

Gendering data care: curators, care, and computers in data-centric biology

Ane Møller Gabrielsen

To cite this article: Ane Møller Gabrielsen (25 Sep 2023): Gendering data care: curators, care, and computers in data-centric biology, Science as Culture, DOI: [10.1080/09505431.2023.2260830](https://doi.org/10.1080/09505431.2023.2260830)

To link to this article: <https://doi.org/10.1080/09505431.2023.2260830>



© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group



Published online: 25 Sep 2023.



Submit your article to this journal [↗](#)



Article views: 565



View related articles [↗](#)



View Crossmark data [↗](#)

Gendering data care: curators, care, and computers in data-centric biology

Ane Møller Gabrielsen 

Department of Philosophy and Religious Studies, NTNU – Norwegian University of Science and Technology, Trondheim, Norway

ABSTRACT

The increase in molecular data and the use of computer technologies in biology have led to the emergence of professional biocurators, who populate biological databases and knowledgebases with high-quality information. Although crucial to life science knowledge production, biocuration is, to a large extent, invisible labour that takes place behind the scenes of data-centric life science. The field suffers from a lack of recognition and status that has been linked to a language of service and a scientific system that is not equipped to recognise and reward new types of scientific practices. However, as the majority of biocurators are highly educated female biologists, biocuration is also reproducing the problematic pattern of women leaving the scientific tenure track in favour of less prestigious positions. Instead of viewing the issue as just another example of ‘the leaky pipeline,’ the gendering of biocuration could be seen as an interplay of gendered structures in science, organisations and society which makes a career in biocuration attractive for female scientists while at the same time positioning the activity as non-scientific low-status work. By illuminating some of the ways gender works in the processes which render certain kinds of technoscientific work invisible, biocuration serves as an example of how existing social structures influence the emerging data-centric science.

ARTICLE HISTORY



Received 26 September 2022
Accepted 15 September 2023

KEYWORDS

Biocuration; data curation; data care; data-centric biology; gender; feminisation

Introduction

In *Matters of care in technoscience: Assembling neglected things* (2011), STS scholar and philosopher Maria Puig de la Bellacasa suggests directing our attention to the practices of care that are taking place in sociotechnical assemblages and often devalued and rendered invisible. In this paper, I will direct the attention toward biocuration, ‘the extraction of knowledge from unstructured biological data into a structured, computable form’ (International Society for

CONTACT Ane Møller Gabrielsen  ane.gabrielsen@ntnu.no  Department of Philosophy and Religious Studies, NTNU – Norwegian University of Science and Technology, Postboks 8900, NO-7491 Trondheim, Trondheim, Norway

© 2023 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. The terms on which this article has been published allow the posting of the Accepted Manuscript in a repository by the author(s) or with their consent.

Biocuration, 2018b, p. 2). The word ‘curation’ comes from the Latin *curare*, meaning ‘taking care of,’ and biocuration can thus quite literally be understood as life science data care.

Despite their limited numbers – the The International Society for Biocuration (ISB) currently counts around 250 members – the impact of professional biocurators is massive. Biological databases have become crucial to life science research (Baxevanis and Bateman, 2009; Hall *et al.*, 2013) and a conservative estimate of the usage of EMBL-EBI resources alone amounts to 88 million data accessions per year (Beagrie and Houghton, 2016). However, despite the increasing attention and efforts directed toward data managing and dissemination, biocuration tends to be overlooked and undervalued by the scientific community (Holinski *et al.*, 2020).

Within STS, there is a strong tradition of foregrounding the importance of necessary and often uncredited care work in terms of technoscientific maintenance work (e.g. Shapin, 1989; Knorr-Cetina, 1999; Mol *et al.*, 2010; de la Bellacasa, 2017). It is tempting to compare biocurators with Steven Shapin’s account of the invisible technician in seventeenth-century science (1989). The role of the technician was that of the servant, and as servants and craftsmen, technicians were effectively written out of the history of knowledge production. Shapin explains this with the political and moral economy at the times, where science was primarily considered to be an activity of the mind, not the body, and where the relationship between master and servant effectively prevented the work of the latter from being recognised as scientific contributions. The ‘gentlemen scientists’ possessed the authority allowing them authorship and acknowledgement and as Shapin comments, the modern equivalent giving scientists their authority, is the PhD (Shapin, 1989). As biocurators generally come from a scientific background and a majority of them hold PhDs, this is clearly not the reason for their scientific invisibility. I will therefore turn my focus towards something which is lacking from Shapin’s analyses of both past and present invisible scientific work: gender.

Not much is written on the topic of gender in data-centric science, but the existing literature suggests that practices of data care tend to be carried out by women (Pinel *et al.*, 2020). This corresponds with biocurator surveys where the majority of respondents are female (Burge *et al.*, 2012; International Society for Biocuration and Vasilevsky, 2021).¹ In other words, biocuration seems to reproduce a well-known gendered pattern in which women occupy the less prestigious jobs in science (Birke, 1986). There are, however, few accounts of *how* gender works to devalue some types of technoscientific work.

Regarding biocuration, there are not many accounts at all. While digital curation and data curation in general have begun to attract some academic attention, mainly within library and information studies (e.g. Faniel *et al.*, 2014; Johnston *et al.*, 2018), the biocuration literature consists mainly of descriptions of curation workflows (e.g. Burkhardt *et al.*, 2006; St. Pierre and McQuilton, 2009) or appeals

to the scientific community for recognition and support (e.g. Howe *et al.*, 2008; Bateman, 2010; International Society for Biocuration, 2018b). One exception to the academic silence surrounding biocuration is philosopher of science Sabina Leonelli, who has written about database curators and their importance in her work on model organism databases, data-centric biology, and life science knowledge production (e.g. Leonelli, 2014, 2016). This article aims to expand on Leonelli's work by exploring the gendering of biocuration and its effects based on interviews with biocurators and theory on gendered work and feminisation. The research questions structuring the paper are as follows: Why are there so many women in biocuration? How do gendering processes relate to the status of biocuration as technoscientific work? How does gendered data care play out in data-centric biology on a broader scale?

Background, method, and material

The specific background for this paper is the interdisciplinary efforts that took place in connection with a responsible research and innovation (RRI) project centred around the notion of a Life Science Knowledge Commons. The knowledge commons was conceptualised as the standardisation and interlinking of existing digital life science infrastructures in a manner that would make the content accessible and understandable for computers. In the project, computer scientists, biologists, and socio-humanists explored the conditions for 'well-constructed knowledge commons while, at the same time, attempting to establish a distributed knowledge resource for a specific type of gene regulation information. This was done within the context of systems biology, which is an interdisciplinary approach to biology applying theories and tools from systems theory, mathematics, and computer science in order to model, understand and, ultimately, predict and control the mechanisms of living systems (Kitano, 2002; Fujimura, 2005). Because the models in question depend on the existence of detailed information about the molecular bits and pieces in the systems, as well as the interactions between them (Kitano, 2002), systems-level modelling would therefore benefit from a unified and interoperable Life Science Knowledge Commons.

As a postdoctoral researcher with background in STS and gender studies, I soon encountered the challenges of navigating the rather complex and confusing territory of systems biology, biological databases, and computational modelling. A sort of unstructured and informal ethnography took place as I travelled along with the life scientists on their missions to negotiate the gene regulation knowledge resource,² and this is where I encountered the biocurators. I had never heard about biocurators before entering the project, and as it turned out, neither had most of the people I talked to, including biologists. However, it soon became clear that biocurators, in themselves, were important to any attempts to create a sustainable and useful digital knowledge resource.

