID: 100007626

Crossref Data:

Name : Magyar Tudományos Akadémia Számítástechnikai és Automatizálási Kutatóintézet

Location : Hungary

Works Count : 61

Also known as: MTA SZTAKI, Computer and Automation Research Institute, Hungarian Academy of Sciences, Számítástechnikai és Automatizálási Kutatóintézet, ...

DNS Comparison: [sztaki.mta.hu](#) and [sztaki.hu](#)

-----  
Domain Relationship Score: 85/100

Matching Name Servers:

- ns4.sztaki.hu

Matching A Records (IPs):

- 195.111.1.80

Matching MX Records:

- mail.sztaki.hu

Related SOA emails found

Email SOA: [hostmaster.sztaki.hu](mailto:hostmaster.sztaki.hu)

Website SOA: [hostmaster.sztaki.hu](mailto:hostmaster.sztaki.hu)

# arXiv - Graph of connections between researchers

Example J

1

**Yan Ai**  
Example Submission

Submission

Activity Log Library

**Workflow**

- Submission
- Review
- Copyediting
- Production

**Publication**

- Title & Abstract
- Contributors
- Metadata
- JATS XML
- Galleys
- Permissions & Disclosure
- Issue
- Bona Fide

### BONA FIDE: AUTHORS AND REVIEWERS

Reload All Get JSON Data

#### Authors

**Yan Ai**  
email: example@email.com

Hide Details

Reload Get relation graph

**Domain (214)**

1	Domain Match (80)	Email is Subdomain of \ Total Score (107)	WHOIS Bonus (16)
---	-------------------	---	------------------

**Yan Ai**

**Name (200)**  
Perfect Match (200)

```
Loading edges from collaboration_network_MihalyHeder_depth_2.csv...
Loading edges from collaboration_network_YoshuaBengio_depth_1.csv...

Total nodes: 359
Total edges: 460

Coauthor Connection Finder
Enter the first researcher's name: Dongwoo Kim
Enter the second researcher's name: Yoshua Bengio

Connection found between 'Dongwoo Kim' and 'Yoshua Bengio':
  Dongwoo Kim
    Sungsoo Ahn
      Yoshua Bengio

Degrees of separation: 2
```

# OJS 3.4 plugin: Previous work

test

19 / alex petrunin et al. / test12

Activity Log Library

Workflow Publication **Trust Score Overview**

New tab in submission

Submission Review Copyediting Production Help

Round 1 New Review Round

Round 1 Status  
Awaiting responses from reviewers.

Bona Fide Journal

23 / Author 1 et al. / Bona Fide Demo submission

Activity Log Library

Workflow Publication **Trust Score Overview**

Authors

Name	Email	Trust Score
Author 1	<a href="mailto:author-1@mail.com">author-1@mail.com</a>	30
Author 2	<a href="mailto:author-2@mail.com">author-2@mail.com</a>	89

Fetching trust data from backend

Reviewers

Name	Email	Trust Score
reviewer	<a href="mailto:reviewer@mail.com">reviewer@mail.com</a>	68
new reviewer	<a href="mailto:newreviewer1@gmail.com">newreviewer1@gmail.com</a>	42

