



# Rethinking Health Recommender Systems for Active Aging: An Autonomy-Based Ethical Analysis

Simona Tiribelli<sup>1,2</sup> · Davide Calvaresi<sup>3</sup>

Received: 8 November 2021 / Accepted: 2 April 2024  
© The Author(s) 2024

## Abstract

Health Recommender Systems are promising Artificial-Intelligence-based tools endowing healthy lifestyles and therapy adherence in healthcare and medicine. Among the most supported areas, it is worth mentioning active aging. However, current HRS supporting AA raise ethical challenges that still need to be properly formalized and explored. This study proposes to rethink HRS for AA through an autonomy-based ethical analysis. In particular, a brief overview of the HRS' technical aspects allows us to shed light on the ethical risks and challenges they might raise on individuals' well-being as they age. Moreover, the study proposes a categorization, understanding, and possible preventive/mitigation actions for the elicited risks and challenges through rethinking the AI ethics core principle of autonomy. Finally, elaborating on autonomy-related ethical theories, the paper proposes an autonomy-based ethical framework and how it can foster the development of autonomy-enabling HRS for AA.

**Keywords** Health recommender systems · AI ethics · Autonomy · Active aging · Persuasion technology

## Abbreviations

HRS Health recommender systems  
ML Machine learning  
AA Active aging

---

✉ Simona Tiribelli  
simona.tiribelli@unimc.it

Davide Calvaresi  
davide.calvaresi@hevs.ch

<sup>1</sup> Department of Political Sciences, Communication, and International Relations, University of Macerata, 62100 Macerata, Italy

<sup>2</sup> Institute for Technology and Global Health, PathCheck Foundation, 955 Massachusetts Ave, Cambridge, MA 02139, USA

<sup>3</sup> University of Applied Sciences and Arts Western Switzerland (HES-SO), Rue de l'Industrie 23, 1950 Sion, Switzerland

|     |                                |
|-----|--------------------------------|
| DL  | Deep-learning                  |
| AI  | Artificial intelligence        |
| CF  | Collaborative filtering        |
| PH  | Physical                       |
| CB  | Content-based filtering        |
| PE  | Personal                       |
| HR  | Hybrid recommendation          |
| PT  | Persuasion technologies        |
| KB  | Knowledge-based recommendation |
| RS  | Recommender systems            |
| EC  | Ethical challenges             |
| SO  | Sociale                        |
| CO  | Cognitive                      |
| WHO | World Health Organization      |
| HCP | Healthcare providers           |

## Introduction

Advances in digital technology and Artificial Intelligence (AI) are fostering the evolution of digital healthcare systems. In this context, continuous and ubiquitous support is enhancing personalization and preventive patient-centered care (Calvaresi et al., 2017a). The users of such systems are supposed to progressively become more autonomous, adopt healthy behaviors, and maintain themselves active and healthier over time (De Croon et al., 2021). Such processes can entail a user's behavioral change (Orji & Moffatt, 2018). To support it, health recommender systems (HRS), often leveraging persuasion technologies (PT)—i.e., rule-based behavioral strategies –, are increasingly considered the most promising approaches (De Croon et al., 2021; Hunter, 2018). In particular, active aging (AA) is a domain in which PT-enabled HRS can play a crucial role. Indeed, aging individuals (approaching frailty) will inevitably face growing health-related matters as they deal with more complex needs (e.g., chronic conditions management), metabolic changes (involving dietary adaptations), reduced resources/information accessibility, possible cognitive degradation, etc. Therefore, assistive technologies (including HRS) target aging individuals as main users/service-consumers (Calvaresi et al., 2017a; Torres et al., 2023; De Croon et al., 2021; Oliva-Felipe et al., 2018). However, to date, there is no unique and reconciling definition of AA. Nevertheless, its multi-faceted nature is generally agreed upon. Indeed, HRS are widely acknowledged as supporting the several dimensions of AA (Fernández-Ballesteros et al., 2013). The underlying contributions span over counteracting loneliness (Holwerda et al., 2012), promoting social support (Tomaka et al., 2006), enhancing social/family relationships (Victor et al., 2000), facilitating social activity and fostering a sense of personal fulfillment (Creecy et al., 1985), strengthening healthy habits (McPhee et al., 2004)—such as smoking cessation and better nutrition (Pirlich & Lochs, 2001), including healthy active motor-habits (Cvecka et al., 2015), stimulating cognitive

activities (Tardif & Simard, 2011), and treating/managing diseases and pathologies (e.g., diabetes and cancer) (Longo et al., 2010; Manzo et al., 2021). Indeed, AA goes beyond the health dimension of well-being in aging. It concerns individual (e.g., well-being derived from living a meaningful life), socio-relational perspectives (e.g., well-being related to participating in social life and being in meaningful relationships), and security (feeling safe) (Fernández-Ballesteros et al., 2013; WHO, 2023).

