Abstract
Recent advances in artificial intelligence (AI) applications have raised concerns about the consequences of the uncontrolled development of AI technology for society and humans. Information and knowledge professionals working in research libraries are in professions that have long existed and have globally applied ethical codes that serve as self-regulatory ethical norms. New AI technologies that penetrate throughout libraries’ operations cause confusion among librarians and challenge the existing ethics. In this paper, we examine these challenges and present a qualitative study that reveals the ethical considerations that research librarians face when they approach new AI technologies. As there are no established AI ethics norms for research librarians, we compared the international code of conduct for libraries against the European AI guidelines to identify relevant themes for our study. We analyzed the data from two Scandinavian workshops for librarians. Our findings highlight the central role of research libraries in making AI-powered research ethical. Our study also indicates a need to update international codes of conduct for libraries for the AI age by including aspects of AI agency and the interests of future generations. This helps librarians better orient themselves and their patrons towards a trustworthy and existentially sustainable future with AI systems.

Keywords
Research libraries, academic libraries, artificial intelligence, ethics.

1. Introduction

Recent and impressive advances in large language models have raised serious concerns about the consequences of the uncontrolled and business-led race to develop artificial intelligence (AI) technologies [1]. Following the initial excitement about the endless-seeming possibilities of these human-competing digital systems, many of us have woken up to the risks to society and humans [2–4]. In other words, the voices calling for existential sustainability are becoming louder, and the need for AI ethics is becoming more imperative.

Before the current burst of public discussion, information and knowledge professionals working in research libraries had anticipated the advent of AI in their work contexts. The earliest scenarios for intelligent machine-operated library tasks were drafted in the mid-1970s. In her article, Smith [5] envisioned automated information retrieval systems and used the term AI. In the past decade, interest in AI in libraries has grown exponentially (see Figure 1).
The literature indicates that the research library community approaches this new technology with uncertainty. Many reports demonstrate how AI offers better opportunities than old technologies for improved library operations and services [7]. The library community also recognizes the risks. The power and obscurity of these new tools can be considered a threat to librarians and library patrons (students, researchers, and citizens) if AI generates biases that exploit users or distort research [6]. The ethical implications of AI are slowly alarming the entire scholarly community. The integrity of the research and the explainability and robustness of the research results gained with the help of AI tools are also viewed as threatened [8].

In addition to ethical norms or regulations set by society’s governing institutions, ethics can be maintained at the individual level through professional codes of conduct. Globally conformed ethical codes that grow from and build the self-image of librarians have existed and been adhered to for many years. Since the first formal ethical code for librarians was published in the 1930s [2], the profession has explicitly expressed its values and principles. The most recent code of conduct that unites librarians and information workers across countries and organization types was published by the International Federation of Library Associations and Institutions (IFLA) in 2012 [9]. This document begins by manifesting the ethical requirements for being a librarian: “Librarianship is, in its very essence, an ethical activity embodying a value-rich approach to professional work with information.”

In this paper, we approach the ethical challenges posed by AI for our society by presenting the results of a qualitative study of the existing and potentially emerging ethical norms of one profession—research librarians. The aim is to examine the self-regulative potential of this profession when AI technologies are emerging forcefully throughout library operations and changing librarians’ work. This study is part of action research [10], where the authors have a dual role: While we serve our local research and library communities by developing services that explore the potential of AI technologies, we also research the phenomenon as information science scholars.

Our research question is: What ethical considerations do research librarians obtain when they explore and approach new AI technologies?

We limited our examination to Europe because our empirical data is from Scandinavian research libraries. The European research library community can be considered homogenous because librarians in Europe have similar educational backgrounds and are well-networked. Academic and other research libraries are also well connected through national and international library associations, such
as LIBER, the Association of European Research Libraries, and librarians have the means to learn the ethics of their profession during their education or through library networks.

In this study, we use the concept of AI in the wide meaning that dominates popular and academic discourses. In these discussions, the acronym AI includes methods, technologies, applications, and research approaches—viewpoints from recent advances in machine learning (ML) to futuristic imaginings. As we focus on a period in which the understanding of this new phenomenon in libraries is vague, we consider it relevant to include all of these diverse viewpoints despite their heterogeneous and even contradictory interpretations.