Negotiations with biocurators thus became an important task for the project as a whole. Although unknown to many, they were the gatekeepers of the biological databases and together with another project member, I decided to interview biocurators about their work. The established connection between our project and several important biological databases at the European Bioinformatics Institute (EMBL-EBI) proved to be of great value in terms of recruiting informants, and through snowball sampling we ultimately recruited eleven informants (eight women and three men) working at various open life science databases. These included several databases at the EMBL-EBI as well as the Gene Ontology knowledgebase and some smaller European biological databases relevant for systems biology. A majority of these databases were so-called knowledge bases or value-added databases where curators manually added relevant information about the biological entities in question. Nine of the interviewed biocurators were employed on temporary contracts, which is the usual practice in the field. For instance, EMBL-EBI operates with the so-called ‘nine year-rule’ which means that the regular three-year contracts are not renewed more than two times.

The interviews were carried out in order to understand more about biocuration as a condition for biological knowledge infrastructures, but also to gain more information about the issue that had gotten my attention: that biocurators were important, yet still almost unknown outside the databases. The semi-structured interviews included topics like the practice of biocuration, becoming biocurators, what they cared about regarding their work, and how they experienced their own status. Due to the noticeable lack of male biocurators, gender also became a topic in these interviews which constitute the main material for the paper. All identifying information has been altered or removed from interview quotes used in the text. This includes specifying the databases further as several of these databases only employ a few biocurators.

Analytical perspectives: gender, work, and feminisation

STS contributions often highlight the uneven distribution of power in technoscience. Examples are Lucy Suchman’s displaced human workers in *Human-Machine Reconfigurations* (2007) and Susan Leigh Star and Anselm Strauss’ ‘Layers of silence, arenas of voice: the ecology of visible and invisible work’ (1999). In the following I will draw on insights from these works and traditions but also add theory about gender, work and organisations in order to analyse how gendered processes are working in life science knowledge infrastructures.

Gendered work, often defined as the division of labour between men and women, could be viewed as work defined, organised, divided, and valued in ways that reflect patterns of relations between women and men or the meanings associated with ‘masculinity’ and ‘femininity’ (Chalmers, 2014). While earlier

accounts of gendered work tended to perceive gender as something static and given, influencing the division of labour from the 'outside,' later feminist accounts have emphasised the dynamic, multidimensional, and interconnected nature of gender in work practices and organisations (Chalmers, 2014). As Joan Acker writes in the seminal work 'Hierarchies, Jobs, Bodies: A Theory of Gendered Organizations' (1990): 'Gender is not an addition to ongoing processes, conceived as gender neutral. Rather, it is an integral part of those processes, which cannot be properly understood without an analysis of gender' (Acker, 1990, p. 146).

According to Acker, gender is a constitutive element in what is termed *organisational logic*, or 'the underlying assumptions and practices that construct most contemporary work organizations' (Acker, 1990, p. 147). Although this logic may seem gender-neutral, there is a gendered substructure that is continuously reproduced through activities and organisational writings and materialised through items such as written rules and management tools. An essential element of the logic consists of hierarchies that are taken for granted, both by managers and by the workers themselves. Different types of jobs are ranked within these hierarchies, and there is an assumed relationship between hierarchical position and the level of complexity and responsibility. A low level in the hierarchy is therefore assumed to correspond with a low level of complexity and responsibility (Acker, 1990).

The term *feminisation* is useful in order to describe some of the processes involved in the gendering of organisations. The term has multiple meanings and may refer to the rising rate of female participation in the workforce or to women entering and coming to dominate a field previously dominated by men (with the effect often being that wages and status are reduced) (Fondas, 1997). A third meaning of feminisation is the one employed by Nanette Fondas in 'Feminization Unveiled: Management Qualities in Contemporary Writings' (1997), in which feminisation refers to 'the spread of traits or qualities that are traditionally associated with females to things or people not usually described that way' (Fondas, 1997, p. 258). This is the understanding of feminisation I will use throughout this text.

Fondas shows how qualities usually held to be feminine have begun to show up in the management literature, in contrast to the traditional masculine leadership style. The management literature is thus spreading a 'feminine ethos' without ever naming it as such (Fondas, 1997). According to Fondas, the management literature legitimises the feminine ethos and contributes to its spread in managerial practices. The term 'feminine' or other explicitly gendered terms are, however, never used to describe the desired qualities. Fondas argues that this is due to the status of masculinity in management discourse, which, again, is closely connected to the existing gender regimes of the corporate world. Naming the desired management qualities as 'feminine' would disrupt the image of the universal male manager. Instead, the feminine is brought

into the discourse under the guise of neutrality, which, again, does nothing to reveal or challenge the already implicitly gendered managerial discourse in which masculinity is the norm.

Fondas' analytical framework draws on a poststructuralist understanding of language as the main constituent of our social reality. According to this view, language consists of signs, which are given meaning by being associated with or contrasted with other signs (Jørgensen and Phillips, 2002). An analysis of feminisation is therefore dependent on some qualities being viewed as masculine and others being viewed as feminine. In 'Gendered skills and unemployed men's resistance to "women's work",' Yavorski *et al.* list the following qualities as connected to either men or women:

Men are viewed as agentic, meaning they are perceived as being achievement-oriented (e.g. competent, ambitious, task-focused), inclined to take charge (e.g. assertive, dominant, forceful), autonomous (e.g. independent, self-reliant, decisive), and rational (e.g. analytical, logical, objective). Feminine-typed skills, on the other hand, emphasize communality and have come to denote concern for others (e.g. kind, caring, considerate), affiliative tendencies (e.g. warm, friendly, collaborative), deference to others (e.g. obedient, respectful, self-effacing), and emotional sensitivity (e.g. perceptive, intuitive, understanding) (Yavorski *et al.*, 2021, p. 1527).

It is important to note that this is not a list of the inherent qualities of men and women but, rather, traits that are culturally ascribed to the categories of 'male' and 'female.' Because the genders are considered complementary and mutually exclusive, something described as feminine is also 'not masculine.' Furthermore, as Yavorski *et al.* notes the skills and qualities associated with men and masculinity are usually those perceived as the most valuable. In the following, I will draw on insights from these theoretical and analytical approaches in analysing the role and effects of gender in biocuration.

Caring for life science data

In 1990, the Human Genome Project was officially launched, and in 1991, molecular biologist Walter Gilbert, at Harvard University, stated that biology was facing a paradigm shift: the soon-to-be-realised knowledge of all the genes would guide all future biological research, and the vessel for the shift would be electronic biological databases (Gilbert, 1991). Because the large numbers of nucleotide sequences could not be published through the usual means, claims were made that biological databases would have to become a new means of primary scientific literature in order to deal with the 'megabytes of archival-quality data' that were being generated every year (Robbins, 1994, p. 3).

Today, the size of archival-quality data has far exceeded megabytes and biologists have joined 'the Big Data club' (Marx, 2013). Other terms used to describe the current situation are 'data-driven' and 'data-intensive,' but as Leonelli notes, the changes are not so much related to a new methodological focus

on data, but rather to the attention given to the ways data is being handled and disseminated. She therefore terms the current situation *data-centric*, where ‘data-centric’ refers to an approach to science where ‘efforts to mobilize, integrate, and visualize data are valued as contributions to discovery in their own right and not as a mere by-product of efforts to create and test scientific theories’ (Leonelli, 2016, p. 2).

Once again filled with promise, the increasing number of life science data also comes with challenges (Attwood *et al.*, 2009; Chadwick and Zwart, 2013). Means of interpretation and integration are lacking, and the issue is often framed as one concerning the way in which data and information are handled and organised. As Attwood *et al.* state:

The real problem is that we have failed to store and organize much of the rapidly accumulating information (whether in databases or documents) in rigorous, principled ways, so that finding what we want and understanding what’s already known become exhausting, frustrating, stressful and increasingly costly experiences (Attwood *et al.*, 2009, p. 318).

This is where biocurators enter the scene. Otherwise known as database curators or annotators (Harding, 2006), biocurators collect and connect biological data and information from various sources, translate it into structured and computable formats and publish it in biological databases (Howe *et al.*, 2008).³ As of 2022, there are 1,645 freely accessible biological databases (Rigden and Fernández, 2022), and it is particularly the ‘added-value databases’ or ‘knowledgebases’ of molecular biology that constitute the domain of professional biocurators (Figure 1).

