Using ORCID to track awardees' progress

Gary McDowell, PhD   10/06/2022   HRA
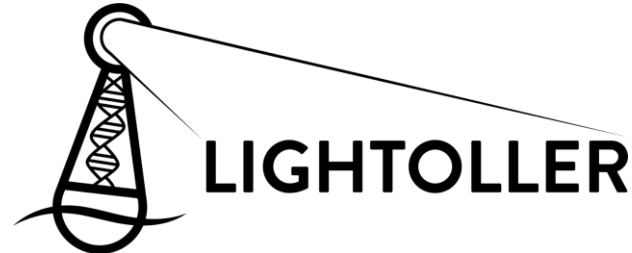
info@lightoller.org

What is ORCID?

ORCID
stands for
Open **R**esearcher and **C**ontributor ID

The ORCID iD: a unique, persistent identifier free of charge to researchers
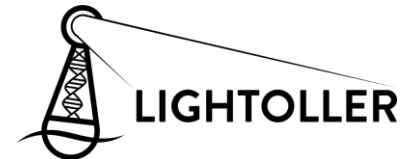
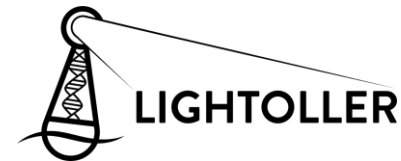An **ORCID record** connected to the ORCID iD, and

A set of **Application Programming Interfaces (APIs)**, as well as the services and support of communities of practice that enable interoperability between an ORCID record and member organizations so researchers can choose to allow connection of their iD with their affiliations and contributions

See: https://info.orcid.org/what-is-orcid/

LIGHTOLLER

Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
    - it's a **unique**, permanent ID for researchers

LIGHTOLLER

Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
  - it's a **unique**, permanent ID for researchers
  - it has a record closely approximating an "academic CV"

LIGHTOLLER

Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
    - it's a **unique**, permanent ID for researchers
    - it has a record closely approximating an "academic CV"
    - users can update and *verify* **their own** information
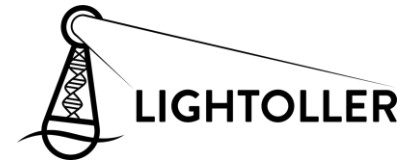
LIGHTOLLER

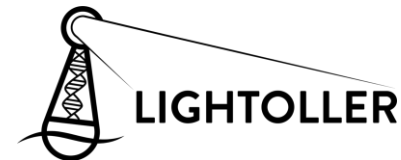# Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
    - it's a **unique**, permanent ID for researchers
    - it has a record closely approximating an "academic CV"
    - users can update and *verify* **their own** information
    - it is increasing in uptake/mandates are coming down the pipeline.

LIGHTOLLER

Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
    - it's a **unique**, permanent ID for researchers
    - it has a record closely approximating an "academic CV"
    - users can update and *verify* **their own** information
    - it is increasing in uptake/mandates are coming down the pipeline.