Elaborating on AA's key elements and dimensions outlined above from an HRS design perspective, we claim that HRS *should* contribute to four aspects (i.e., physical, personal, social, and cognitive) to foster AA. With respect to such aspects, HRS contribute as follows.

*Physical* (PH)—provide knowledge-based health-related motor stimuli and recommendations tailored to the user's needs and capabilities while preventing human deskilling; *personal* (PE)—ensure a fair set of meaningful choices and options that do not constrain or erode user's freedom of choice; *social* (SO)—promote interactions among individuals in similar conditions to foster motivation, support, and empathy, preventing self-isolation and social pressure; and *cognitive* (CO)—provide user-friendly, intuitive, and pervasive interfaces and interactions that do not understimulate nor overwhelm the user perception, elaboration, and actuation capabilities (most relevant causes for early tech abandon).

The effectiveness and efficiency of current HRS can also come at the expense of the users' autonomy (see Sect. "Rethinking HRS Challenges Based on the Ethics Principle of Autonomy"). Indeed, although HRS can empower users' autonomy in AA, especially in terms of self-control and independence, they can also raise ethical risks and concerns that can undermine AA. An example can be the unfair reduction of meaningful opportunities due to biased correlations and socio-relational exposure, possibly leading to alienation/isolation (Pariser, 2011; Knees et al., 2024; Valentine et al., 2023). Such risks can only be understood through a more complex *multidimensional* concept of autonomy (Prunkl, 2022; Tiribelli et al., 2023b). This paper (i) investigates HRS and proposes (ii) an understanding of the AI ethics principle of autonomy to elicit and circumvent the threats affecting AA, and (iii) possible mitigation strategies.

The remainder of the paper is organized as follows. Section "Health Recommender Systems: Technical Aspects and Ethical Challenges" briefly introduces HRS' main technical features and their role in health behavioral change in AA. In turn, it highlights potential ethical risks that HRS can generate for individuals pursuing health-related goals and contextualizes them into the AA's key aspects. Section "Rethinking HRS Challenges Based on the Ethics Principle of Autonomy" draws on ethical theory to show how an adequate understanding of the AI ethics principle of autonomy helps to categorize, prevent, and/or mitigate such ethical risks and design HRS for AA. Section "From Theory to Practice: Enacting Autonomy in Next-Gen HRS" shows how to rethink the next-generation HRS via the ethical principle of autonomy and how the latter helps to consider a series of strategies to counteract HRS-related autonomy threats. Finally, Sect. "Conclusions" concludes the paper.

## Health Recommender Systems: Technical Aspects and Ethical Challenges

According to De Croon et al. (2021), recommender systems (RS) are software tools and techniques providing suggestions for items of interest for a given user. We experience RS daily, receiving suggestions affecting several decision-making processes, such as items to buy, music to listen to, and news to read. HRS are a specialization of RS. As they operate in the health context, HRS recommend items of interest that generally consist of non-confidential, scientifically proven, or at least generally accepted health-related and/or medical information.

Conventional RS have boosted, among others, marketing, e-commerce, and overall consumerist behaviors (Calvaresi et al., 2022). However, ingrained in modern life, commercial RS have remarkably contributed to the development of unhealthy habits (i.e., excessive consumption of energy/sugar-dense food and alcohol and use of unhealthy sets of recipes) that can last over time (Trattner & Elswelier, 2017; Calvaresi et al., 2022). HRS should fight such tendencies by promoting prevention, awareness, and overall healthy behaviors (De Croon et al., 2021). Indeed, HRS mainly operate in lifestyle, nutrition, general health, and condition-specific application scenarios (De Croon et al., 2021). Overall, HRS span from wellness and well-being to medical application scenarios. Clearly, the implications and responsibilities vary remarkably in such a broad spectrum. Nevertheless, they can all be concerned when focusing on HRS applications in AA. For instance, there are HRS designed to enable AA by counteracting elderly social isolation (e.g., by providing multi-modal coaching for social activities (Dimitrov et al., 2019)), improving nutrition habits (e.g., providing tracking functionalities and “guiding” recommendations to observe a good macro-nutrients balance) (Espín et al., 2016; Calvaresi et al., 2021b), promoting psycho-physical activity (e.g., proposing general-purpose and personalized physical tasks (Nassabi et al., 2014)), including the cognitive activity (e.g., for equilibrium maintenance: HRS that provide reminders and personalized exercises to stimulate muscles and cognition skills (Calvaresi et al., 2021)), and allowing tele-rehabilitation (e.g., HRS providing monitoring and guidance to recover a good range of motion, such as in the case of older individuals who have undergone lower-limbs surgery) (Buonocunto et al., 2018). Finally, several solutions have tried to provide longitudinal (multi-dimension) contributions, employing argumentation and rewarding mechanisms (Herpich et al., 2017), fighting addictions (e.g., smoking cessation) from both a social and physical health perspectives - by providing support in the monitoring and the craving phase (Calvaresi et al., 2019), and via automated (AI-based) and manual (formulated by doctors) recommendations and guidelines (e.g., for cancer survivors (Manzo et al., 2021)).