Although researchers may perform curation activities as part of their data practices, the majority of the work is carried out by professionals, with the largest clusters working with public databases at institutions such as the European Bioinformatics Institute (EMBL-EBI), the Swiss Institute for Bioinformatics (SIB) or the US National Center for Biotechnology Information (NCBI) (Harding, 2006; Sanderson, 2011; Holinski *et al.*, 2020).

Scientific picture straighteners

In ‘The Art of Biocuration,’ Vivienne Baillie Gerritsen and Marie-Claude Blatter write that biocurators ‘have been referred to – whether endearingly or not – as ‘museum cataloguers of the internet age,’ ‘those who prefer computers to pipettes,’ ‘self-confessed bookworms,’ or ‘monk copyists’ (Gerritsen and Blatter, 2016). The notion of biocurators as obsessed with rules and order seems to be common. One biocurator job description asks, ‘Have you ever been called pedantic or precise? Did you take that as a compliment, not a criticism? If so, we have the job for you!’ (International Society for Biocuration, 2018b).



Figure 1. Biocurator. Illustration of a biocurator in action uploaded to Wikimedia Commons by user Andrawaag and licenced under CC0 1.0 Universal (CC0 1.0) Public Domain Dedication.

Although it is time-consuming, biocurators seem to enjoy this careful tracing and assembling of information and the feeling of ‘getting it right’ in the end. As one biocurator stated, ‘I’ve always said we’re the picture straighteners and nit-pickers of the universe. I think, if you’re not that sort of person, then you’re not going to curate particularly well. You have to care about the detail, about getting things correct’ [BC10]. Furthermore, ‘getting things correct’ requires more than care for details and order. Biocurators facilitate what Leonelli (2016) refers to as ‘data journeys,’ in which data travel from their original site of production to be reused in new settings. One important part of the work consists of detaching data and information from their original context and rearticulating them for reuse, which Leonelli (2016) terms ‘decontextualization’ and ‘recontextualization.’ These processes consist of labelling data using standardised terms and providing sufficient metadata, i.e. additional information about the data and

how they were produced in order for the user to make informed decisions about quality and relevance.

The decontextualisation and recontextualisation of data are active translation processes requiring substantial knowledge and experience, and a majority of biocurators have PhDs in biology and are trained in experimental science (International Society for Biocuration, 2018b). As one biocurator commented, there are also many aspects of biocuration that correspond with traditional biological research: ‘Even though it’s just reading papers, you are doing research. Because you develop the ontology, you are developing things, and you are developing ideas’ [BC09]. The ‘ontology’ in question refers to bio-ontologies, standardised classificatory systems that allow biological information to be represented in computable formats and facilitate integration and comparison across species and databases (Blake and Bult, 2006; Boem, 2016). Ontologies entered biology from computer science (Boem, 2016) and an important aspect of biocuration is also to facilitate sense-making for computers by translating the ‘natural’ language in papers into structured formats that computers recognise and are able to process.

‘There are certainly a lot of women’

A quick glance at the names of biocurators employed at different databases or at the members of the executive committee of the International Society for Biocuration (ISB) through the years reveals an overwhelming majority of women 2023. Why do female biologists choose biocuration? An interesting observation is that none of the biocurators we interviewed, regardless of gender, wanted to become a biocurator from the outset of their careers. They usually did a PhD and perhaps a couple of postdocs in biology before changing direction and as a male biocurator explained, the motivation has more to do with personality and interests than with gender:

I think it’s more like a character thing, rather than the gender thing. Because, for me, if you are not the bookwormish type that I described beforehand, if you need more action, then this job is not for you, because it’s a quiet job. For me, it’s more like the character thing. If you need more oomph, then you do something else, but that could be true for a woman as well as for a man, so I don’t really know why there is an imbalance in terms of gender [BC04].

A female biocurator noted that the large number of women was typical of biology in general: ‘There are certainly a lot of women, but that might be a reflection of the fact that it’s biology and more of them tend to be women in the biological sciences anyway’ [BC06]. The gender ratio in the biosciences is known to be more balanced than in other STEM disciplines (science, technology, engineering, and mathematics), with women accounting for about half of the awarded PhDs (Hill *et al.*, 2010; Bonham and Stefan, 2017). After the PhD, however, the number of female biologists drops rapidly (Hill *et al.*, 2010),

making biology a typical example of ‘the leaky pipeline,’ i.e. the phenomenon of women leaving science before reaching tenure (Goulden *et al.*, 2011). Because most biocurators are women with PhDs in biological science, they could be viewed as examples of the leakage from the life science pipeline, but thus far, the reasons for the gender disparity have not been explored.

The quiet and meticulous work of biocuration is not for everyone, but none of the interviewed biocurators described biocuration as a something women would be better suited to than men. One of the female biocurators commented that the more technical databases seemed to be more balanced:

The databases that are a bit more technical, that you need more programming skills, for instance, the European Nucleotide Archive, which is all about just taking big files and fixing them a bit and putting them into a big archive, it’s probably more gender balance there. It’s probably more men [BC10].

The same biocurator noted that there were always more women applying for curation posts than for developer posts and suggested that this could be related to female communication skills: ‘[C]uration posts, particularly the ones where there’s a lot of writing free text, always seems to get more women applying than men. And, I mean, I don’t know ... female communication skills are supposed to be better [...]. I don’t know whether it’s just coincidence ...’ [BC10].

The comment illustrates the binary between what are considered to be feminine, social activities and masculine, technological performance (Gansmo *et al.*, 2003). However, she did not seem quite convinced by her own argument and added ‘Or women feel more pressure to leave the lab. I mean, this sort of work you can do at home, so if you got a baby, you can take a laptop home’ [BC10].

When we followed up on the topic in other interviews, almost all the female biocurators mentioned having children as a reason for going into biocuration. This included one who initially stated that she simply grew weary of failing experiments: ‘I really enjoyed working with plants, and growing them in greenhouses was very therapeutic, but when it came down to doing PCRs and repeating them constantly and just not getting results ... [...] Plus, I just had my first child as well and wanted something a bit, not easier, but, yeah, I don’t know, 9 to 5’ [BC08]. The most explicitly gendered story came when a female biocurator asked if we wanted ‘the real story’ of how she ended up in biocuration. She then recalled how she had approached a PI and told him that she wanted to do her PhD in the ‘wet lab’ (laboratory). He first approved, but when he found out she had children, he changed his mind:

Then, he stopped and said, ‘I have to think about it because you have to decide whether you want to stay in the lab until night or if you want to stay with your family.’ I said that I wanted to stay with both, and he answered, ‘Ok, you can have half or your thesis as wet biologist and the other half as curator’, and I agreed. And

after two weeks, he came and said, ‘Ok, I’ve decided that you will just be a curator’ [BC05].

Although she later came to feel that biocuration had been ‘the greatest part of my life’ [BC05], the story of how she entered the profession demonstrates that biocuration is perceived as more attainable than lab work for women with children.

The leaky pipeline is a complex phenomenon, with several suggested causes ranging from a lack of ambition to cultural expectations and traditional gender roles (Schiebinger, 1999; Hill *et al.*, 2010; Goulden *et al.*, 2011). However, according to Goulden *et al.* (2011), family formation, in terms of marriage and children, accounts for the largest leaks of women. The large number of women in biocuration is therefore likely to be a result of the flexible work hours and the biological sciences’ tendency to exclude persons with child care responsibilities. In theory, working from home with a flexible schedule should be just as relevant for men with children. However, as a male biocurator’s experience at EMBL-EBI demonstrates, caring for children still often ends up being the woman’s responsibility due to factors like uneven distribution of parental leave:

So, in our case, for example, my wife and I, we are both curators, we both took part time and we both shared the load of work. But since she was the one that had the children, that had to go through the medical step, and also, the paternity leave here is not great, it is not equal with maternity leave. So, she ended up taking more of a load of family care than I did [BC07].

