

# Careers in Biocuration: 2023 Workshop Report

Randi Vita<sup>1</sup>, Maria Cristina Aspromonte<sup>2</sup>, Susan M. Bello<sup>3</sup>, Nomi L. Harris<sup>4</sup>, J. Harry Caufield<sup>4</sup>, Melissa Haendel<sup>5</sup>, Charles Tapley Hoyt<sup>6</sup>, Federica Quaglia<sup>2</sup>, Joyeuse Mujambere<sup>7</sup>, Stephen P. Panossian, T. B. K. Reddy<sup>4</sup>, Mary Ann Tuli<sup>8</sup>, Varsha K. Khodiyar<sup>9</sup>, Nicole Vasilevsky<sup>7</sup>.

<sup>1</sup>La Jolla Institute, San Diego, CA, USA; <sup>2</sup>University of Padova, Padova, Italy; <sup>3</sup>Jackson Laboratories, Bar Harbor, ME, USA; <sup>4</sup>Lawrence Berkeley National Laboratory, Berkeley, CA, USA; <sup>5</sup>University of Colorado Anschutz Medical Campus, Aurora, CO, USA; <sup>6</sup>Northeastern University, Boston, MA, USA; <sup>7</sup>Critical Path Institute, Tucson, AZ, USA; <sup>8</sup>GigaScience Press, Hong Kong; <sup>9</sup>Echobyte Data Solutions, UK

## Introduction

Biocuration is a professional field comprising biologists, computer scientists, and bioinformaticians who function as data stewards, to develop and apply data standards to make data FAIR: findable, accessible, interoperable, and reusable [1–3]. Biocuration is an intersection of biology, basic sciences, medicine, public health, ecology, data sciences, and other applied biological sciences, and has been described as an intermediary between bench science and bioinformatics. The International Society for Biocuration (ISB; <https://www.biocuration.org>) is a professional society for biocuration, founded in 2009, that aims to promote the field and our work through outreach, community events, networking, and training opportunities, and provide opportunities for biocurators to share results, insights, and advice with each other, both via the annual conference and on a more informal ongoing basis. The ISB website hosts a list of open job postings (<https://www.biocuration.org/community/jobs>), as well as a generic job description (<https://www.biocuration.org/community/biocuration-generic-job-description>), which was created as an outcome of the first career workshop.

At the 16<sup>th</sup> Annual International Biocuration Conference held in Padua (Italy) from April 23rd to April 26th, 2023 (<https://biocuration2023.github.io/>), we hosted the 6<sup>th</sup> Annual Careers Workshop (<https://biocuration2023.github.io/workshops#isb>) to explore ways the ISB can assist biocurators with career progression. At this interactive workshop, we discussed relevant skill sets, how to tailor a resume or curriculum vitae (CV) for finding a position or for career progression, as well as tips for recruitment. The workshop targeted