From a technological perspective, RS and HRS partially share some core elements. For example, both RS and HRS can rely on symbolic-based reasoners (i.e., rule-based state machines), data-driven engines (i.e., machine learning/deep learning—ML/DL models), and hybrid infrastructures leveraging both the approaches (i.e., via neuro-symbolic integration (Spillo et al., 2022)).

Moreover, RS can be non-personalized (i.e., not requiring any prior knowledge about any specific user (Falk, 2019) or personalized (i.e., requiring a remarkable amount of knowledge about the targeted user). Instead, HRS are intended to be strongly personalized due to profile- and context-crucial information. Non-personalized RS leverage generic information (i.e., items' popularity, novelty, price, and distance) to sort the possible recommendations. For example, basket modeling & analysis is used in retail to identify complementarity (i.e., items to recommend as usually bought together) (Yıldırım et al., 2020; Lin et al., 2002). Personalized (H)RS leverage users' details, including demographics, location (i.e., geofencing), language, preferences, and possibly aggregated behaviors to have a deeper understanding of the user (Schäfer et al., 2017; Logesh et al., 2019). Such data are usually obtained by tracking and profiling the users (Wang et al., 2021; Iwata et al., 2007). Here, *rating* is a popular mechanism to determine users' preferences and qualify an item explicitly or implicitly (if the preference is deduced from the user's behavior) (Guo et al., 2014; Oard et al., 1998). Moreover, as summarized in Calvaresi et al. (2021b), the most adopted data-driven techniques in HRS are: *collaborative filtering*(CF), which consists of leveraging users' similarities and ratings (Jannach et al., 2010); *content-based filtering*(CB), which consists of recommending similar items based on similar profiles' previously liked items Sánchez Bocanegra, Sánchez Laguna, and Sevillano (2015); *knowledge-based recommendation*(KB), where the recommendation is based on the user preferences and constraints (Felfernig & Burke, 2008), and *hybrid recommendation* (HR), which combines the techniques mentioned above (Ricci et al., 2011)). In short:

- CF: Uses the collected data to identify similarities between items and user profiles and, in turn, computes the expected ratings of unseen items (Aggarwal et al., 2016; Chen et al., 2018). CF algorithms can be classified as *memory-based* (Karabadjji et al., 2018), *model-based* (matrix factorization, tensor completion) (Luo et al., 2014), *hybrid CF* (combining model-based and memory-based) (Zhang et al., 2017), and *deep learning-based* (Wei et al., 2017) approaches. For example, Sahoo et al. (2019) utilized CF to correlate patients (i.e., to propose physical exercises) based on their personal interests, feedback on therapies, etc.).
- CB: Uses the items' underlying characteristics to provide new recommendations—mostly suitable if the user is directly interested in them (Wang et al., 2018). For example, CB-based HRS are used to support health consumers looking for multiple sources when searching for health information online (Longo et al., 2010).
- KB: Conceals the knowledge about a given domain—mostly suitable if variability and personalization are broad (requiring both domain and item-specific knowledge) (Tarus et al., 2018). For example, it is employed by systems

monitoring emotional health to detect users with potential psychological disturbances (i.e., depression and stress) and to send happy, calm, relaxing, or motivational messages contextually (Rosa et al., 2018).

**HR:** Combines two or more approaches. For example, in a CF approach, the dependency on ratings entails disfavoring unrated items. Nevertheless, combining CF and KB approaches (if the items contain attributes) could solve the limitation (Dong et al., 2017). For example, it is employed by applications (i.e., smartphone-based virtual assistants) that process a broad range of vital/wellness parameters to understand if contacting a doctor or performing further analysis is recommended (Jamshidi et al., 2018).