Or as a female biocurator put it: ‘it is very appealing to working mums, really. And unfortunately, there is this kind of disparity where the mothers are always the first to drop everything and look after their kids’ [BC02].

Unsung heroes

As Schiebinger (1999) emphasises, leaving the laboratory should not necessarily be viewed as a failure. There is important scientific work to be done outside the realms of academic science, and biocuration is a field where women and men can practice science beyond the competitive demands of traditional biology. One female biocurator commented that biocuration is also a field where women may have more opportunities than in the more male-dominated academic STEM disciplines: ‘[A] lot of PIs and people who have pushed biocuration are women as well. So, in terms of the executive structure, compared to, you know, your university structure, there are a lot of women there. So, it’s actually a place where career progression is quite possible’ [BC02].

The possibilities for career progression are, however, perhaps not as promising as this quote suggests. Despite the increasing attention and effort directed toward data managing and dissemination, there is limited allocation of

resources for biocuration (Howe *et al.*, 2008; Leonelli, 2016), and as the above-mentioned biocurator also explained, '[Funding agencies] always want a connection with the wet lab and the translational aspects of the work. So, it's really hard to get funding agencies to fund the pure curation. You always have to link it to something' [BC02]. The lack of long-term funding makes permanent positions scarce, and many biocurators are employed on temporary contracts that will only be renewed a limited number of times. Career planning is therefore difficult and, according to a 2017 survey, only 18% of the respondents were 'very satisfied' with their career progression (International Society for Biocuration, 2017), while an earlier survey showed that a majority of biocurators expressed concern about future work opportunities and saw the lack of opportunities for career progression as a barrier to remaining in biocuration (Burge *et al.*, 2012).

In addition to the lack of funding, the status of biocuration and biocurators in the scientific community seems to be low. Biocurators have been termed 'the unsung heroes of molecular biology' (Bateman, 2010, p. 991), and for good reason. As Bateman (2010) notes, users of biological databases tend to take the availability of information for granted, and according to one biocurator, many researchers are surprised to learn that there are people doing the work of extracting, annotating and integrating information: 'They just assume it's there, it's always been there and, if anything, it's automatically added. Although there are actually people reading papers and adding things, it's like, 'that sounds crazy'' [BC08].

Biocuration is currently not acknowledged or rewarded by established academic mechanisms (Ankeny and Leonelli, 2015; Pinel *et al.*, 2020), and as Ankeny and Leonelli note, biocuration is viewed as something useful but not scientifically important:

Curation is more often viewed as background or routine work whose results are crucial to making data available but do not influence the ways in which data are analyzed and interpreted. Hence, data curation remains largely invisible to data users, who do not view the gathering and formatting of data as 'real' scientific work involving conceptual decisions (Ankeny and Leonelli, 2015, p. 141).

Despite being an activity that actively influences the results that can be drawn from the data and information, the notion of biocuration as non-scientific service work seems to be prevalent. As one experienced biocurator reported, this also influences career progression: 'I think that, in the research areas, people see it as a service. I actually applied for a job recently in a university, and the head of the department wrote back to me and said, 'We don't want somebody doing this, providing this service. We want someone doing research'' [BC09]. The general lack of status was further confirmed by a biocurator who recollected how her former supervisor seemed to be disappointed over her career choice when she decided to leave academic biology:

When I told her that I wanted to be a biocurator, she kind of felt a bit disappointed and thought, ‘oh, you could go on to do so much better.’ So, I think she thought it was kind of a step down for me, which was kind of disheartening. And I kind of have that, all the time, I have that feeling, that maybe, if I chose differently ... I kind of have that feeling for myself, maybe I could have done a bit better [BC08].

As this last quote shows, when biocuration is considered a step down from ‘real’ science, this also has consequences for the practitioners’ view of themselves.

Gendering data care

At first glance, biocuration seems to be almost ‘accidentally gendered’ because it is not usually perceived as a gendered occupation and the large number of women seems to be a direct result of the flexibility of the work, which makes it compatible with childcare responsibilities. The appeal of biocuration for female biologists who need to combine work and family responsibilities was commented upon in 2000 by Amos Bairoch, the founder of the Swiss-Prot protein database (Bairoch, 2000):

It is also interesting to note that for almost 10 years Jean- Pierre was the only male annotator in the Geneva SWISS PROT group. One of the key reasons that made and still makes the SWISS-PROT group attractive to women scientists with children is that it is possible to work part time, with a flexible schedule and that part of the work can be done from home. All of which is not possible with practical laboratory work (Bairoch, 2000, p. 54).

According to Acker, the association with childcare could in itself be a contributing factor in keeping biocuration at the lower end of the hierarchy. Even the concept of ‘a job’ is gendered in itself because it ‘assumes a particular organization of domestic life and social production’ (Acker, 1990, p. 149), and women’s jobs are devalued due to the association with ‘childbearing and domestic life’ (Acker, 1990, p. 152). However, there are also other ways that gender can influence the status and positioning of an occupation and in the following I will take a closer look at the language of biocuration.

In *Data-centric Biology*, Leonelli (2016) argues that part of the reason for biocuration’s lack of scientific status lies in the language of service characterising the way curation activities are communicated to research communities. This language is important in order to make the databases attractive to users but also means that the scientific contribution of biocuration is downplayed. There is, however, also a gendered dimension of the rhetoric used in communicating biocuration to the wider scientific community. Like the managerial literature Fondas analyses, biocuration is never explicitly described as feminine or something women should be better suited to than men. Still, there are many similarities between the desired qualities for biocurators and Fondas’ description of qualities traditionally ascribed to women:

[E]mpathy, helpfulness, caring, and nurturance; interpersonal sensitivity, attentiveness to and acceptance of others, responsiveness to their needs and motivations; and orientation toward the collective interest and toward integrative goals such as group cohesiveness and stability; a preference for open, egalitarian, and cooperative relationships, rather than hierarchical ones; and an interest in actualizing values and relationships of great importance to community (Fondas, 1997, p. 260).

An example is the previously mentioned generic biocurator job description on the ISB webpage suggesting how biocuration could be advertised. Although subject matter expertise is listed at the top, three out of the next four typical job requirements in the job description are culturally coded as feminine: collaborative skills, communication skills and interpersonal skills:

Typical job requirements

- Subject matter expertise – typically a PhD, although not a requirement
- Ability to collaborate and work on a team
- Defines and refines rules and standards (as data types and user requirements evolve)
- Can communicate well with computer programmers, bioinformaticians and biologists alike
- Liaise with all stakeholders regarding the data, from the producers/submitters to the consumers

(International Society for Biocuration, 2018a)

The job description goes on to list several other types of skills and competencies, but there are not any mention of qualities that are traditionally coded as masculine, like analytical, logical, independent or ambitious (Yavorsky *et al.*, 2021). Furthermore, biocurators are often described in ways that allude to caring capacities. As mentioned above, the word ‘curation’ means to take care of, and biocuration could be seen as the practical caretaking needed to enhance the value of the data and make it useful for others. There is, however, also an affective dimension of care, that is, ‘the activity of attending to others and responding to their emotions and needs’ (Coltrane and Galt, 2002, p. 16). In the job advertisement, the best biocurators are described as ‘adaptable to the needs of the community and/or to the needs of the software systems’ (International Society for Biocuration, 2018a). The ideal biocurators are thus described as caring in the affective and nurturing sense, which is a dimension of care often attributed to women (Tronto, 1993; Baez *et al.*, 2017).