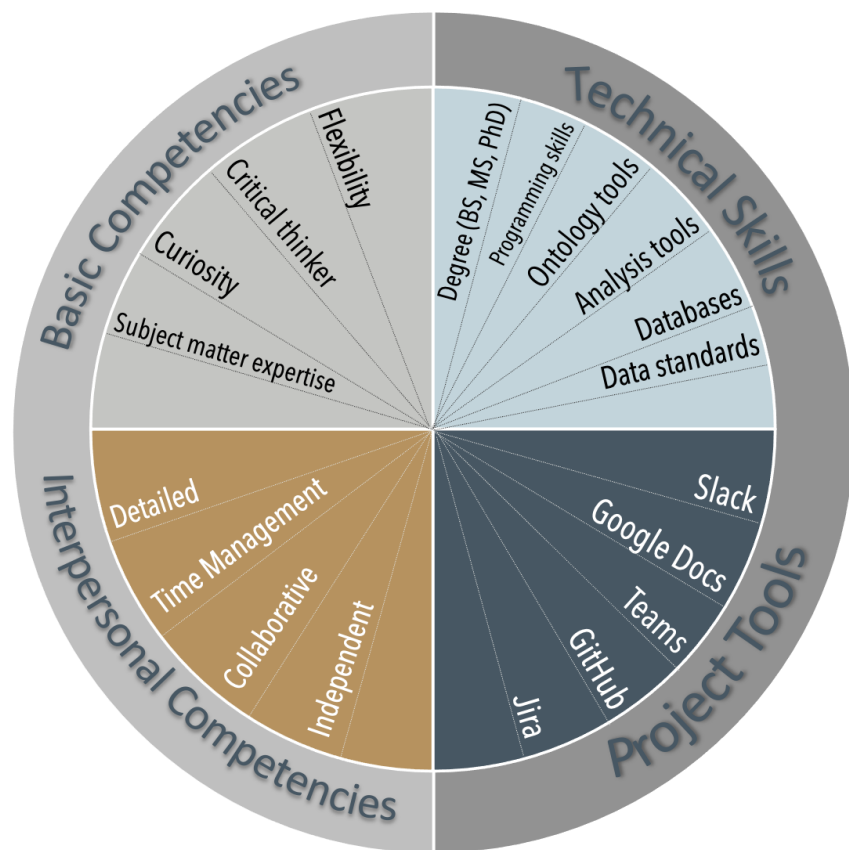
current professionals working in the biocuration field, future biocurators, funders, and policymakers.

## Acquiring the necessary skills

### Core Competencies for Biocurators

Successful biocurators combine domain and subject matter expertise with attention to detail. Biocuration requires meticulous attention and application of standards to unstructured information from disparate sources and therefore requires diligence. A recent publication by members of the biocuration community enumerated ‘10 tips for biocuration’, which provides guidance on successful curation pursuits, and also serves as an entry point for researchers to align their competencies and interests with the field [2]. Generally, a domain-specific degree is required, and many biocurators have an advanced degree (M.Sc. or Ph.D.) in biology or a related field. Anecdotally, being a successful biocurator does not depend on having an advanced degree; curators with coursework in relevant undergraduate courses can be trained to do biocuration in a subject area, especially with a thorough understanding of FAIR principles (<https://www.go-fair.org/fair-principles/>). For a given biological context, biocurators should understand how to create and adhere to standards for data content, meta-data, units, terminology, and file headers and formats. Furthermore, a post-graduate biocurator may progress to assist in the curation of the very same data sets he/she created. Having a Ph.D. or equivalent training/experience can translate to the ability to bring a deep knowledge of an area of biology to the work. Related degrees may also be relevant, such as Library Science, Computer Science, or Bioinformatics.

While not every biocuration position necessitates technical proficiency at the onset, the desire to learn and acquire technical skills is needed for most positions. Programming skills such as Python, Perl, and Java are currently relevant, and proficiency in statistical analysis software like R or SAS may be helpful. Other technical skills may include experience in relational databases, SQL, data transformation, or data visualization tools. Familiarity with ontology tools such as Protege, ROBOT, Ontology Development Kit (ODK), and ROBOT may be pertinent. Given that biocuration is performed in a team environment, and often in remote or distributed work environments; familiarity with, and the ability to use, collaborative tools like Slack, Google Docs, GitHub, and/or Jira is needed. General competencies and skills for biocurators are outlined in Figure 1.



**Figure 1: Relevant competencies and skills for biocurators.** Basic competencies include subject matter expertise, curiosity, and the ability to think critically and be flexible. Additional interpersonal competencies include attention to detail, the ability to manage your time, and to work collaboratively and independently. Skills include technical skills and familiarity with project tools. A degree is typically required, and while not a strict requirement, many biocurators often have an advanced degree in a relevant field. More advanced technical experience is often helpful, but not always required, including experience with building and using data standards and ontologies, databases, programming skills, and experience with analysis tools. Project tools will change, but some relevant tools are highlighted here.

## Paths to a Biocuration Career

The path to professional biocuration is not as straightforward as many disciplines; there are multiple ways to become a biocurator. Formal training programs in the field are few but are becoming more prominent [4,5]. In responses to our survey of the biocuration community, many noted they were bench scientists first and later turned to curation. Most had a PhD in a biological field. Some reported they found their way into the field simply

by responding to an open job opportunity or transitioning within their company. Some became curators while working as computational scientists and became more interested in data completeness and quality. Networking was also beneficial to finding a job in biocuration.