Overall, although hybrid approaches seem to be the most promising option from current literature in the field, to date, in practice, RS mainly rely on data-driven approaches, and HRS mainly rely on rule-based finite state machines—mainly given the complexity and sensitivity of the input data and effects on the users.

Such a tendency is particularly motivated by the presence of PT within the HRS reasoning engine. It is worth recalling that persuasion is intended as “an activity that involves one party trying to induce another party to believe or to do something” (Hunter, 2018). In HRS, the persuasion usually targets abandoning unhealthy behaviors in favor of what is suggested (what is considered healthy and a good solution for the given user condition). Persuasion strategies are mostly formulated as symbolic connections (i.e., state machines) that ensure more control of the process. Moreover, it is necessary to distinguish PT from other forms of influence on freedom of choice, which cannot be considered legitimate. Unlike manipulation or coercion, persuasion affects the architecture of choices by providing pre-known (possibly extended) alternatives without undermining freedom of choice (Carli et al., 2022).

While such techniques have been shown to be particularly functional in triggering a behavioral change in health and for HRS’ tasks, as outlined (Calvaresi et al., 2022), such techniques can also raise a number of ethical risks and challenges that, if untackled, can undermine individuals’ well-being—in particular in aging conditions. The ethical risks underlying such challenges indeed become increasingly relevant with respect to those raised by conventional RS (Milano et al., 2020), considering their high-sensitive application field.

As follows, we highlight some of the major ethical challenges that HRS, as increasingly prototyped as HR in current literature (De Croon et al., 2021), ought to address to ensure their ethical development (i.e., their design and implementation to ensure people’s well-being). Then, we show how such ethical challenges (EC), if unaddressed, can undermine AA’s key dimensions previously outlined. HRS should:

**EC1:** Avoid inaccurate recommendations. HRS must be benevolent. This means that HRS target users’ health and therefore provide suggestions that must not harm the users’ health. An example can be recommending too complex or

advanced physical exercises, just leveraging user self-reporting inputs, insofar as they can lead to serious injuries. Benevolence includes accuracy with respect to individuals' core beliefs and values.

- EC2: Ensure privacy: the generation of personalized health recommendations entails access to (sensitive) personal information. As mentioned above, CF has been shown to be vulnerable to data leakage in the inference phase (Calandrino et al., 2011; Ji et al., 2020). HRS are strongly characterized by a trade-off between recommendations' accuracy and how much the users are willing to share about themselves. The current research direction is a layered notion of privacy per user group (Xin & Jaakkola, 2014). For example, systems leveraging third-party services might run into data leakage, exposing data to partially authorized personnel (i.e., within a given HRS, a doctor must only be allowed to visualize the info of patients who gave them specific consent) or using personal data as not initially declared/intended.
- EC3: Safeguard personal identity and freedom of choice: HRS might (a) intentionally restrict users' freedom of choice with biased recommendations and/or a reduced (sub)set of options and (b) reshape the users' community, leading to echo chambers and filter bubbles—a risk to which are exposed HRS providing social-integration/support features (Kampik et al., 2018). Examples can be (i) the over-simplifying of the food items list to be recommended, (ii) vehiculating the community's interest towards third-party products/services pushing for mere consumeristic interests, and (iv) proposing a recommendation (medically correct) that might counter the user's moral/cultural values.
- EC4: Reduce the HRS opacity: current HRS engines employ ML/DL predictors "as is" and do not provide transparency on their internal mechanisms (i.e., no info about how/why a given recommendation is/has been produced). Such opacity instills mistrust and a lack of accountability (Graziani et al., 2022; Anjomshoae et al., 2019). For example, as of today, users cannot ask how/why a given recommendation (e.g., medication, food item, activity, therapy, nearby shop, or restaurant) is proposed and have no power over it.
- EC5: Avoid unfairness: skewed and biased data sets and inappropriate contextual constraints can yield unfair recommendations. To mention a few, a user might receive recommendations (i) (un)intentionally committed to brands, activities, or doctors, including getting proposed over-the-counter drugs brands—possibly more expensive—over others less known/publicized—and possibly cheaper—on a regular basis), (ii) that are unattainable due the possible user's limited access to healthcare services or information, physical disabilities, lack of internet access, or language barriers, and (iii) that suggest solutions that are insensitive of the user's culture.

**Table 1** Mapping active aging aspects—HRS ethical challenges

| AA aspects | HRS ECs                      |
|------------|------------------------------|
| PH         | EC1, EC3                     |
| PE         | EC1, EC2, EC3, EC4, EC5, EC6 |
| SO         | EC1, EC5, EC6                |
| CO         | EC1, EC2, EC5                |

Unfortunately, the identification, isolation, and exclusion of such biases to avoid unfair recommendations are strongly hampered by the opacity of today's HRS and their related training mechanisms.