The descriptions above could be seen as examples of feminisation. In Fondas’s use of the term, ‘feminisation’ refers to the use of qualities that are culturally coded as feminine to describe something that is not usually described in that manner (Fondas, 1997). One might argue that collaboration, communication and care are crucial elements of biocuration that always have been

and *should* be present in descriptions of the practice and its practitioners. The issue is, however, not that biocurators are described in this way. Rather, it is that scientists are usually not. According to Schiebinger, science ‘was part of the territory that fell to the masculine party in the struggles that divided social and intellectual labour between the sexes in European society’ in the eighteenth century (Schiebinger, 1989, p. 233). When women were excluded from the spaces where science was conducted, this also led to the exclusion of ‘a whole set of values, qualities, and characteristics subsumed under the term *femininity*’ (Schiebinger, 1989, p. 234). Because the traits in question are associated with femininity, they thus serve as a contrast for ‘science,’ where, just as in the managerial discourse in Fondas’s example, masculinity is the norm. In other words, ‘science’ is already gendered, and due to the assumed differences in and complementarity of the genders, naming the desired qualities for biocurators also implicitly articulates biocurators as non-scientists despite their PhDs and the scientific aspects of their work.

Another potential effect of feminised language in advertising biocurator positions could be the reinforcement of the actual gendering in terms of the number of women applying. Although the large number of women seems to be an unintended result of structural inequalities concerning laboratory biology and the organisation of parental leave, several studies have shown that gendered language in job advertisements matters in terms of who applies, as well as that men tend to avoid applying for jobs emphasising feminine-typed skills (Gaucher *et al.*, 2011; Yavorsky *et al.*, 2021). As one biocurator stated in the above, curation positions involving a great deal of writing, always attracts more women than men, while the technical databases seem to be more balanced in terms of gender. Although seemingly reflecting inherent differences between men and women, this could also be an issue of how the positions are advertised and communicated. The effect is, nonetheless, that the gendered construction of a biocurator materialises as real women working as biocurators.

It might be interesting to compare biocuration with related fields like bioinformatics, a rather recent interdisciplinary field applying techniques from maths, computer science, and statistics in order to analyse and interpret large-scale omics-data (Luscombe *et al.*, 2001). Although bioinformatics also suffers from a lack of acknowledgement (Lewis *et al.*, 2016), there are important differences. Unlike biocuration, which only exists as postgraduate courses, bioinformatics has been institutionalised as an academic field and the lack of acknowledgement is rather a matter of not getting *sufficient* credit. As Lewis *et al.* writes in ‘Hidden in the Middle: Culture, Value and Reward in Bioinformatics’ (2016), bioinformaticians might ‘not be as invisible as Shapin’s technicians’ but are ‘rarely afforded centre stage’ (p. 487). Biocurators, on the other hand, are usually hidden behind the stage altogether. The reason for this difference in status is hardly the level of scientific importance. Curated biological databases play a crucial role in defining what counts as biological

knowledge for the fields they cover and could be said to constitute *obligatory passage points* (Callon, 1984); i.e. the position which defines what is considered as true knowledge about a field (Johnsen, 2004). Still, the curators creating and maintaining these knowledge resources are struggling to be acknowledged or granted scientific status. Again, the statistics are limited, but Lewis *et al.* (2016) report that 80% of the respondents to their survey amongst UK bioinformaticians were male while other sources state a male ratio of around 65% (Frontiers, 2023). As bioinformatics thus seems to be quite male dominated it does not seem unreasonable to argue that gender plays a part in the positioning of the fields.

Gendering data-centric biology

The gendering of biocuration is also part of a broader picture. In *Memory Practices in the Sciences*, Geoff Bowker notes that databases are not the product of the computer revolution. In fact, it is the other way around (Bowker, 2005). The sequencing of the genome ‘brought computational and informatics approaches to the forefront of life sciences research’ (National Research Council, 2005, p. 12), and the careful collection and standardisation of information has enabled an increasing amount of biological research to take place *in silico* in the form of computer modelling and simulations (Lewis *et al.*, 2016). It could therefore be argued that data-centric biology is also a matter of *computer-centric* biology.

In her account of the epistemic agency of technology, philosopher Federica Russo (2016) notes that knowledge is ‘distributed not just across the ‘brains’ of the scientists but also across the instruments that scientists use’ (Russo, 2016, p. 166). Technology thus becomes an epistemic agent that ‘partakes in [the] production of data, in their analysis, and thereby in their interpretation’ (Russo, 2016, p. 166). Although the epistemic agency of technology tends to be downplayed in traditional scientific accounts, it is often highlighted in data- and computer-centric science where the ideal digital formats are not only computer-readable but also computer-*interpretable* (Aranguren *et al.*, 2011), meaning that computers should ideally be able to make sense of them on their own.

The envisioned futures of computer-centric biology often take shape as automated processes where machines and artificial intelligence makes sense of the increasing amounts of life science data with little or no human intervention (e.g. Ginsparg, 2009; Evans and Rzhetsky, 2010; Nielsen, 2012; Alkhateeb, 2017). Although these automated scenarios are far from the current reality, they are what Sheila Jasanoff terms *sociotechnical imaginaries*; ‘collectively held and performed visions of desirable futures’ (Jasanoff, 2015, p. 4) and as such, they inform current approaches and policies. An example is the FAIR principles for scientific data management and stewardship (Wilkinson *et al.*, 2016). The FAIR principles provide guidelines for how to improve the

Findability, Accessibility, Interoperability, and Reuse of digital assets, and are currently gaining importance in science policies worldwide (Mons *et al.*, 2017). Within the FAIR discourse, the goal is ‘machine-actionability’ (Wilkinson *et al.*, 2016), and the optimal state is described as a state ‘where machines fully “understand” and can autonomously and correctly operate-on a digital object’ (Wilkinson *et al.*, 2016, p. 3).

However, computers are also entities in need of human facilitation and assisting them is seen as ‘a critical consideration for all participants in the data management and stewardship process’ (Wilkinson *et al.*, 2016, p. 3). Constituting machines both as active agents in knowledge creation as well as entities in need of human assistance, influences the way subjects are positioned in terms of power. As Russo notes, ‘behind the machines, the softwares, or any any other piece of technology, there always is the (techno)scientist, or actually, many (techno)scientists’ (Russo, 2016, p. 160), meaning that there are always human epistemic agents utilising the machines to produce knowledge. The FAIR literature clearly states that the computational agents are working on someone’s behalf, but also that someone is working on behalf of the computational agents, a category that implicitly includes biocurators who enable actionable machines through scientific translation work. Through imaginaries of machine actionability, biocurators are thus positioned as assistants and service providers for computers.

It should be noted that there have been numerous attempts to automate biocuration, and biocurators themselves are often avid supporters of automating parts of the workflows (Tang *et al.*, 2019). However, as Leonelli (2016) notes, biocuration is difficult to automate completely due to the level of judgement involved in extracting and rearticulating information. The curation process requires substantial knowledge about biology as well as the use of tacit knowledge and subjective judgement developed through training and experience (Ankeny and Leonelli, 2015) and the skills of human biocurators will therefore still be necessary in the foreseeable future. Still, the discourses surrounding data-centric biology seldom mention the scientific expertise and cognitive effort required in order to facilitate machine-actionability. As Suchmann notes, discourses of information technology tend to erase human labour in favour of a utopian fantasy of the perfect, invisible infrastructure (Suchman, 2007), and while machines are granted epistemic agency in the emerging computer-centric biology, those who facilitate the machines are usually not.

The rise of computer technology in the life sciences resonates with historian Thomas Haigh’s account of how the computer’s role, in the corporate world during the 1960s, was transformed from being a processor of data to ‘a mighty information system sitting at the very heart of management, serving executives with vital intelligence about every aspect of their firm’s past, present, and future’ (Haigh, 2001, p. 16). According to Haigh, the computer was intimately connected with ‘the systems men,’ who were attempting to

gain managerial control. This was also a gendered process, and as Haigh notes, the use of the term ‘systems men’ could reflect an attempt ‘to separate themselves from the appreciable number of women working in the lower-status job of office manager’ (Haigh, 2001, p. 22).