In the following chapters, we present the conduct of our study, starting with a description of our methods. We describe our findings in detail, discuss their implications, and summarise our conclusions.

1.1. Trustworthy and ethical AI

The ethical development and use of AI systems in libraries is grounded on trust. First, the trustworthiness and integrity of AI systems and their output is an ethical premise that entails AI technologies working "for the good of humanity, individuals, societies and the environment and ecosystems" [11]. Second, the trustworthiness of information has been the ethical mainstay of libraries for millennia. The digital shift, or, as Ovenden [12] describes it, the “digital deluge”, poses an existential challenge to the role and mission of libraries.

Trust is principally based on the interest of the trusted [13]. Trust requires a willingness to be vulnerable to another human and to believe that the trustee will fulfil the agreed-upon commitments [9]. As Wheeless and Grotz [14] point out, “trust is antecedent to a willingness to disclose.”

To build trust in AI-powered systems, one needs to show theoretical guarantees, such as algorithmic provenance and dependencies [15,16], so that the outcome is interpretable for non-experts. Lipton [17] argues that trustworthiness is based on how often the output (that is, the model) is correct, and in which context and for which examples it is correct. One way to solve this issue is to present AI-powered outcomes to users with visual or textual artefacts [18].

When using library services, patrons need to trust the quality and reliability of the information provided by a library as an organization and the integrity of its staff [19]. Only then can users disclose their needs and lack of skills. Luckily, libraries still have a high level of trust in society [20,21].

2. Methodology

As there are no established or widely available AI ethics guidelines for research librarians, we began our study by examining two existing norms provided by organizations with a position of authority among European research libraries. The first reveals the viewpoint of libraries, and the second concerns AI technology. Then, to see whether these norms play any role in practice—if librarians have ethical considerations regarding AI in their work—we studied the data from two workshops we conducted in 2022. In these workshops, Scandinavian academic librarians explained and shared their understanding of AI in their work.

In the first part of the study, we compared the contents of two documents that describe current ethical norms: The first document is the *IFLA Code of Ethics for Librarians and Other Information Workers (full version)*, published by IFLA in 2012 (hereafter IFLA Codes) [9]. This document captures and explicates librarians’ ethical norms, some of which have been applied for centuries, such as the moral imperative to provide access to information [12].

The second document that we inspected was the *Ethics guidelines for trustworthy AI*, created by the European Commission (EC) and published in 2019 (hereafter AI Guidelines) [22]. Although these guidelines emphasize the responsibility of AI technology developers, they are also intended to cover the deployment and use of AI systems. As the document addresses “researchers” and “institutions” [22], among other stakeholders, it is also intended for research librarians.

Other relevant guidelines, such as the OECD AI principles [23] and UNESCO’s AI ethics [11], offer recommendations for the public sector and mention (public) libraries as examples. We limited
our inspection to the IFLA and EC documents, as, based on our professional experience, the OECD and UNESCO and their documents lack normative power within the European research library community.

We aimed to find and dissect the key themes and, in particular, those ethical norms that are characteristic of AI technologies. Drawing from recent research [8], we assumed that the essence of contemporary AI applications sets specific requirements for ethical considerations. Our goal was to identify potential differences that may influence librarians’ abilities to consider ethical questions specific to AI technology adoption and use.

We compared the contents of the two documents using an ethnographic content analysis approach, which enables the detection of significance and meaning in a particular context and culture [24]. As an essential element in our action research methodology, we used visual representations to extract relevant concepts, themes, and patterns from the texts, as well as to elaborate our findings. We used the Miro online platform for visual analysis in both parts of the study. As an outcome of the first part of the study, we generated a list of themes that indicated relevant ethical considerations about AI in the context of research libraries.

In the second part of the study, we analyzed the outcome of two workshops conducted in 2022. Altogether, 45 Scandinavian librarians or professionals working for library services participated in these on-site workshops. The goal was to help librarians approach the AI phenomenon that has emerged in their realm and lead to sentiments of a fundamental change. In these workshops, we facilitated the discussions and future planning exercises using designerly approaches and methods. The workshops provided qualitative data in the form of handwritten texts and photographs. We transcribed the participants’ notes on two exercises. The first encouraged participants to think about AI strategy failures and remedies in their libraries (the so-called sabotage method), and the second steered them to plan feasible AI activities in their work (the so-called back-casting method). We analyzed the transcripts and elaborated our analysis visually on the Miro platform.