- EC6: Avoid social pressure and socio-relational impoverishment: as already mentioned in EC3, polarization is considered one of the most dangerous side-effects of using HRS (Knijnenburg et al., 2012; Konstan & Riedl, 2012)—unfortunate (H)RS byproduct since their early adoption (Van Alstyne et al., 1996). Social networks have exacerbated (H)RS' effects and set themselves as an (almost unique) source of news and information, dramatically boosting filter-bubbles (Bozdag & Van Den Hoven, 2015) and echo-chambers (Levy & Razin, 2019) and worsened the already existing polarization (i.e., societal). For example, aging individuals coping with smoking cessation campaigns that include social aspects conducted over social media are extremely emotionally vulnerable. Hence, they are dramatically exposed to manipulation/coercion. Moreover, HRS can end up strengthening socio-relational isolation and impoverishment. For example, promoting excessively isolating activities—possibly due to users' mobility restrictions—can lead users to loneliness and malaise and harm AA's social dimension.

Finally, Table 1 presents a possible mapping of AA aspects with the elicited ethical challenges. In particular,

- PH: HRS might push the user beyond their capabilities and safety boundaries to reach a goal. Indeed, many HRS rely on self-reported and possibly missing data, potentially presenting a more "optimistic" picture than reality. Therefore, naive or overly mechanistic approaches can endanger the user rather than promoting their best interest. Moreover, having limited or unbalanced sets of practices (solely focusing on given tasks or entirely ignoring others) can cause the user's progression to drift in unforeseen/unwanted directions.
- PE: Recommendations might include elements asking users to violate or infringe their personal values and beliefs. This can occur if the user model disregards some elements (i.e., religion, nutritional orientation, social commitments,



and societal customs) or if the user is misclassified. On the one hand, using users' sensitive information might violate personal standing. For example, most users usually agree to share some or all of their information to improve the system or enhance medical understanding. Nevertheless, data misuse could represent a risk of fostering commercial interests at the user's expense (extending/shifting from what they have originally agreed with). On the other hand, assuming coherency in the suggestions, the user might not perceive the rationale behind them, and current HRS fail to engage in argumentative interactions to explain the system's recommendations and decisions. Eventually, users can come to question HRS' fairness (i.e., hidden promotion of consumeristic interests—brands over others—via recommendation), and, to date, HRS provides no exhaustive answers or clarifications regarding their decision-making.

- SO: Personal information can be used by HRS to foster users' adherence (e.g., via group chats over social media or gamification—leveraging the social factor). However, despite the possible positive outcomes obtained in most cases, more attention must be paid to shielding specific aspects that outlier users would like to protect rather than share. Indeed, such information might be misused by individuals of the same “community/society” against their owners, inducing self-isolation rather than the hoped/expected integration.
- CO: HRS can provide both physical and mental assistance. From a cognitive perspective, some tasks, procedures, and work/data flows must be more flexible and adaptable to a specific audience. For example, the language, formulation, and length of the consent forms are often more confusing than informative, leading the user to blindly accept them to be able to use the system's functionalities. Once again, such formulations are static and generic (one-size-fits-all) and are not processed by all users in the same way.

## **Rethinking HRS Challenges Based on the Ethics Principle of Autonomy**

We have outlined that HRS can be a valid technological ally supporting AA. Indeed, several studies mentioned above have shown that HRS can provide tangible support for several AA-related tasks and activities (e.g., enabling healthier habits and health-and-care self-management) (De Croon et al., 2021; Torres et al., 2023; Calvarresi et al., 2017a). However, we have also shown that HRS can raise several ethical risks and challenges. Such risk exposure can jeopardize the HRS benefits, producing effects damaging individuals' well-being, especially during aging. Thus, it seems crucial to understand how to address the ethical challenges (EC1-to-EC6) and, therefore, prevent the aforementioned ethical risks to fully harness HRS potential in AA. Indeed, as previously outlined, AA concerns people's well-being in aging beyond

the sole health domain. Rather than considering HRS as a *panacea* solving all the issues related to aging, it is vital to adopt a balanced perspective on HRS in terms of their opportunities and downsides to understand their role in supporting AA.

Therefore, the question that arises concerns how HRS engineers can design such tools to take advantage of their value for AA while preventing the ethical risks mentioned. Also, insofar as ethics is concerned, it sounds legitimate to ask whether it is possible to rely on ethical principles, guidelines, and/or criteria to tackle the ethical risks and challenges outlined above.