The gendering of biocuration seems to reinforce a similar gendered divide between those who control the computers and those who are seen as merely assisting them in data-centric biology. As Russo noted, there are always technoscientists behind the machines and as women are underrepresented in computational biology, i.e. the field using computational models to study biological functions and systems (Bonham and Stefan, 2017), there is reason to assume that a large proportion of them are male. When biocurators are positioned as caretakers and service providers for computers, they are therefore reproducing a well-known gendered pattern dating back to the female keypunchers who, in the 1960s, punched the hand-written computer programs into the punch cards for the machines to read, while the operators of the machines were male (Parolini, 2015).

Conclusion

By carefully attending to digital data and information, biocurators constitute an intrinsic part of the systems that represents and shapes life science data and knowledge. Perhaps not surprisingly, much of this work is carried out by women, and the number of female biocurators reproduces well-known gendered patterns in science. Throughout this paper I have argued that while the large number of women in biocuration is an effect of gendered structures of biological laboratory science and inequalities concerning childcare responsibilities, the discursive feminisation of biocuration through emphasis on care, communication, and collaboration implicitly contrasts biocurators with scientific actors. It is important to note that the problem is not that there are many female biocurators. Furthermore, biocuration *does* require care, collaboration, and communication, regardless of the gender of the practitioner. Although there are attempts to counteract gendered language in job descriptions, like Gender Decoder (2023), the problem with the way biocuration is communicated is not that the language has feminine connotations, but rather that the qualities in question are marginalised in representations of science in general. Being represented in a manner that emphasises qualities often interpreted as ‘feminine’ rather than capacities that are perceived as ‘scientific’ or even ‘masculine’ could therefore contribute to the positioning biocuration as non-scientific service work.

This means that the low status of biocuration and biocurators is not simply a matter of a lack of awareness or hidden work but must be viewed in the light of how the organisational logic of science is already gendered. By attracting women with childcare responsibilities and by communicating a ‘feminine ethos’ which traditionally is seen as incompatible with scientific work, biocuration becomes both materially and discursively gendered in a way that places the

field at the bottom of both the scientific and the organisational hierarchies of the data-centric biosciences. While Shapin's technicians (Shapin, 1989) lacked the societal status granting them scientific authority, gendered dynamics position biocurators as non-scientific service providers who are effectively written out of the official accounts of life science knowledge production.

In the final section of the paper, I have explored how gendered data care is intertwined with devaluation of human effort in sociotechnical life science imaginaries. As Starr and Strauss note (1999), work becomes visible through a selection of changing indicators. When technology is granted epistemic agency through envisioned machine actionability, the work that goes into facilitating the machines is further removed from scientific recognition. The effect is that biocuration is positioned as service work for machines and biocurators become parts of an invisible gendered infrastructure enabling computational biology. Although the results of their work are visible through the databases, biocurators fade further into the background and become what Star and Strauss term *non-persons*: people producing visible results while remaining invisible – a position more commonly associated with domestic service personnel, but increasingly relevant for digital infrastructures (Star and Strauss, 1999).

The biosciences are often perceived as being in the forefront of a computerised data-centric scientific revolution but have not taken into account how this development tends to reproduce and reinforce traditional gendered patterns of inequality. The constant dripping from the leaky pipelines suggests that more female biologists will seek towards the uncredited fields of data care while male biologists continue to pursue fields like bioinformatics and computational biology. There are, however, also signs of change towards more recognition of data care. In the midst of the commotion surrounding machine-centric data-driven science, there are also calls for a more open and collaborative science which recognises the need to credit and reward data curation (e.g. European Commission, 2018). Furthermore, biocurators are working toward increasing the awareness and funding of biocuration through the International Society for Biocuration (ISB). By emphasising the importance of biocuration efforts as well as their role as professional scientists, the ISB is working towards increased visibility and acknowledgement (e.g. International Society for Biocuration, 2018b). To avoid reproducing and reinforcing existing gendered inequalities in emergent data-centric science, these efforts may benefit from addressing both the gendering processes currently taking place as well as the obvious need for qualities traditionally coded as 'feminine' in the emerging data-centric science.

Notes

1. Of 130 respondents in the 2021 survey, 81 listed their gender as female, 31 listed their gender as male and one reported themselves to be non-binary. Seventeen of the respondents did not indicate gender.

2. For more about the background for this endeavor, see Tripathi *et al.* (2016): ‘Gene regulation knowledge commons: Community action takes care of DNA binding transcription factors.’
3. ‘Biocurator’ is the most common job title, but there is still a lack of standardised names and titles for biocuration positions (Vasilevsky, 2021).

Acknowledgements

I would like to thank the editors of *Science as Culture* and the anonymous reviewers for their helpful and constructive comments. I also wish to thank my colleagues, Anamika Chatterjee and Sophia Efstathiou, for their valuable input, feedback, and support; my Crossover collaborators, Astrid Læg Reid, Martin Kuiper, Steven Vercreusse, Marcio Acencio, and Rune Nydal, for all the valuable discussions and insights; Sandra Orchard for the kind help and all the biocurators for generously giving me their time and stories.

Disclosure statement

No potential conflict of interest was reported by the author(s).

Funding

This research was funded by the Research Council of Norway through the project Crossover Research 2.0: Well-constructed Knowledge Commons under grant number 247727.

Notes on contributor

Ane Møller Gabrielsen holds a PhD in Interdisciplinary Studies of Culture and is currently working as a senior research librarian at the Norwegian University of Science and Technology. She has experience from STS, Gender Studies, and Human-Animal studies and has previously published on dogs as well as the topic of trust in the biosciences.

ORCID

Ane Møller Gabrielsen  <http://orcid.org/0000-0002-2470-3451>

References

- Acker, J. (1990) Hierarchies, jobs, bodies: A theory of gendered organizations, *Gender & Society*, 4(2), pp. 139–158. doi:10.1177/089124390004002002
- Alkhateeb, A. 2017. Can scientific discovery be automated. *The Atlantic*. Available at <https://www.theatlantic.com/science/archive/2017/04/can-scientific-discovery-be-automated/524136/> (accessed March 30 2023).
- Ankeny, R. A. and Leonelli, S. (2015) Valuing data in postgenomic biology: How data donation and curation practices challenge the scientific publication system, in: S. S. Richardson, and H. Stevens (Eds) *Postgenomics: Perspectives on Biology After the Genome*, pp. 126–149 (Durham and London: Duke University Press).