The second analysis aimed to reveal the kind of ethical considerations that librarians express when they are in the process of exploring and approaching new AI technologies. We used the theme list from the first part of the study as a reference tool and detected equivalent expressions from the workshop data.

3. Findings

3.1. Comparison of ethics guidelines

To reveal whether there are differences in ethical principles specifically for the development and use of AI technologies for librarians, we compared two key documents: the IFLA Codes and the AI Guidelines. The comparison showed that the documents were conformant in many points that addressed the importance of basic ethical norms such as privacy, protection of personal data, and transparency of governance processes. Both documents also emphasize diversity and open access to information for all (see the grey connectors in Figure 2).

We identified the following technology-related ethical aspects from the AI Guidelines that had neither equivalent nor adjacent expressions in the IFLA Codes. Still, we considered that these aspects could be related to existing ethical principles in librarians’ work (see the blue connectors in Figure 2):

- **Robustness and safety**, both technical and social, affect the ethical use of information and service quality. The current IFLA Codes do not mention technical systems and their obvious role as instruments of access, although the IFLA Code number 2 mentions potential “barriers”. There is also no indication that technology affects the “highest standards of service quality,” although the quality depends on the systems’ robustness and safety.

- The highest standards of service quality are also dependent on **responsible mechanisms** that ensure data governance, auditability, and accountability “for AI systems and their outcomes.” This is important because AI systems are inclined toward “unfair bias”.


In the era of AI, the highest service quality also requires an awareness of AI. This means that if libraries are using or promoting AI-powered systems and services, they should inform their users of the existence of AI and take care of traceability mechanisms.

Informing users of AI-powered systems’ capabilities and limitations and educating them to become aware of AI systems can be considered a new aspect of information literacy. This is one of the core services that research libraries provide to their patrons.

We detected three discrepancies between the two ethics documents, which reflect ambivalent value propositions regarding research libraries and AI (see the red connectors in Figure 2):

- Acknowledgement of the agency of AI technology: The IFLA document does not acknowledge the autonomous agency of technology. The IFLA Codes number 2 uses the expression “autonomous users,” but this refers to individual humans who access library services without external help. The first requirement in the AI Guidelines concerns human agency, which can be interpreted as an opponent to AI agency.
- Governance versus citizens’ freedoms: The AI Guidelines imply the importance of governance by society when it uses expressions such as “data governance mechanisms” and “legitimate access to data.” On the other hand, the IFLA document emphasizes a citizen’s freedom through expressions such as “scrutiny of the general public” and “reject… censorship … by states, governments… or civil society institutions.”.
- Interests of future generations: The AI Guidelines number 6 explicitly expresses the importance of addressing the needs of future generations, whereas the IFLA Codes seem to address the needs of current library patrons, employees, and stakeholders. None of the IFLA codes expresses a concern about the future.

We also observed that the final IFLA code does not have an equivalent in the AI Guidelines. This point focuses on collegial relations within a library and a library community, emphasizing fairness and respect for colleagues. However, we interpreted a close connection between this code and other IFLA Codes, as well as the AI Guidelines, as it expresses a general ethical principle: “Librarians and other information workers strive to earn a reputation and status based on their professionalism and ethical behavior.”
### 3.2. Librarians’ considerations

From the first part of the study, we extracted 11 themes that we used to analyze the data from the two workshops. We used these themes to code expressions of ethical considerations from the workshop transcripts. The first column of Table 1 shows the themes (codes), the two columns in the middle indicate the occurrences of coded expressions in the data from the workshops, and the rightmost column provides examples of participants’ notes.