In this section, we claim that HRS ethical challenges can be understood as challenges to the AI ethics principle of autonomy if the latter is adequately understood. Specifically, we show how autonomy requires diverse sub-conditions to be met, whose consideration helps to categorize such risks and challenges. Therefore, we will unpack the ethical principle of autonomy, drawing insights from ethical theory rooted in moral philosophy. By doing so, we will show how our ethical inquiry can offer an autonomy-based ethical framework to steer the design of near-future HRS that prevent and/or mitigate the above-mentioned ethical risks and challenges and eventually promote AA.

### **Why is Autonomy Crucial in HRS for AA?**

Autonomy plays a central role in AA. As pointed out by the World Health Organization (WHO, 2023), the maintenance and the promotion of autonomy are crucial for enabling AA. Overall, autonomy is referred to as (i) the ability to control and make personal decisions about one's life according to one's own rules and preferences (autonomy as control and self-determination), (ii) independence - that is, the capacity to live independently in the community, with no or little help from others (WHO, 2023; Paúl et al., 2012), (iii) connected to the quality of life - as people age, their quality of life is determined by their ability to maintain autonomy and independence, where the quality of life also includes physical health, psychological state, personal beliefs, and social relationships (Paúl et al., 2012), and by (iv) the healthy life expectancy (how long people can expect to live autonomously), that is, without disabilities (p. 3).

Moreover, autonomy is recognized as a paramount issue in the age-tech field based on AI and algorithms (Ho, 2020; Rubeis, 2020; Rubeis et al., 2018; Sixsmith et al., 2013). From this literature, a widely embraced idea is to design self-tracking and personalized health systems by considering, on the one hand, how they can promote the autonomy of the elderly (Ho, 2020), especially in terms of engagement, self-management, and self-determination on health-related issues (Sharon, 2017; Müller et al., 2010; McLean et al., 2013; Vassilev et al., 2015; Topol, 2015), and on the other hand, how they can undermine it—by raising risks such as profit-driven health surveillance and/or social pressure, users' informed consent bypassing, and privacy infringements (Lupton, 2013; Rubeis, 2020).

Similar ethical challenges are also at the core of the nascent debate on the ethics of RS (Milano et al., 2020; Varshney, 2020). Here, autonomy is mainly understood

as human self-determination and is claimed to be the core ethical principle for the future design of ethical RS. This is because the benefits, main opportunities, and risks raised by RS are considered to be intertwined with autonomy (Varshney, 2020). Indeed, RS can empower people's autonomy as they help users to navigate informational overload saving time and cognitive resources. This is also relevant for HRS. For instance, consider HRS used as personal health virtual assistants (Varshney, 2020) or as embedded in web-based platforms to access health information/recommendations. According to Sezgin and Özkan (2013), such tools have a significant role in filtering information for users' self-diagnostic searches. Some contextual examples proposed in such a study are HRS used by physicians for diagnostic and educational purposes, such as HRS suggesting online health resources (HealthyHarlem) and HRS-based educational resources with patient records (MyHealthEducator). Other examples include HRS in mHealth apps connecting users with doctors or healthcare providers (HCP) using filtering options based on people's tracked profiles. Other examples in the telemedicine field include HRS used to recommend health treatments in support of medical diagnoses, as in the case of the UK Babylon health app (Iacobucci, 2020; Tiribelli et al., 2023b). Such opportunities are even more relevant to enhancing the autonomy of aging individuals (with weakening cognitive skills). However, such users both benefit from HRS as well as might be more affected by the residual ethical risks they pose. As highlighted in studies (Milano et al., 2020; Varshney, 2020), the risks raised by RS are also strictly intertwined with autonomy (Whittlestone et al., 2019; Varshney, 2020). Such risks range from human deskilling or cognitive atrophy (independence loss), due to tasks' over-delegation, to social homogenization and cultural diversity impoverishment. The latter happens when RS suggest similar or identical options/items for certain segments of populations and thus reducing the qualitative diversity of options and diversified social exposure that are critical for adequate and genuine decision-making (Fleder & Hosanagar, 2009; Tiribelli, 2023) — creating filtering or epistemic bubbles (Pariser, 2011).