- Aranguren, M. E., Stevens, R., Antezana, E., Fernández-Breis, J. T., Kuiper, M. and Mironov, V. (2011) Technologies and best practices for building bio-ontologies, in: G. Alterovitz, and M. Ramoni (Eds) *Knowledge-based Bioinformatics: From Analysis to Interpretation*, pp. 67–86 (Chichester: John Wiley & Sons).
- Attwood, T. K., Kell, D. B., McDermott, P., Marsh, J., Pettifer, S. R. and Thorne, D. (2009) Calling international rescue: Knowledge lost in literature and data landslide!, *Biochemical Journal*, 424(3), pp. 317–333. doi:[10.1042/BJ20091474](https://doi.org/10.1042/BJ20091474)
- Baez, S., Flichtentrei, D., Prats, M., Mastandueno, R., García, A. M., Cetkovich, M. and Ibáñez, A. (2017) Men, women ... who cares? A population-based study on sex differences and gender roles in empathy and moral cognition, *PLoS One*, 12(6), pp. e0179336. doi:[10.1371/journal.pone.0179336](https://doi.org/10.1371/journal.pone.0179336)
- Bairoch, A. (2000) Serendipity in bioinformatics, the tribulations of a Swiss bioinformatician through exciting times!, *Bioinformatics (Oxford, England)*, 16(1), pp. 48–64. doi:[10.1093/bioinformatics/16.1.48](https://doi.org/10.1093/bioinformatics/16.1.48)
- Bateman, A. (2010) Curators of the world unite: The international society of biocuration, *Bioinformatics (Oxford, England)*, 26(8), pp. 991. doi:[10.1093/bioinformatics/btq101](https://doi.org/10.1093/bioinformatics/btq101)
- Baxevanis, A. D. and Bateman, A. (2009) The importance of biological databases in biological discovery, *Current Protocols in Bioinformatics*, 50(1), pp. 1.1.1–1.1.8. doi:[10.1002/0471250953.bi0101s27](https://doi.org/10.1002/0471250953.bi0101s27)
- Beagrie, N. and Houghton, J. (2016) *The value and impact of the European Bioinformatics Institute, Hinxton: EMBL-EBI*. Available at www.beagrie.com/EBI-impact-report.pdf.
- Birke, L. (1986) *Women, Feminism and Biology: The Feminist Challenge* (Brighton: Harvester Press).
- Blake, J. A. and Bult, C. J. (2006) Beyond the data deluge: Data integration and bio-ontologies, *Journal of Biomedical Informatics*, 39(3), pp. 314–320. doi:[10.1016/j.jbi.2006.01.003](https://doi.org/10.1016/j.jbi.2006.01.003)
- Boem, F. (2016) Orienteering tools: Biomedical research with ontologies, *Humana. Mente Journal of Philosophical Studies*, 9(30), pp. 37–65.
- Bonham, K. S. and Stefan, M. I. (2017) Women are underrepresented in computational biology: An analysis of the scholarly literature in biology, computer science and computational biology, *PLoS Computational Biology*, 13(10), pp. e1005134. doi:[10.1371/journal.pcbi.1005134](https://doi.org/10.1371/journal.pcbi.1005134)
- Bowker, G. C. (2005) *Memory Practices in the Sciences* (Cambridge, MA: MIT Press).
- Burge, S., Attwood, T. K., Bateman, A., Berardini, T. Z., Cherry, M., O'Donovan, C. and Gaudet, P. J. D. (2012) Biocurators and biocuration: surveying the 21st century challenges, *Database*, 2012, pp. bar059. doi:[10.1093/database/bar059](https://doi.org/10.1093/database/bar059)
- Burkhardt, K., Schneider, B. and Ory, J. (2006) A biocurator perspective: Annotation at the research collaboratory for structural bioinformatics protein data bank, *PLoS Computational Biology*, 2(10), pp. e99. doi:[10.1371/journal.pcbi.0020099](https://doi.org/10.1371/journal.pcbi.0020099)
- Callon, M. (1984) Some elements of a sociology of translation: domestication of the scallops and the fishermen of St Brieuc Bay, *Sociological Review*, 32(1_suppl), pp. 196–233. doi:[10.1111/j.1467-954X.1984.tb00113.x](https://doi.org/10.1111/j.1467-954X.1984.tb00113.x)
- Chadwick, R. and Zwart, H. (2013) From ELSA to responsible research and promismomics, *Life Sciences, Society and Policy*, 9(1), pp. 3–3. doi:[10.1186/2195-7819-9-3](https://doi.org/10.1186/2195-7819-9-3)
- Chalmers, L. (2014) Gendered work, in: A. C. Michalos (Eds) *Encyclopedia of Quality of Life and Well-Being Research*, pp. 2472–2475 (Dordrecht: Springer Netherlands).
- Coltrane, S. and Galt, J. (2002) The history of men's caring, in: M. H. Meyer (Eds) *Care Work: Gender, Labor, and the Welfare State*, pp. 15–36 (New York: Routledge).
- de la Bellacasa, M. P. (2011) Matters of care in technoscience: Assembling neglected things, *Social Studies of Science*, 41(1), pp. 85–106. doi:[10.1177/0306312710380301](https://doi.org/10.1177/0306312710380301)

- de la Bellacasa, M. P. (2017) *Matters of Care: Speculative Ethics in More Than Human Worlds* (Minneapolis: University of Minnesota Press).
- European Commission. (2018) *Turning FAIR into Reality. Final Report and Action Plan from the European Commission Expert Group on FAIR Data* (Luxembourg: Publications Office of the European Union). doi:10.2777/1524
- Evans, J. and Rzhetsky, A. (2010) Philosophy of science. Machine science, *Science*, 329 (5990), pp. 399–400. doi:10.1126/science.1189416
- Faniel, I. M., Minor, D. and Palmer, C. L. (2014) Putting research data into context: Scholarly, professional, and educational approaches to curating data for reuse, *Proceedings of the American Society for Information Science and Technology*, 51(1), pp. 1–4. doi:10.1002/meet.2014.14505101016
- Fondas, N. (1997) Feminization unveiled: Management qualities in contemporary writings, *The Academy of Management Review*, 22(1), pp. 257–282. doi:10.5465/AMR.1997.9707180266
- Frontiers. 2023. Women in Bioinformatics – 2023, Frontiers Research Topic. Available at <https://www.frontiersin.org/research-topics/54952/women-in-bioinformatics—2023> (accessed 01 April 2023).
- Fujimura, J. H. (2005) Postgenomic futures: Translations across the machine-nature border in systems biology, *New Genetics and Society*, 24(2), pp. 195–226. doi:10.1080/14636770500184826
- Gansmo, H. J., Lagesen, V. A. and Sørensen, K. H. (2003) Out of the boy's room? A critical analysis of the understanding of gender and ICT in Norway, *NORA-Nordic Journal of Feminist Research*, 11(3), pp. 130–139.
- Gaucher, D., Friesen, J. and Kay, A. C. (2011) Evidence that gendered wording in job advertisements exists and sustains gender inequality, *Journal of Personality and Social Psychology*, 101(1), pp. 109–128. doi:10.1037/a0022530
- Gender Decoder. 2023. Available at <http://gender-decoder.katmatfield.com/about> (accessed 20 February 2023).
- Gerritsen, V. B. and Blatter, M.-C. 2016. The Art of Biocuration, *Protein Spotlight*, 177. Available at https://web.expasy.org/spotlight/back_issues/177/ (accessed 30 March 2023).
- Gilbert, W. (1991) Towards a paradigm shift in biology, *Nature*, 349(6305), pp. 99.
- Ginsparg, P. (2009) Text in a data-centric world, in: T. Hey, S. Tansley, and K. Tolle (Eds) *The Fourth Paradigm: Data-Intensive Scientific Discovery*, pp. 185–192 (Redmond, Wash: Microsoft Research).
- Goulden, M., Mason, M. A. and Frasch, K. (2011) Keeping women in the science pipeline, *The Annals of the American Academy of Political and Social Science*, 638(1), pp. 141–162. doi:10.1177/0002716211416925
- Haigh, T. (2001)) inventing information systems: The systems men and the computer, 1950–1968, *The Business History Review*, 75(1), pp. 15–61. doi:10.2307/3116556
- Hall, A. S., Shan, Y., Lushington, G. and Visvanathan, M. (2013) An overview of computational life science databases & exchange formats of relevance to chemical biology research, *Combinatorial Chemistry & High Throughput Screening*, 16(3), pp. 189–198.
- Harding, A. (2006) Rise of the Bio-librarian: The field of biocuration expands as the data grows, *The Scientist*, 20(4), pp. 82–84. Available at <https://www.the-scientist.com/pulse-oximeter/rise-of-the-bio-librarian-47688> (accessed 30 March 2023).
- Hill, C., Corbett, C. and St Rose, A. (2010) *Why So Few? Women in Science, Technology, Engineering, and Mathematics Research Report* (Washington, DC: American Association of University Women (AAUW)). Available at <https://eric.ed.gov/?id=ED509653> (accessed 30 March 2023).