<table>
<thead>
<tr>
<th>Themes</th>
<th>Occurrences in workshop 1</th>
<th>Occurrences in workshop 2</th>
<th>Examples</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Common ethics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>E1 Privacy, data protection and transparency</td>
<td>1</td>
<td>1</td>
<td>• Protect sensitive data</td>
</tr>
<tr>
<td>E2 Diversity, fairness, and open access to all</td>
<td>5</td>
<td>8</td>
<td>• Understand AI in the context of users</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Consider the diversity of library staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Easy access</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Small languages</td>
</tr>
<tr>
<td>E3 Societal well-being</td>
<td>-</td>
<td>3</td>
<td>• Something great for everyone</td>
</tr>
<tr>
<td><strong>Technology-related ethics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>T1 Robustness and safety</td>
<td>10</td>
<td>14</td>
<td>• Thorough testing and analysis</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• AI-competence building</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Unsustainable/sustainable deployment of tools</td>
</tr>
<tr>
<td>T2 Responsible mechanisms</td>
<td>10</td>
<td>49</td>
<td>• Ignore AI in society, bury our heads in the sand</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Explore together with users</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• AI strategy for the library</td>
</tr>
<tr>
<td>T3 Awareness of AI</td>
<td>3</td>
<td>6</td>
<td>• Raise awareness of AI among staff and patrons</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Provide possibilities in the library</td>
</tr>
<tr>
<td>T4 Information on the system’s capabilities and limitations</td>
<td>7</td>
<td>18</td>
<td>• Expect AI to solve research questions</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Ignore limitations or capabilities</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Understand the strengths and weaknesses of AI</td>
</tr>
<tr>
<td><strong>Ambivalent ethics</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>A1 Acknowledgement of AI agency</td>
<td>-</td>
<td>3</td>
<td>• Human control of knowledge</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>• Dedicate time for human learning</td>
</tr>
<tr>
<td>A2 Governance vs. citizen freedom</td>
<td>-</td>
<td>4</td>
<td>• Challenge top-down decisions</td>
</tr>
</tbody>
</table>

Table 1: Analysis of workshop notes
Themes | Occurrences in workshop 1 | Occurrences in workshop 2 | Examples
--- | --- | --- | ---
A3 Interests of future generations | - | - | -
A4 Collegial relations | 4 | 8 | • Maintain research freedom  
• Encourage individual initiatives
• Professional identity with AI  
• AI is for the IT department, not for libraries  
• Generate AI knowledge hubs  
• Collaborate  
• Build networks across boundaries

We identified a few expressions in the thematic category of common ethics, such as a need to protect users’ privacy or sensitive data, the provision of easy access for users, or a need to understand users’ or librarians’ diverse interests and skills. Workshop 2 provided notes we interpreted as an ethical consideration of societal well-being. For example, in its context, the expression “something great for everyone” reflected the same idea as AI Guideline number 6: “Social and societal impact should be carefully considered”.

Most of the notes from both workshops concerned ethical considerations for technology. In the exercise that involved thinking about failures and their counteractions, the workshop participants considered aspects related to AI robustness and safety, responsible mechanisms that ensure the trustworthy deployment of AI technologies, wise distribution of activities between stakeholders, and accountability of processes and their outcomes. There were several suggestions to establish an AI strategy for the participants’ own libraries. It is also worth mentioning that only two notes were made about algorithmic bias in workshop 2.

Under the ambivalent ethics theme, there was a greater dispersion of notes. We could not identify any expressions from workshop 1 that reflected the discrepancies we recognized between the IFLA Codes and the AI Guidelines. However, both workshops resulted in notes with collegial relations with suggestions for building librarians’ professional identities through AI competence and collaboration across organizational boundaries. In workshop 2, participants made notes about human agency over AI, and about maintaining individuals’ freedom in the context of research communities. Again, it is noteworthy that the notes provided no suggestions for future generations, although all the workshop participants were from libraries with cultural heritage collections.

4. Discussion

Research libraries stand for trustworthiness and reliability in their information and knowledge services. Sharing and transparency are among the ethical principles of the profession, as manifested in the IFLA Codes. Exploiting the value of academic libraries as trustworthy institutions in the context of AI has implications for the identities of these organizations.

As our workshop participants noted, librarians should aim to build their own AI literacy to supervise their patrons (see T2 and T3 in Table 1). This competence could then be used to highlight AI tools’ possibilities and ethical pitfalls. For the time being, few libraries can offer this kind of service. A set of ethical principles on how to govern the human use of AI in different contexts and the implications thereof have emerged from several institutions and groups [25]. Auditability, satisfaction, effectiveness, persuasiveness, efficiency, and trust are values that underpin transparency and explanations of results in AI-based systems [26].