- It acts as a database of affiliations, and subsequent achievements, likely of interest to those measuring career/research outcomes.
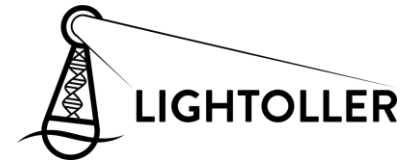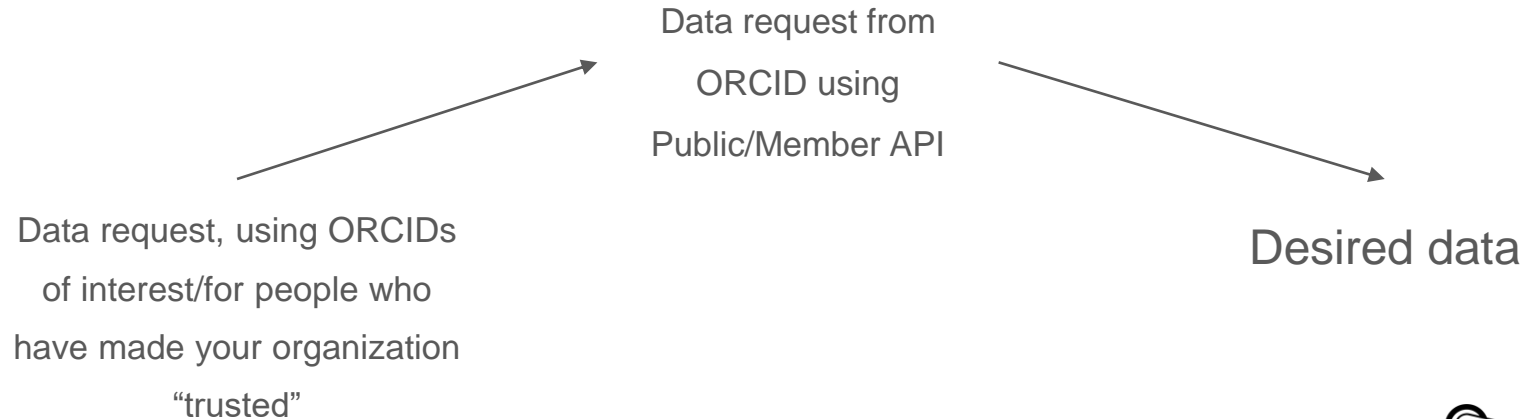
LIGHTOLLER

Why ORCID is suitable for outcomes tracking

- ORCID is useful because:
  - it's a **unique**, permanent ID for researchers
  - it has a record closely approximating an "academic CV"
  - users can update and *verify* **their own** information
  - it is increasing in uptake/mandates are coming down the pipeline.

- It acts as a database of affiliations, and subsequent achievements, of interest to those measuring career/research outcomes.

- Member organizations can push data to records of users who authorize them, and also retrieve data *made available to them*.

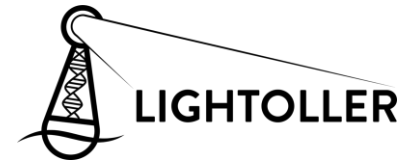LIGHTOLLER

Workflow to get ORCID data

Anyone can pull data from the public API; ORCID Member organizations have the ability to pull, as well as push, data from/to ORCID using the Member API

Data request from
ORCID using
Public/Member API

Data request, using ORCIDs
of interest/for people who
have made your organization
"trusted"

Desired data

LIGHTOLLER

Data potentially available in ORCID

Information recorded:

- Person
  - **Names** (except other credit names)
  - **Email**
  - **Bio**
  - Websites/Social Media
  - Country
  - Keywords
  - *External IDs*
- Note: Demographics Not Included

LIGHTOLLER

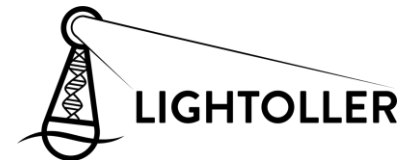# Data potentially available in ORCID

Information recorded:

- Activities
  - Education
  - Employment
  - Funding
  - Works
  - Distinctions, Invited Positions, Qualifications, Service, Membership
  - *(Journal Article) Peer Review; Research Resources*
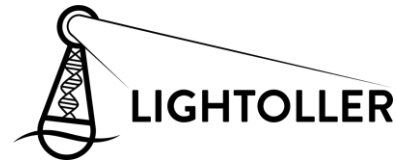
LIGHTOLLER

Using ORCID to look at a small dataset

- Working with Dr. Alberto Roca, looking at career outcomes for people in the DiverseScholar Doctoral Directory
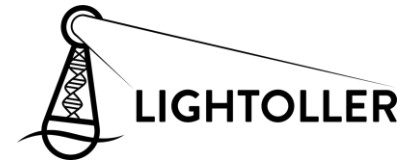
LIGHTOLLER

Using ORCID to look at a small dataset

- Working with Dr. Alberto Roca, looking at career outcomes for people in the DiverseScholar Doctoral Directory

- Database of grads and postdocs from underrepresented groups that includes demographic information and articulation of career intention

LIGHTOLLER

Using ORCID to look at a small dataset

- Working with Dr. Alberto Roca, looking at career outcomes for people in the DiverseScholar Doctoral Directory