- Holinski, A., Burke, M., Morgan, S., McQuilton, P. and Palagi, P. (2020) Biocuration - mapping resources and needs [version 2], *F1000Research*, 9(1094), pp. 1094. doi:10.12688/f1000research.25413.2
- Howe, D., Costanzo, M., Fey, P., Gojobori, T., Hannick, L., Hide, W., Hill, D. P., Kania, K., Schaeffer, M., Stusan, S. P., Twigger, S., White, O. and Rhee, S. Y. (2008) Big data: The future of biocuration, *Nature*, 455, pp. 47–50. doi:10.1038/455047a
- International Society for Biocuration. (2017). Biocuration Career Description Survey. Available at <https://www.biocuration.org/dissemination/survey-results/> (accessed 30 April 2022).
- International Society for Biocuration. (2018a). Biocuration Generic Job Description. Available at <https://www.biocuration.org/community/biocuration-generic-job-description/> (accessed 30 April 2022).
- International Society for Biocuration. (2018b) Biocuration: distilling data into knowledge, *PLoS Biology*, 16(4), pp. e2002846. doi:10.1371/journal.pbio.2002846
- International Society for Biocuration. 2023. Executive Committee (EC). Available at <https://www.biocuration.org/about/executive-committee/> (accessed 30 April 2023).
- International Society for Biocuration [Equity, Diversity, and Inclusion Committee] and Vasilevsky, N. 2021. Survey results [Data set]. doi:10.5281/zenodo.5504477
- Jasanoff, C. (2015) Future imperfect: science, technology and the imaginations of modernity, in: S. Kim, and S. Jasanoff (Eds) *Dreamscapes of Modernity*, pp. 1–33 (Chicago and London: The University of Chicago Press).
- Johnsen, J. P. (2004) Latour, natur, og havforskere: hvordan produsere natur?, *Sosiologi i dag*, 2004(2), pp. 47–67.
- Johnston, L. R., Carlson, J., Hudson-Vitale, C., Imker, H., Kozlowski, W., Olendorf, R. and Stewart, C. (2018) How important is data curation? Gaps and opportunities for academic libraries, *Journal of Librarianship and Scholarly Communication*, 6(1), pp. 2198. doi:10.7710/2162-3309.2198
- Jørgensen, M. and Phillips, L. (2002) *Discourse Analysis as Theory and Method* (London: Sage).
- Kitano, H. (2002) Systems biology: A brief overview, *Science*, 295(5560), pp. 1662–1664. doi:10.1126/science.1069492
- Knorr-Cetina, K. (1999) *Epistemic Cultures: How the Sciences Make Knowledge* (Cambridge, MA: Harvard University Press).
- Leonelli, S. (2014) Data interpretation in the digital age, *Perspectives on Science*, 22(3), pp. 397–417. doi:10.1162/POSC_a_00140
- Leonelli, S. (2016) *Data-Centric Biology: A Philosophical Study* (Chicago: University of Chicago Press).
- Lewis, J., Bartlett, A. and Atkinson, P. (2016) Hidden in the middle: Culture, value and reward in bioinformatics, *Minerva*, 54(4), pp. 471–490. doi:10.1007/s11024-016-9304-y
- Luscombe, N. M., Greenbaum, D. and Gerstein, M. (2001) What is bioinformatics? A proposed definition and overview of the field, *Methods of Information in Medicine*, 40(4), pp. 346–358.
- Marx, V. (2013) The Big challenges of Big data, *Nature*, 498, pp. 255–260. doi:10.1038/498255a
- Mol, A., Moser, I. and Pols, J. (2010) *Care in Practice: On Tinkering in Clinics, Homes and Farms* (Bielefeld: Transcript Verlag).
- Mons, B., Neylon, C., Velterop, J., Dumontier, M., Da Silva Santos, L. O. B. and Wilkinson, M. D. (2017) Cloudy, increasingly FAIR; revisiting the FAIR data guiding principles for

- the European open science cloud, *Information Services and Use*, 37(1), pp. 49–56. doi:10.3233/ISU-170824
- National Research Council. (2005) *Catalyzing Inquiry at the Interface of Computing and Biology* (Washington, DC: National Academies Press). doi:10.17226/11480
- Nielsen, M. (2012) *Reinventing Discovery - The New Era of Networked Science* (Princeton, NJ and Oxford: Princeton University Press). doi:10.2307/j.ctt7s4vx
- Parolini, G. (2015) From computing girls to data processors: Women assistants in the Rothamsted Statistics Department, in: V. Schafer, and B. Thierry (Eds) *Connecting Women. History of Computing*, pp. 103–117 (Cham: Springer International Publishing). doi:10.1007/978-3-319-20837-4_7
- Pinel, C., Prainsack, B. and McKeivitt, C. (2020) Caring for data: Value creation in a data-intensive research laboratory, *Social Studies of Science*, 50(2), pp. 175–197. doi:10.1177/0306312720906567
- Rigden, D. J. and Fernández, X. M. (2022) The 2022 nucleic acids research database issue and the online molecular biology database collection, *Nucleic Acids Research*, 50(D1), pp. D1–D10. doi:10.1093/nar/gkab1195
- Robbins, R. (1994) Biological databases: A new scientific literature, *Publishing Research Quarterly*, 10(1), pp. 3–27. doi:10.1007/BF02680434
- Russo, F. (2016) On the poietic character of technology, *Humana. Mente Journal of Philosophical Studies*, 9(30), pp. 147–174.
- Sanderson, K. (2011) Bioinformatics: curation generation, *Nature*, 470, pp. 295–296. doi:10.1038/nj7333-295a
- Schiebinger, L. (1989) *The Mind has No Sex? Women in the Origins of Modern Science* (Cambridge, MA: Harvard University Press).
- Schiebinger, L. (1999) *Has Feminism Changed Science?* (Cambridge, MA: Harvard University Press).
- Shapin, S. (1989) The invisible technician. (The role of assistants in the making of science), *American Scientist*, 77(6), pp. 554–563.
- St. Pierre, S. and McQuilton, P. (2009) Inside FlyBase: biocuration as a career, *Fly*, 3(1), pp. 112–114. doi:10.4161/fly.3.1.7769
- Star, S. L. and Strauss, A. (1999) Layers of silence, arenas of voice: The ecology of visible and invisible work, *Computer Supported Cooperative Work*, 8(1-2), pp. 9–30. doi:10.1023/a:1008651105359
- Suchman, L. A. (2007) *Human-machine Reconfigurations: Plans and Situated Actions*, 2nd ed., (Cambridge: Cambridge University Press).
- Tang, Y. A., Pichler, K., Füllgrabe, A., Lomax, J., Malone, J., Munoz-Torres, M. C., Vasant, D. V., Williams, E. and Haendel, M. (2019) Ten quick tips for biocuration, *PLoS Computational Biology*, 15(5), pp. e1006906. doi:10.1371/journal.pcbi.1006906
- Tripathi, S., Vercruyse, S., Chawla, K., Christie, K. R., Blake, J. A., Huntley, R. P., Orchard, S., Hermjakob, H., Thommesen, L., Læg Reid, A. and Kuiper, M. (2016) Gene regulation knowledge commons: community action takes care of DNA binding transcription factors, *Database: The Journal of Biological Databases and Curation*, 2016, pp. baw088. doi:10.1093/database/baw088
- Tronto, J. C. (1993) *Moral Boundaries: A Political Argument for an Ethic of Care* (New York: Routledge).
- Vasilevsky, N. 2021. Addressing implicit or unconscious bias: Equity, diversity and inclusion workshop at the Biocuration2021 virtual conference. Available at <https://www.biocuration.org/addressing-implicit-or-unconscious-bias-equity-diversity-and-inclusion-workshop-at-the-biocuration2021-virtual-conference> (accessed 24 August 2022).

- Wilkinson, M. D., Dumontier, M., Aalbersberg J. I., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J., Da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., Hoen, P. A. C. T., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S., J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., Van Schaik, R., Sansone, S-A, Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J. and Mons, B. (2016) The FAIR guiding principles for scientific data management and stewardship, *Scientific Data*, 3, pp. 160018. doi:[10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18)
- Yavorsky, J. E., Ruggs, E. N. and Dill, J. S. (2021) Gendered skills and unemployed men's resistance to "women's work", *Gender, Work, and Organization*, 28(4), pp. 1524–1545. doi:[10.1111/gwao.12694](https://doi.org/10.1111/gwao.12694)