Generating PDF trust summary

Export PDF

test

Submissions

My Queue (17) Unassigned (5) Active (22) Archived (1) Help

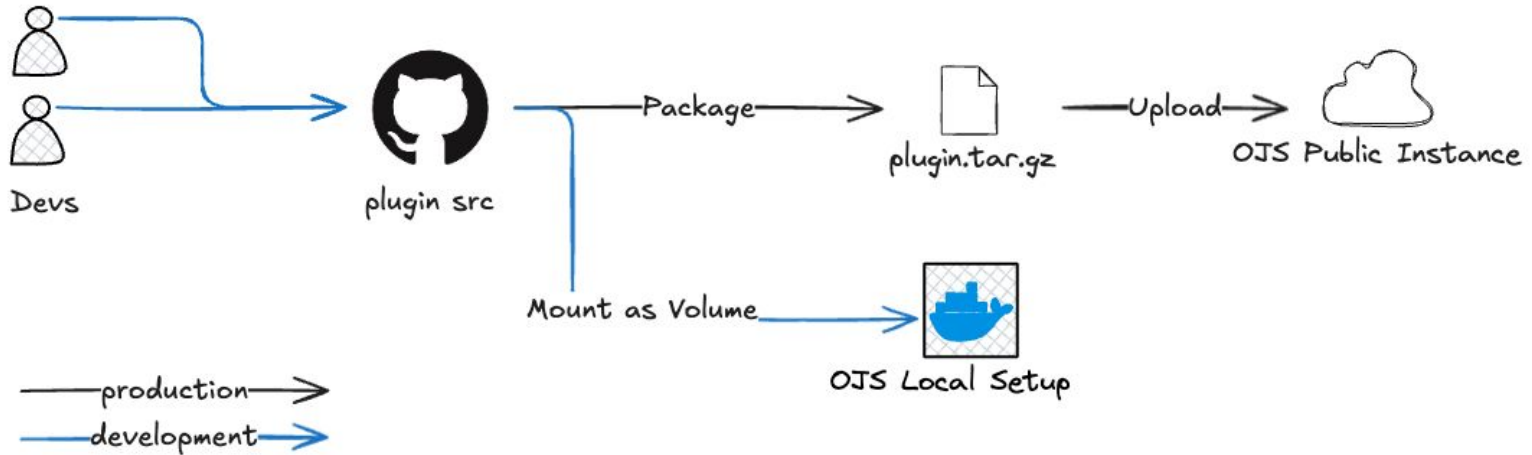
My Assigned

ID	Author	Status	Actions
23	Author 1 et al. Bona Fide Demo submission	No editor has been assigned to this submission. <a href="#">Assign Editor</a>	Submission View
13	Family Name test7	Waiting for reviewers to be assigned.	Review View
22	author retrieval check	No editor has been assigned to this submission. <a href="#">Assign Editor</a>	Submission View
20	Family Name et al. test13	No editor has been assigned to this submission. <a href="#">Assign Editor</a>	Submission View
19	alex petrunin et al. test12		Review View
18	test11	Incomplete	Incomplete View
17	test10	Incomplete	Incomplete View
16	test10	Incomplete	Incomplete View

- **OJS Upgrade 3.4 → 3.5: Vue3 frontend framework, features deprecation**
- Enhanced DX, Artefact generation workflow
- New features
  - Subscores UI
  - Graph representation
  - JSON reports

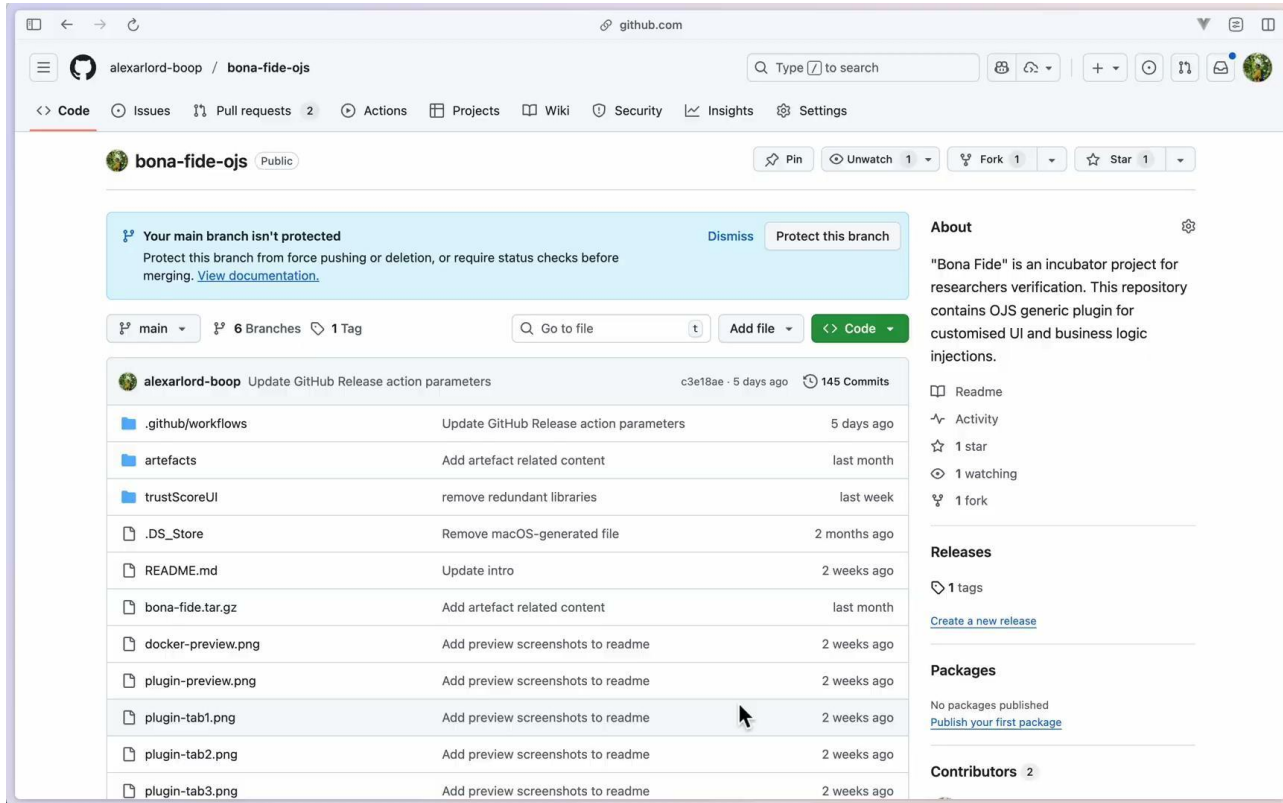
Source repository: <https://github.com/alexarlord-boop/bona-fide-ojs>