## Finding a job in biocuration

### What to put in a resume or curriculum vitae

Biocuration positions are available in numerous types of institutions, including academia, non-profit, government organizations, and industry. In preparation for the biocuration workshop, the organizers conducted interviews with six US-based and two UK-based biocurators from each type of organization noted above to identify the key features of a biocuration-specific resume or curriculum vitae (CV). Depending on the type of position and geographical location, a resume or CV may be best suited for applying for a position or seeking a promotion.

It should be noted that the specific type of document and document name varies in different countries. While the ISB represents biocurators globally, the guidance provided in this report is biased toward US-based and UK-based practices. We welcome additional perspectives and encourage further discussion on our Slack channel or mailing lists that are open to all (<https://www.biocuration.org/contact-the-isb/>).

In the US, a resume is a succinct summary of education, work history, skill list, and accomplishments, and is typically 1-2 pages and a CV is usually a comprehensive description of one's scholarly and professional achievements, including awards, publications, and invited presentations [6].

In the UK, resumes are referred to as "CVs" and range in length, from short, concise documents, to more extensive and comprehensive documents. A long CV is only required when applying for academic positions, although it may also be acceptable for scientific positions at other institutions, such as government agencies or non-profit organizations.

Biocuration-specific recommended content specific to resumes or CVs is outlined in Table 1. For certain roles, such as domain-specific literature and database curation, emphasis on previous lab experience is highly valuable, if applicable. Previous experience in bench science or lab work offers an invaluable understanding of the experimental data in the published literature, which is difficult to replicate in on-the-job training. In general, highlighting relevant field expertise, whether it be bench science, genetic counseling, social work, etc is useful.

A notable difference between a resume and a CV is the publication list. A resume could highlight some select publications (if applicable), while a CV typically includes a comprehensive list of publications. Instead of including an extensive publication list, authors could include a link to professional profiles such as a Google Scholar profile, ORCID iD (<https://orcid.org/>), or a personal website to enumerate the remainder of the publications. These are described in more detail below.

Specific content for different positions may vary based on the institution type (academia vs. industry vs. non-profit, etc.) and geographic location. Many biocuration positions are available within academic institutions, typically in research track positions. In addition, faculty biocuration positions may also be attainable. In such instances, an extensive curriculum vitae (CV) of unlimited length encompasses an exhaustive list of scholarly activities, including publications, presentations, and awards. Skills are not typically enumerated in an academic CV but may be incorporated or expounded upon within the accompanying cover letter.

**Table 1: Recommended content specific to biocuration to include in your (1-2 page) resume and/or longer CV.**

Content	Resume	CV
Experience/knowledge of the domain/subject matter	Yes	Yes
Experience with frequently used communication and project management tools, potentially including: <ul style="list-style-type: none"> <li>• Slack</li> <li>• Google Docs</li> <li>• Microsoft Teams</li> <li>• GitHub</li> <li>• Jira</li> </ul>	Yes	No
Previous data experience	Yes	Yes
Professional profiles: <ul style="list-style-type: none"> <li>• LinkedIn</li> <li>• Google Scholar</li> <li>• GitHub</li> <li>• ORCID</li> <li>• APICURON</li> <li>• SlideShare</li> </ul>	Yes	Yes
Analytical skills	Yes	Not usually
Websites/dashboards with analyses or visualization	Yes	Yes
Project web pages	Yes	Yes
Website development experience	If applicable	Yes

Content	Resume	CV
Demonstrable communication skills such as presentations, webinars, workshops, documentation, training, etc.	Yes	Yes
Examples of being meticulous with detail and working in a time-sensitive manner	Yes	Yes
Demonstrate ability to keep up to date with new developments in the field	Yes	Yes
Membership in professional associations	Yes	Yes
Major committee and service responsibilities	No	Yes
Grant or manuscript peer review experience	No	Yes
Presentations	No	Yes
Teaching experience	If applicable	Yes
Biocuration contributions to databases and data standards	If applicable	Yes
Contributions to or knowledge of major ontologies	If applicable	Yes
Shared datasets	If applicable	Yes
In the news	No	Yes