- Database of grads and postdocs from underrepresented groups that includes demographic information and articulation of career intention

- I searched for ORCIDs for everyone in database; also LinkedIn search; where possible also bringing in "citation" e.g. lab website

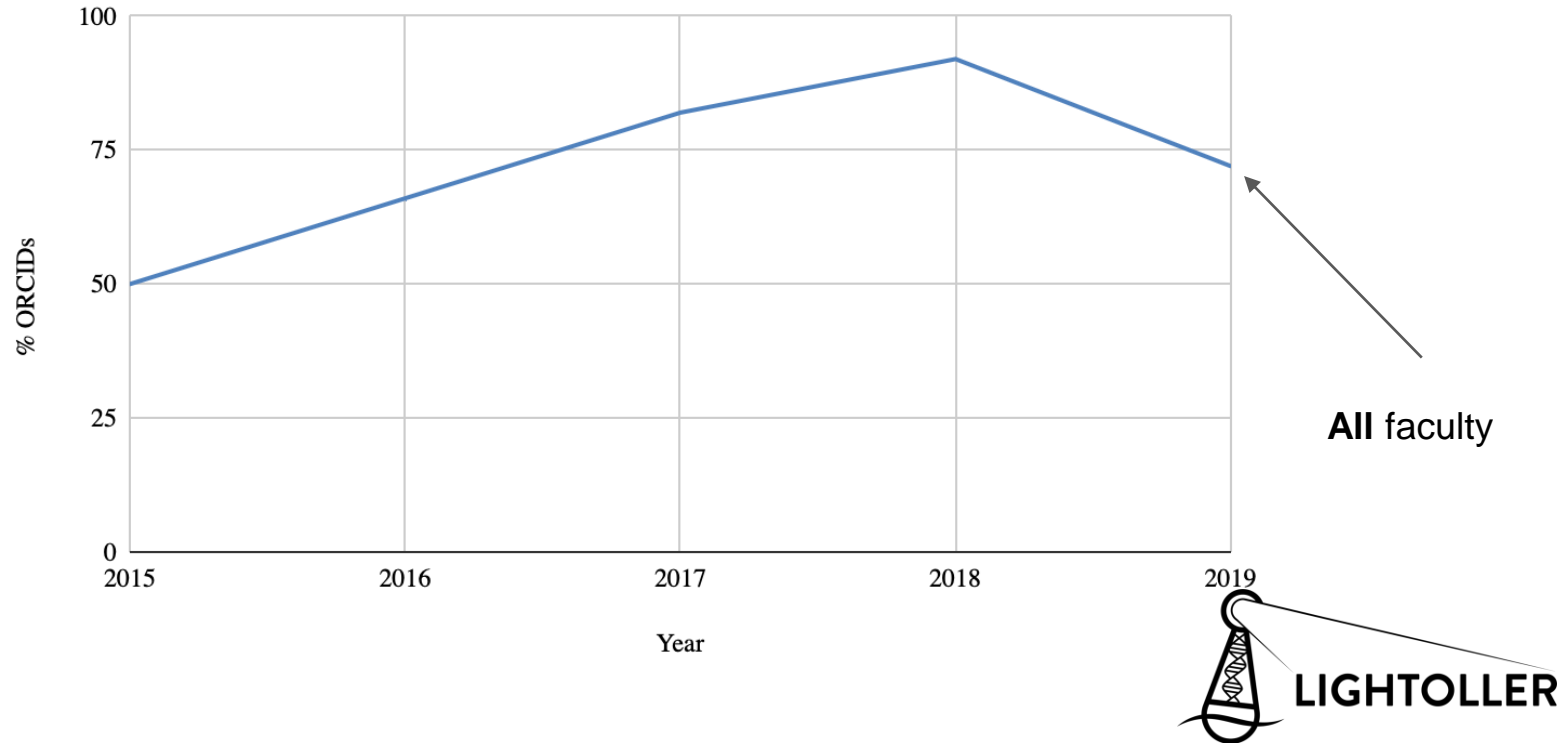LIGHTOLLER

Using ORCID to look at a small dataset

% ORCIDs vs. Year

LIGHTOLLER

Using ORCID to look at a small dataset

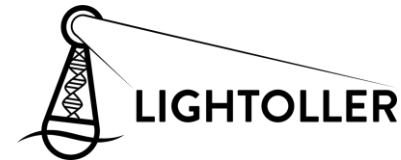% ORCIDs vs. Year

All faculty

LIGHTOLLER
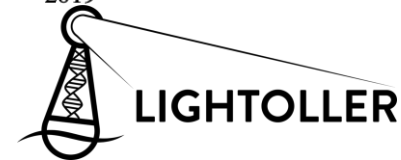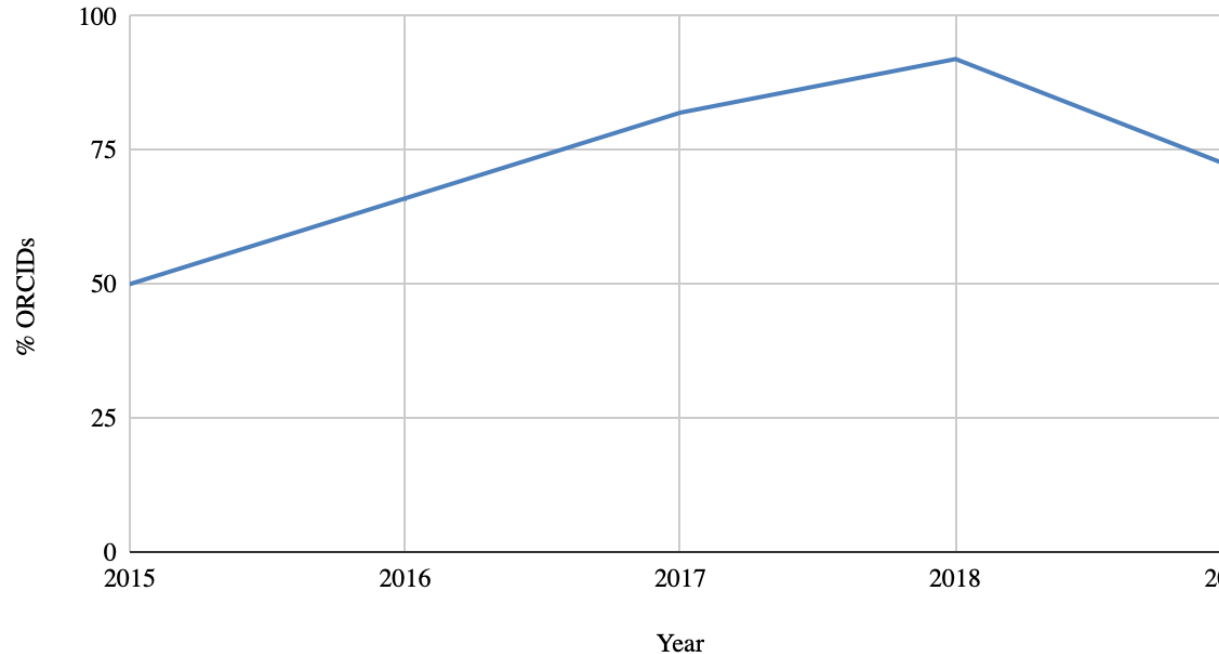
Using ORCID to look at a small dataset

Overall, 47% of 2015-2019 database members are currently in faculty positions

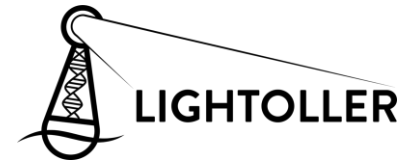56% of those with ORCIDs are in faculty positions

   (i.e. ORCID is more useful for looking at faculty)

LIGHTOLLER

Initial findings working with small public ORCID datasets

Kinds of Information most generally *available* and somewhat *standardized*, for tracking interest:

- Current location and job title
  - Dates

LIGHTOLLER

Initial findings working with small public ORCID datasets

Kinds of Information most generally *available* and somewhat *standardized*, for tracking interest:

- Current location and job title
  - Dates

- "Works" e.g. publications (after job info, number of single publication entries (mandate?))
  - Dates
  - Venue
  - Title

LIGHTOLLER

Initial findings working with small public ORCID datasets

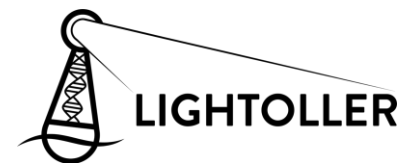Kinds of Information most generally *available* and somewhat *standardized*, for tracking interest:
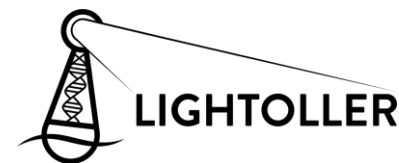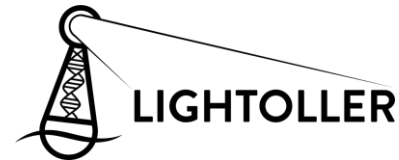
- Current location and job title
  - Dates
- "Works" e.g. publications (after job info, number of single publication entries (mandate?))
  - Dates
  - Venue
  - Title
- Funding
  - Source
  - Amount
  - Title
  - Description

LIGHTOLLER

Initial findings working with small public ORCID datasets

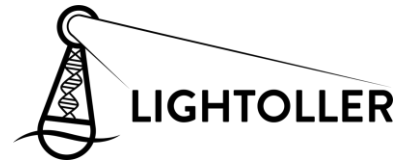ORCID is particularly useful for disambiguation

Currently a reasonable amount of data to work with, potentially more in member API

LIGHTOLLER

Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

LIGHTOLLER

Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

- Variations on data availability, based on what person set as private/public

LIGHTOLLER

Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

- Variations on data availability, based on what person set as private/public
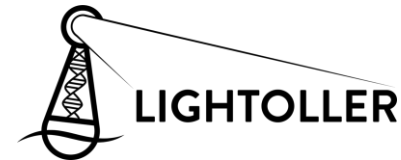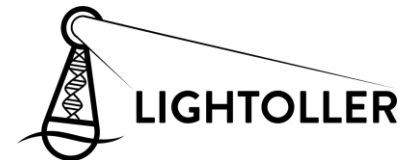
- Variations on current location e.g. move to LinkedIn

LIGHTOLLER

Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

- Variations on data availability, based on what person set as private/public

- Variations on current location e.g. move to LinkedIn

- While data is more likely to be verified by author and *correctly attributed*, some may be missing purely by *omission*, or have inaccuracies

LIGHTOLLER

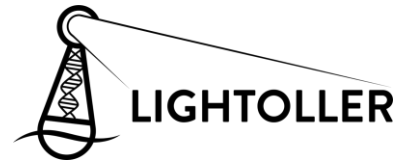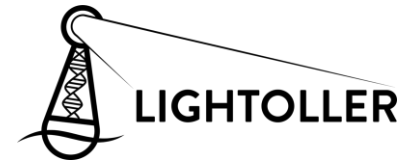# Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

- Variations on data availability, based on what person set as private/public

- Variations on current location e.g. move to LinkedIn

- While data is more likely to be verified by author and *correctly attributed*, some may be missing purely by *omission*, or have inaccuracies
    - Dependent on how often author maintains account

LIGHTOLLER

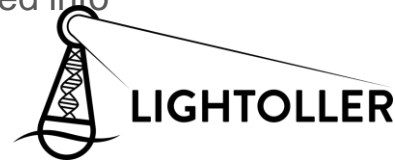# Initial findings working with small public ORCID datasets

Caveats

- No emails (publicly)

- Variations on data availability, based on what person set as private/public

- Variations on current location e.g. move to LinkedIn

- While data is more likely to be verified by author and *correctly attributed*,
  some may be missing purely by *omission*, or have inaccuracies

  - Dependent on how often author maintains account

  - However, may be likely to change with increasing focus on externally verified info

LIGHTOLLER

Initial findings working with small public ORCID datasets

Opportunities

- Member orgs have access to potentially more data through member API
  - Last update date very recent i.e. same day (likely through API connections) - which **is public**, even if no other information is, suggesting more data likely to be available from member API.