# OJS 3.5 plugin: Development



Source repository: <https://github.com/alexarlord-boop/bona-fide-ojs>

# OJS 3.5 plugin: Deployment



The screenshot shows the GitHub interface for the repository 'alexarlord-boop / bona-fide-ojs'. The repository is public and has 1 star and 1 fork. A notification banner indicates that the main branch is not protected. The repository contains several files and folders, including workflows, artefacts, and source code files. The 'plugin-tab1.png' file is highlighted by a mouse cursor.

**alexarlord-boop** / **bona-fide-ojs** (Public)

Your main branch isn't protected  
Protect this branch from force pushing or deletion, or require status checks before merging. [View documentation.](#)

main 6 Branches 1 Tag

alexarlord-boop Update GitHub Release action parameters c3e18ae · 5 days ago 145 Commits

File/Folder	Description	Last Commit
<code>.github/workflows</code>	Update GitHub Release action parameters	5 days ago
<code>artefacts</code>	Add artefact related content	last month
<code>trustScoreUI</code>	remove redundant libraries	last week
<code>.DS_Store</code>	Remove macOS-generated file	2 months ago
<code>README.md</code>	Update intro	2 weeks ago
<code>bona-fide.tar.gz</code>	Add artefact related content	last month
<code>docker-preview.png</code>	Add preview screenshots to readme	2 weeks ago
<code>plugin-preview.png</code>	Add preview screenshots to readme	2 weeks ago
<code>plugin-tab1.png</code>	Add preview screenshots to readme	2 weeks ago
<code>plugin-tab2.png</code>	Add preview screenshots to readme	2 weeks ago
<code>plugin-tab3.png</code>	Add preview screenshots to readme	2 weeks ago

**About**

"Bona Fide" is an incubator project for researchers verification. This repository contains OJS generic plugin for customised UI and business logic injections.

**Releases**

1 tags

[Create a new release](#)

**Packages**

No packages published

[Publish your first package](#)

**Contributors** 2

# OJS 3.5 plugin: Scoring Feature

The screenshot displays a web browser window with the URL `http://localhost:8081/TJ/dashboard/editorial?currentViewId=initial-review&workFlowSubmissionId=1&workFlowMenuKey=overview`. The page title is "Test Journal". The main content area shows a navigation menu on the left with options: Workflow, Publication, Bona Fide, Authors and Reviewers (selected), Report Preview, and Settings. The main content area is titled "BONA FIDE: AUTHORS AND REVIEWERS" and contains two sections: "Authors" and "Reviewers". The "Authors" section shows "Loading Authors..." and the "Reviewers" section shows "No reviewers found". There are also buttons for "Reload All" and "Get JSON Data" in the "BONA FIDE: AUTHORS AND REVIEWERS" section.