## Advice for job seekers

### Take every opportunity to network

Networking involves meeting and exchanging information with people, and can often be mutually beneficial. Job seekers should take advantage of conferences and workshops to meet people connected to their field. Job seekers should talk with people they know or with good contacts who may know people at target organizations. If possible, job seekers should talk directly to hiring managers about positions they are interested in. Moreover, contacting labs, groups or companies of interest, even if they don't have any current openings can lead to opportunities. They may keep the job seeker in mind and contact them directly when they are hiring and/or pass their name on to a different group that is hiring. Many job seekers and hiring managers network online through LinkedIn. The ISB has a LinkedIn page here: <https://www.linkedin.com/groups/1851280>, which provides online networking opportunities and promotion of community resources and events.

Informational interviews are unofficial discussions with a professional working in a field of interest. Arranging a meeting with a professional biocurator can offer insight into the daily activities and work environment, and provide an invaluable networking opportunity. Given that the field of biocuration is less known in the broader biological research community, it can be very satisfying for biocurators to assist the next generation and help grow the field. As

many biocurators entered the profession via circuitous paths, they are generally very empathetic and amenable to sharing insights and information with informational interviewees.

A show of gratitude after an informational interview can be very meaningful; a thank you email or a handwritten thank you card will make a very good impression.

## Gain experience and build a portfolio

A great application for a biocuration job includes a specific demonstration of the ability to do the job. New biocurators or seasoned biocurators who have exclusively worked on private/internal resources or resources that do not track attribution often will not have any specific contributions to point to. These individuals could consider contributing to an open-source project, which demonstrates the ability to learn and apply contribution guidelines, use a specific technology stack (e.g., using social workflows from GitHub), share expertise in a specific domain, and engage with a community.

In contrast to large contributions, many resources are responsive to so-called "drive-by curations" [7] which include small, actionable improvements. This can include small, novel contributions or even minor typographical fixes. Further, meaningful contributions are not only limited to code and data, but also include bug reports, data quality reports, documentation, and more. Often, contributions lead to co-authorships on scientific articles or other presentations and collaborative work will often result in new collaborations and further growing the biocurator's network.

Many biocurators use or contribute to OBO Foundry ontologies (<https://obofoundry.org>), which is a community of ontologists who build and maintain biological and biological ontologies under an agreed-upon set of guiding principles. Participation in the OBO operations committee and other subcommittees is open to the public and is a notable way to contribute to the community and network. ORCID iDs are often included in new term requests or other requested changes in ontologies, to provide nano-attributions for those contributions.

Contributions to open-source projects can be showcased in several ways. Many open-source projects are adopting more granular attribution models, including curation attribution using ORCID identifiers that can be tracked on sites such as APICURON [8], directly on ORCID, or on the sites that show the curated data. APICURON (<https://apicuron.org>) is a resource specifically developed to recognize and credit the work of biocurators. APICURON is a web server that provides biological databases and organizations with a real-time automatic tracking system of biocuration activities, allowing them to attribute and quantify the effort of biocurators. Even in the absence of formal recognition, contributions are visible on an individual GitHub profile (if applicable). Including a link to a GitHub profile in the header portion of a resume or CV allows interested parties to peruse a biocurator's work.

Even in the absence of formal recognition, a biocurator's contributions can be readily observed on their individual GitHub profile. Incorporating a link to this profile in the header section of a resume or CV enables interested parties to explore the biocurator's body of work.

To help curators find projects to contribute to, the ISB website promotes volunteer opportunities (<https://www.biocuration.org/curate-now>). Two exemplary resources are the WikiPathways Academy (<https://new.wikipathways.org/academy>) and the OBO Academy (<https://oboacademy.github.io/oobook>), which both have detailed contribution guidelines and onboarding to help people with some domain knowledge become curators. The ISB offers volunteer opportunities, including conference organization, conference scientific committee (which entails abstract review), and various subcommittees like the Executive Committee Nomination Committee (for the EC yearly election) (<https://www.biocuration.org/community/isb-committees/>).