LIGHTOLLER

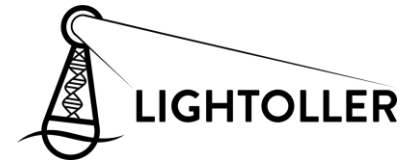# Initial findings working with small public ORCID datasets

Opportunities

- Member orgs have access to potentially more data through member API
  - Last update date very recent i.e. same day (likely through API connections) - which **is public**, even if no other information is, suggesting more data likely to be available from member API.
- Communication efforts could drastically increase the information available to member organizations

LIGHTOLLER

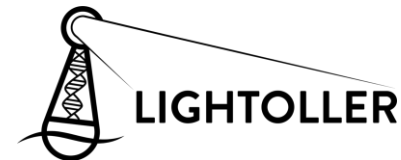# Initial findings working with small public ORCID datasets

Opportunities

- Member orgs have access to potentially more data through member API
  - Last update date very recent i.e. same day (likely through API connections) - which **is public**, even if no other information is, suggesting more data likely to be available from member API.
- Communication efforts could drastically increase the information available to member organizations
- For various reasons, ORCID data is only likely to continue to get *better*

LIGHTOLLER

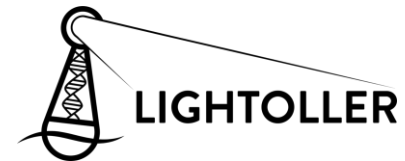# Initial findings working with small public ORCID datasets

Opportunities

- Member orgs have access to potentially more data through member API
  - Last update date very recent i.e. same day (likely through API connections) - which **is public**, even if no other information is, suggesting more data likely to be available from member API.
- Communication efforts could drastically increase the information available to member organizations
- For various reasons, ORCID data is only likely to continue to get *better*
- Opportunities to push "service" etc into unused categories (perhaps standardize?)

LIGHTOLLER

Future Work

Suggested future directions for the HRA community

- Move towards long-term reliance on ORCID data

LIGHTOLLER

Future Work

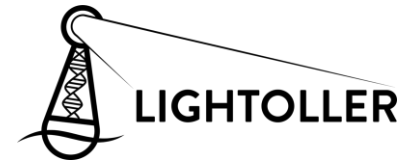Suggested future directions for the HRA community

- Move towards long-term reliance on ORCID data

- Work with developer to set up what your organizations need

LIGHTOLLER

Future Work

Suggested future directions for the HRA community

- Move towards long-term reliance on ORCID data

- Work with developer to set up what your organizations need

- Work with ORCID on specific outcomes tracking needs

LIGHTOLLER

Future Work

Suggested future directions for the HRA community

- Move towards long-term reliance on ORCID data

- Work with developer to set up what your organizations need

- Work with ORCID on specific outcomes tracking needs
 - what is your ideal dataset?

LIGHTOLLER

Future Work

Suggested future directions for the HRA community

- Move towards long-term reliance on ORCID data

- Work with developer to set up what your organizations need

- Work with ORCID on specific outcomes tracking needs
  - what is your ideal dataset?
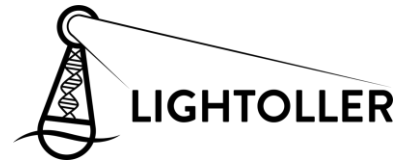 - advocate for what is needed, that they could deliver

LIGHTOLLER

General Personal Thoughts on Use of Evaluation Tools

LIGHTOLLER
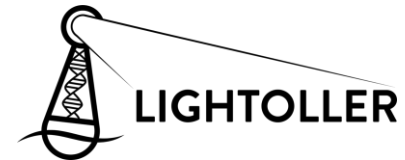
General Personal Thoughts on Use of Evaluation Tools

Tools that allow you to understand where the data is coming from/how it is aggregated in a transparent fashion should be a priority - so that you can be fully aware of the caveats.

Greater focus on data about *people* in concert with e.g. publications data will allow for appreciation of context and (dis)advantages of different groups

LIGHTOLLER

Acknowledgements and Questions