Test Journal

1

**Héder**  
Test  
Submission

Activity Log Library

Workflow  
Publication  
Bona Fide  
Authors and Reviewers  
Report Preview  
Settings

**BONA FIDE: AUTHORS AND REVIEWERS**

Reload All Get JSON Data

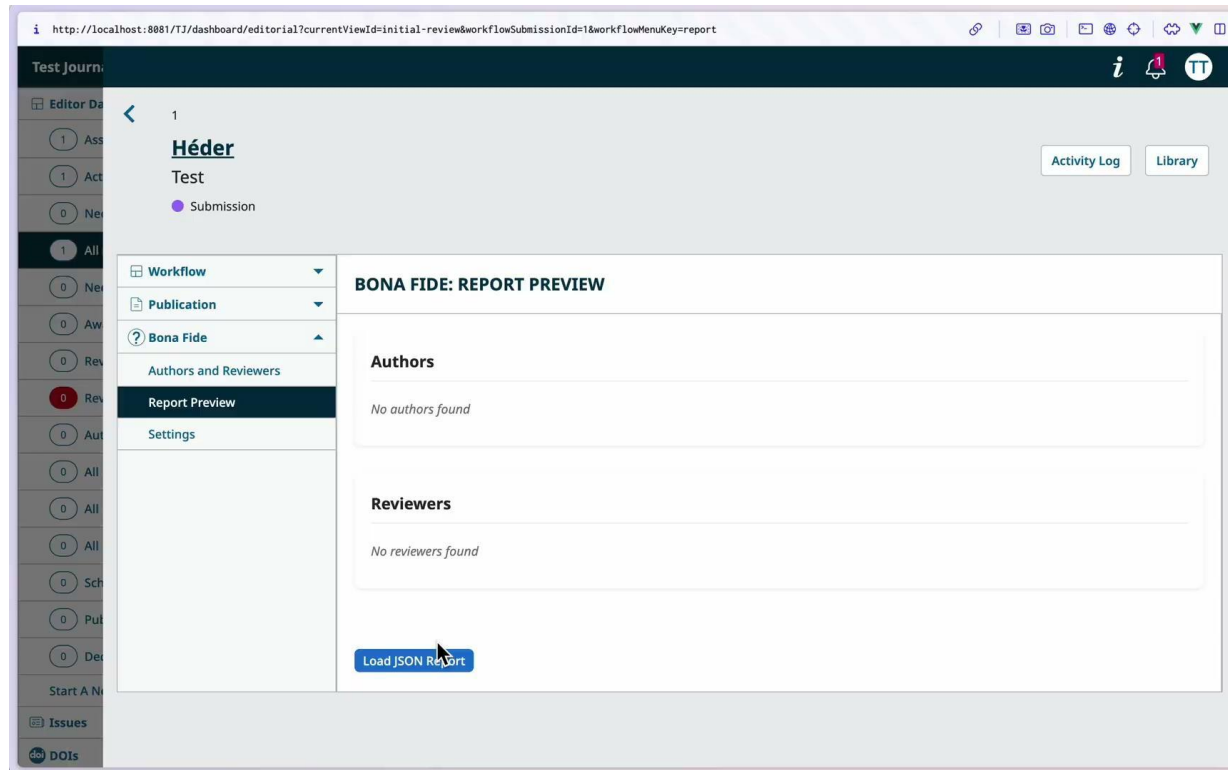
**Authors**

Loading Authors...