Overall, the principle of autonomy is central to the design of RS. Indeed, to be accurate and effective, RS cannot avoid intervening in some ways on users' autonomy to some extent (Cowls et al., 2019), such as reducing options or extracting users' crucial information to contextualize recommendations. Such consideration is pivotal for RS used in the health industry and for health or medical purposes. As personal health information is key to developing accurate recommendations, the risks of undermining users' autonomy increase - especially for already vulnerable groups (Milano et al., 2020; Varshney, 2020; Mittelstadt et al., 2016). Nonetheless, people might be willing to take such risks by voluntarily choosing to use such technology and consciously ceding/renouncing some of their autonomy in exchange for other potential benefits stemming from such use.

In addition to being recognized as a key principle for AA (WHO, 2023), autonomy is one of the most acknowledged ethical principles in the field of ethics of AI and algorithms (Jobin et al., 2019) in general and as applied to health (Guidance,

2021)<sup>1</sup>. While the relevance of autonomy in the fields intersected by our inquiry is evident, some outline how such a core AI ethics principle still tends to be vague or undefined. In other words, it is not clear what the respect and promotion of autonomy truly demand for the ethical design of AI (Prunkl, 2022) and RS (Prunkl, 2022).

In summary, autonomy is acknowledged as a crucial ethical principle for AA, as well as for the design of AI and RS (in general) or health RS (i.e., HRS). However, the lack of its adequate conceptualization hampers its concrete operationalization. Autonomy runs deeply in moral philosophy as any other ethical principle (Beauchamp, 2019). Therefore, conceptualize it properly, that is, as an ethical principle, asks to rely on moral philosophy (Bietti, 2020; Giovanola & Tiribelli, 2022a, 2022). The latter indeed helps us to shed light on its multidimensionality and bridge the gap between high-level theory and practical feasibility, namely, to prevent ethics principles from being misunderstood or used as mere labels and therefore becoming trivial, useless, and toothless (Munn, 2022; Rességuier & Rodrigues, 2020).

### The Ethics Principle of Autonomy

As previously mentioned, the current AI ethics principle of autonomy is often vague and opaque. Jobin et al. (2019) provided evidence of this vagueness by analyzing more than 80 ethical frameworks for AI. From this analysis, the ethical principle of autonomy emerges as a core one and it is mainly understood as (i) informational self-determination, (ii) privacy-preserving human control over AI systems, (iii) freedom to withdraw consent, and (iv) freedom from exploitation, manipulation, and surveillance (p.11). Consequently, the most commonly adopted methods of operationalizing it consist of “giving people notice and consent”, “refraining from data collecting and spreading data in the absence of informed consent”, “not reducing options and knowledge of citizens”, “increasing people’s knowledge of AI”, and ensuring “transparency and predictable AI” (p.11). While all these definitions are pertinent, such fragmentation hinders the understanding of what autonomy consists of, what it demands, and how it can be empowered or violated by AI systems. We claim that this opacity is due to the lack of systematic ethical inquiry into the concept of autonomy, which is a notoriously highly complex ethical concept rooted in moral philosophy. To untangle this complexity, in this section, we draw insights on theories on autonomy developed in moral philosophy to unpack the concept of autonomy, highlighting the dimensions it encompasses and the sub-conditions it requires.

In moral philosophy, a prominent notion—also known as the standard conception—of autonomy is that developed within the liberal tradition (Christman & Anderson, 2005). Notwithstanding some conceptual variations, the liberal view

---

<sup>1</sup> It is important to highlight that autonomy is an ethical principle that tends to be prominent in Western cultures more than in other parts of the world. However, recent contributions have highlighted the centrality of autonomy, and especially of relational autonomy (see Sect. 3.2), also in non-Western perspectives (e.g., Ubuntu Ethics), showing the importance of analyses that consider it beyond its mainstream Western understanding as individual self-determination and independence for a truly inclusive and ethical design of AI-based systems (Mhlambi & Tiribelli, 2023).

on autonomy mainly describes it as rational self-government, control, and independence. From this perspective, autonomy is rooted in the individuals' capacity to choose and act according to their own interests, reasons, beliefs, and values. More specifically, autonomy is respected and promoted if and only if people can embrace or endorse (i.e., reflectively assess, approve or identify with, or refuse and modify) what guides their choices and, by doing so, self-determine themselves. As follows, the liberal or traditional view on autonomy tends to overlap with self-determination, which in turn is grounded in the individuals' rational and deliberative capacity to be in control of their actions (or self-govern themselves), insofar as such actions are chosen according to motives that are somehow their own (Frankfurt, 1971; Dworkin, 1988; Ekstrom, 1993; Bratman, 2017; Korsgaard, 2014; Pink, 2016). However, while there is a broad agreement on what autonomy consists of, there is less convergence on what autonomy requires. Within this debate, at least two main conditions are mainly understood as crucial to meeting autonomy: competency conditions and authenticity conditions (Killmister, 2017).