This is the critical moment for research libraries to build their AI strategies, not least because of the ethical considerations. If libraries do not take an active role in AI ecosystems and remain passive
users of AI-powered services, they may become obsolete. Academic and other research libraries must
decide whether to keep their role as mediators of trustworthy information based on ethical principles
(see E1 and E2 in Table 1). Various observations from our study pinpoint transparency, open access
to information, and responsible mechanisms as the most important ethical values to consider (see E1,
E2, and T2 in Table 1). These observations conform with voices from the field calling to safeguard
libraries’ values [3,27,28].

Another observation is that librarians must acquire new skills and competencies to cope with
ethical issues when using AI-powered tools and providing AI-enhanced services. These may include
the copyright of the output of a service, possible hidden biases in training data for algorithms, and an
understanding of where training data originates (see T3 and T4 in Table 1). New competencies have
become more critical because the entire knowledge ecosystem, including research communities,
research libraries, and academic publishers, is in the middle of AI transformation. While research
libraries are probing their role with AI technologies, academic publishers are improving their
production processes and services using the power of algorithms [29]. Software companies are also
capitalizing on AI technologies and marketing their innovations to the wide academic community,
from individual researchers to national libraries—a fact also mentioned by participants of workshop
number 2: “[A road to failure:] Forget that most AI tools are commercial”.

While ethics and trust are closely related in libraries, they often remain undiscussed because they
are taken-for-granted components of a librarian’s identity. In general, research libraries need to
consider ethical problems when using AI “so that research libraries will continue to serve as trusted
advisors to our users, and as responsible collectors, disseminators, and preservers of knowledge” [30].
Algorithmic bias is an intensively debated issue since libraries are brokers of a large amount of data,
both their own and that produced by university users [8].

Finally, we observe that neither the IFLA Codes nor the librarians who attended our workshops
considered the interests of future generations. A growing body of research on societal and cultural
sustainability challenges existing work practices, including in libraries. Research libraries can be
considered fortresses of cultural sustainability if they comply with “development that meets the needs
of the present without compromising the ability of future generations to meet their own needs” [31].
To reach this ethical norm, librarians need to consider their future patrons. It may not be feasible to
anticipate the expectations of several generations ahead, but consideration of the next generation
would be an ethical act.

5. Conclusion

In this study, we examined the ethical challenges AI poses in the context of research libraries and
from the perspectives of research librarians. We examined two ethics norms that provide an
understanding of established library ethics, on the one hand, and amended AI ethics, on the other.
From the comparison of these normative documents, we extracted 11 themes for ethical consideration.
Then, we analyzed the empirical data from two Scandinavian workshops and reflected the data
against these themes.

Our findings indicate that the context of libraries is crucial to maximizing the ethical
trustworthiness of AI-powered services for research communities. Librarians will need to guarantee
the reliability of the knowledge provided by AI systems. Our findings also show that librarians want
to understand the “black box” where AI “magic” happens, and involving them during future projects
is mandatory.

In the past year, there has been growing concern in the research library profession about the ethical
implications of new AI-based tools. It has been observed that these tools are sometimes utilized in
ways that may compromise research integrity. This notion highlights the need for caution and ethical
considerations. Lack of trust and insecurity in how these new tools work underpin the necessity of
ethical codes for libraries. Existing AI guidelines and regulations need to be revised. International
library institutions, such as the IFLA and the LIBER, should lead their professional communities with
up-to-date norms. However, as it is necessary to consider regional differences, we encourage research
libraries to establish local AI strategies.
Our findings also indicate a need to update library ethics and codes of conduct to meet the needs of the age of AI. While the existing main principles align with general AI ethics, and the core mission to provide open and fair access to information withstands, librarians should also consider independently acting algorithms as new users of library services. Future generations will be the judges, beneficiaries, or victims of today’s ethical decisions. Anticipating their needs is also an ethical consideration worth doing.

6. Acknowledgements

We thank all the workshop participants who contributed to this research.

7. References