**Reviewers**

No reviewers found

# OJS 3.5 plugin: Report Feature

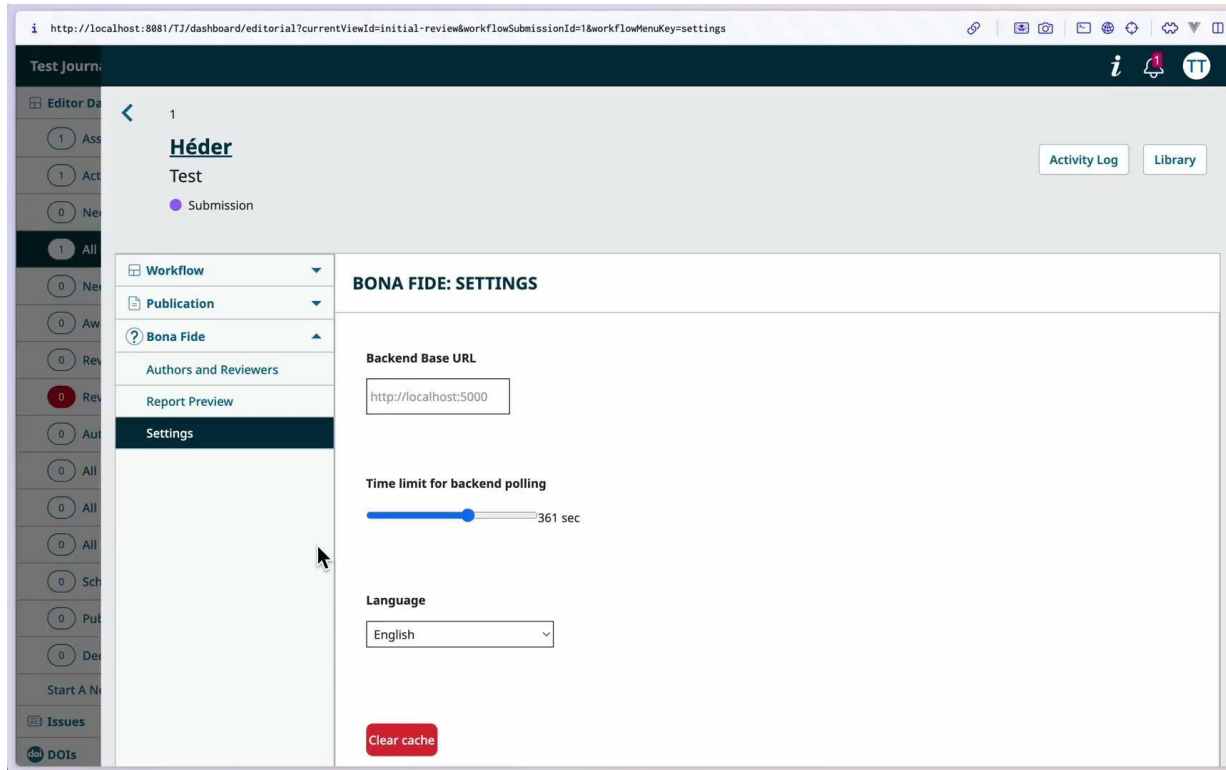


The screenshot shows a web browser window displaying the OJS 3.5 plugin report feature. The browser's address bar shows the URL: `http://localhost:8881/TJ/dashboard/editorial?currentViewId=initial-review&workflowSubmissionId=1&workflowMenuKey=report`. The page title is "Test Journal".

The interface includes a sidebar on the left with a navigation menu. The main content area displays the following elements:

- Editorial Dashboard:** Shows a list of items with counts: 1 Ass, 1 Act, 0 Ne, 1 All, 0 Ne, 0 Aw, 0 Rev, 0 Rev, 0 Aut, 0 All, 0 All, 0 All, 0 Sch, 0 Put, 0 De, Start A N.
- Editorial Information:** Shows "1" and "Héder Test" with a "Submission" indicator.
- Activity Log and Library:** Two buttons are visible in the top right.
- Workflow and Publication Menus:** A dropdown menu is open, showing options: "Workflow", "Publication", "Bona Fide", "Authors and Reviewers", "Report Preview" (highlighted), and "Settings".
- BONA FIDE: REPORT PREVIEW:** The main content area displays the title "BONA FIDE: REPORT PREVIEW" and two sections: "Authors" and "Reviewers". Both sections show "No authors found" and "No reviewers found" respectively.
- Load JSON Report:** A blue button is located at the bottom of the report preview area.

# OJS 3.5 plugin: Settings Feature



The screenshot shows a web browser window displaying the OJS 3.5 plugin settings page. The browser's address bar shows the URL: `http://localhost:8881/TJ/dashboard/editorial?currentViewId=initial-review&workflowSubmissionId=1&workflowMenuKey=settings`. The page title is "Test Journal". The main content area is titled "Héder Test" and includes a "Submission" indicator. A sidebar on the left contains a navigation menu with options: Workflow, Publication, Bona Fide (selected), Authors and Reviewers, Report Preview, and Settings. The "BONA FIDE: SETTINGS" section contains the following configuration options:

- Backend Base URL:** A text input field containing `http://localhost:5000`.
- Time limit for backend polling:** A slider control set to 361 seconds.
- Language:** A dropdown menu currently set to "English".

A red "Clear cache" button is located at the bottom of the settings panel. The browser's top toolbar includes icons for information, notifications, and user profile.

- Wording: trust/assurance about the information (not the person!)
  - multi-dimensional as per CoARA principles
- Limitation: does not work for fresh/new academics
  - Not better than eduGAIN+R&S
- Future work
  - Check the wording/language on the UX
  - Further data sources
  - OJS- revisit federated login, data harmonization
  - Testing
- Links
  - OJS frontend: <https://github.com/alexarlord-boop/bona-fide-ojs>
  - RoR script: <https://github.com/JovanSimonoski/GEANT-TIM-BonaFide>
  - ArXive script: <https://github.com/JovanSimonoski/GEANT-TIM-BonaFide>

Please get involved:  
mihaly.heder@sztaki.hu