Further, the ISB is open to community ideas and offers funding for community projects through dedicated 'microgrant' funding (<https://www.biocuration.org/community/microgrants>). Additionally, the ISB offers Exchange Fellowships for curators to visit another group and learn new skills (<https://www.biocuration.org/community/fellowships>).

Demonstrating collaboration efforts, including volunteer efforts in application materials will demonstrate proficiency in team science, which is necessary to succeed in this field.

## The importance of an online presence

Biocurators should ensure their online visibility is optimized, taking into consideration that a lack of online presence may be suboptimal. It is advisable to obtain an ORCID iD and fill out the user profile in detail, including job history, education history, and publications. Users with names that overlap with others in the field, should consider using a middle initial or full middle name in their profile. Similar information should be posted to LinkedIn, ResearchGate, or other professional networking software, as these are frequently leveraged by employers and recruiters to seek or vet candidates.

Creating a personal website is a great way to manage an online presence and highlight work experience. This can be linked to a succinct resume or CV to expand upon accomplishments and sell relevant expertise. A personal website can be created using GitHub for free (thousands of repositories are available to fork to create a personal website: <https://github.com/topics/personal-website>). A personal website goes beyond a resume or CV and can allow biocurators to write in more detail about their interests and experiences. Moreover, demonstrating the ability to create a website will make the biocurator more attractive and competitive to potential employers. Presentations and posters can be shared on sites such as FigShare (<https://figshare.com/>), Zenodo (<https://zenodo.org/>), or an institutional repository.



## The benefits of a mentor

Mentorship is valuable at every career stage. There may be formal mentoring programs available through certain organizations; asking a biocurator for informal mentorship is often well received. Generally, people want to help and are happy to share their expertise. One path towards finding a mentor is by contributing to open-source projects, which typically comprise individuals who are delighted to have external contributors and will offer guidance both in the curation work itself and also in navigating the surrounding community.

## Tailoring a cover letter for the job

While this is generally applicable to any job application, it is critically important that application materials speak to the specific qualifications for a biocuration position. While not always a requirement, a personalized cover letter can help make up for deficits in qualifications, for example, expressing a gap in experience but excitement to learn a new skill can be perceived very favorably. Survey participants recommended reviewing the organization's website and the job posting to ensure the application materials are tailored to the opportunity. For example, they should include specific qualifications and keywords that appear in the advertisement. A recommendation is to generate a word cloud of the job's description and listed responsibilities to determine the top 10 to 15 keywords needed in a CV or resume. Anecdotally, in the US, matching at least 80% of the job's requirements will usually get an application noticed.

## Be prepared for the interview

To prepare for an interview for a biocuration position, first and foremost, the candidate should familiarize themselves with the projects and resources of the organization they are applying to and the domain area in general. Candidates should do background research on the resource they are applying to and familiarize themselves with their website. Additionally, candidates should look into key papers on biocuration and ontology development, as well as key players in the field. The ISB offers yearly awards to exceptional contributors to biocuration; candidates should familiarize themselves with the influential curators in the community (such as the ISB career award winners: <https://www.biocuration.org/community/biocuration-career-awards>). Candidates should consider some key skills that make biocurators successful and market their skills and achievements. They should be prepared to provide examples of being proactive and flexible and explain how their experience can translate to the area covered by the position. It is not necessary for the candidate's background to exactly match the position as long as they can make a case for how they can translate previous experience to the current role.

## Be persistent

Many biocuration positions are grant-funded, and hiring may ebb and flow with the changing funding landscape, political climate, and other external factors, depending on location. Many opportunities are remote but may require staff to reside within a particular country or continent. Biocuration as a recognized appears to be growing, the demand for data management and data science is increasing, and it is expected that this field will continue to evolve [3,9]. As one workshop attendee advised, maintain your contacts, don't lose hope, and keep applying.

## Recruiting a Biocurator

### Writing a job description

While the field of biocuration has been around for decades, it is still a relatively new field and underrecognized. The ISB created a generic job description that is available online (<https://www.biocuration.org/community/biocuration-generic-job-description/>). The keywords in the job description are important to enable suitable applicants to find an advert among the thousands on the internet. It was noted that not all job postings use the word 'biocuration', something that hiring managers may want to consider including, even if the position title is something else. Related keywords and position titles include but are not limited to, Informaticist, Data Analyst, Data Wrangler, Data Manager, Data Engineer, and a number of variations on the term "Curator" (such as biocurator, scientific data curator, etc).

### Job titles

Upon surveying the ISB community prior to the workshop, the respondents provided nearly 30 different titles for their current roles, including (bio)curator, (Data) Wrangler, Analyst, Bioinformatician/Bioinformatics Engineer, Data Editor, Data Manager, Information Specialist, Ontologist, Project Manager, and Quality Analyst. The biocuration community lacks standardized titles for specific roles, which can be problematic with respect to career progression and pay equity. This was discussed at last year's Biocuration Careers Workshop and is an opportunity for further work by the ISB or the community [10].

## Conclusion

Biocuration is a diverse field that continues to grow. The ISB is a community-led organization that focuses on promoting the field and bringing together biocurators and

organizations that want to hire them. In this article, we have summarized some of the advice for job seekers and hiring managers that was collected in a survey of ISB members (67 people completed the survey).

ISB welcomes contributions to promote the field and grow the workforce. Current biocurators, people interested in joining the field, and those wanting to hire biocurators are all encouraged to join the ISB and contribute to these discussions. Future projects may include formally defining the core competencies for biocurators, developing training programs for these skills, or additional aspects of career progression such as opportunities for career advancement, how academic jobs differ from industry jobs, and obtaining tenure.

Participation in the community and ISB is welcomed; please visit <https://www.biocuration.org/> to discover ways to get involved.

## Additional information:

Workshop Slides:  [Biocuration Careers Workshop\\_2023-04-26](#)

Biocuration 2023 Conference Video:  [Workshop - Careers In Biocuration](#)

## References

1. Wilkinson MD, Dumontier M, Aalbersberg IJJ, Appleton G, Axton M, Baak A, et al. The FAIR Guiding Principles for scientific data management and stewardship. *Sci Data*. 2016;3: 160018.
2. Tang YA, Pichler K, Füllgrabe A, Lomax J, Malone J, Munoz-Torres MC, et al. Ten quick tips for biocuration. *PLoS Comput Biol*. 2019;15: e1006906.
3. International Society for Biocuration. Biocuration: Distilling data into knowledge. *PLoS Biol*. 2018;16: e2002846.
4. Palmer CL, Heidorn BP, Wright D, Cragin MH. Graduate Curriculum for Biological Information Specialists: A Key to Integration of Scale in Biology. 2008;2: 31–40.
5. Postgraduate certificate in biocuration. 30 Oct 2019 [cited 6 Oct 2023]. Available: <https://www.ice.cam.ac.uk/course/postgraduate-certificate-biocuration>
6. Christenbery TL. The curriculum vitae: gateway to academia. *Nurse Educ*. 2014;39: 267–268.
7. Hoyt CT. Drive-by Curation. *Qeios*. 2022 [cited 27 Nov 2023].

doi:10.32388/KBX9VO

8. Hatos A, Quaglia F, Piovesan D, Tosatto SCE. APICURON: a database to credit and acknowledge the work of biocurators. Database . 2021;2021: baab019.
9. Burge S, Attwood TK, Bateman A, Berardini TZ, Cherry M, O'Donovan C, et al. Biocurators and biocuration: surveying the 21st century challenges. Database . 2012;2012: bar059.
10. Creators Nicole Vasilevsky<sup>1</sup> Mohammad Hosseini<sup>2</sup> Kristi Holmes<sup>2</sup> Mary Ann Tuli<sup>3</sup> Randi Vita<sup>4</sup> Show affiliations 1. University of Colorado Anschutz Medical Campus 2. Northwestern University 3. GigaScience 4. La Jolla Institute. Biocuration Careers Workshop. doi:10.5281/zenodo.7102460