*Competency conditions* refer to a set of cognitive capacities (including psychophysical abilities) that individuals should own to self-govern or control themselves, which include rational thought to freedom from pathology or systematic self-deception. To foster autonomy, such capacities ought to be respected and/or supported. For example, in the medical field, these conditions are those that enable the individual, once they are provided with certain health information, to be able to act on the basis of that (e.g., information x "smoke damages your health, by affecting a, b, c, etc." requires competency as the capacity for self-control). In the field of AI, such conditions can be undermined directly when AI decreases such capacities. Some examples in the AI and ethics debate on autonomy's infringements in this sense include using AI such as RS to enable systematic self-deception (Susser et al., 2019; Natale, 2021), or to steer human behavior according to third-party goals by exploiting cognitive biases or non-rational elements (e.g., health vulnerabilities, traumas, or emotionally-loaded events) (Tiribelli, 2023). Competencies can be violated by AI systems indirectly when there is a human deskilling or loss of capacity due to over-reliance on AI for tasks and decisions or when the information on which one should act is not intelligible to the end-user.

*Authenticity conditions* refer instead to the individuals' capacity to form authentically or genuinely (i.e., to critically reflect upon and endorse) the reasons, values, interests, preferences, and beliefs motivating their actions. This means that such beliefs, reasons, values, and motivations should not be the product of external manipulative or constraining influences but should be embraced and reflectively endorsed by the subjects as expressing their personal identity. Like competency, authenticity also centers the element of self-reflection, but in a more substantial way.

Competency conditions have been indeed criticized both in moral philosophy as well as in applied ethics, especially in bioethics and medical ethics, as expressing just a formal or procedural account of autonomy. Indeed, they focus solely on the procedure and cognitive means through which a person can come to rationally endorse certain options (i.e., reasons, values, preferences, etc.). Exercising that set of skills is deemed enough to be autonomous. In this sense, they center self-reflection but in a way that is described as procedural independence: they just focus on

the procedure while neglecting the importance of questioning whether the options from which a specific person chooses embed what that person (i.e., preferences, values, reasons, etc.) deems truly meaningful. As it has been claimed widely by scholars from feminist ethics (Mackenzie & Stoljar, 2000a, b; Oshana, 2006; Benson, 2005), overlooking the substance in favor of the procedure can lead to easily legitimizing options that embed (historically rooted) values of oppression like unfair and discriminating biases and thereby perpetuating social conditions that are inherently autonomy-undermining as oppressive (i.e., inherently curtail certain options), especially for the most vulnerable and marginalized. Moreover, such conditions fail to account for people who live in constraining health and/or socio-economic conditions, such as those affected by epistemic injustice (Fricker, 2007) and/or physical decline or debilitating pathology but anyway reclaim their autonomy and related dignity (Jaworska, 2009).

This is why competency conditions are not the sole to be considered. The conditions of authenticity express the possibility of the individuals to be autonomous if and only if they can identify with those beliefs, values, preferences, and reasons steering their behavior, endorsing them as motives of their choices and actions (Mackenzie, 2014). In this sense, autonomy is respected and promoted if the options from which a person can choose embed what (value, commitment, project, etc.) that person feels as deeply meaningful (i.e., substantial self-reflection and independence). Looking at the literature on the ethics of AI, such conditions are described to be undermined when AI and RS are used to manipulate individuals (Milano et al., 2020). This is why autonomy is often formalized as “freedom from manipulation and undesired external influences” (Jobin et al., 2019), and is often strictly connected to privacy (Mittelstadt et al., 2016). Such a state of “freedom from interferences” is mainly understood as the protection of individuals’ meaningful space of independence, that is, their informational privacy (Floridi, 2011). Such conditions can also be undermined when AI is used to trigger phenomena such as adaptive preference or belief formation (e.g., RS that pre-select contents or options can lead people to adapt their preferences to match the options available to them) (Adomavicius et al., 2013). Moreover, such conditions can be harmed when the—filtered or recommended—options available are not truly meaningful for the individuals as a result of biased correlations and stereotypes, resulting in a loss of meaningful opportunities.

However, authenticity conditions are also broadly criticized within moral philosophy and ethical theory. Overall, liberal accounts of autonomy as procedural and/or substantial independence are criticized by scholars coming from feminist ethics and other ethical traditions such as communitarianism (Oshana, 2006; Mackenzie & Stoljar, 2000a, b; Westlund, 2009; Gutmann, 1985; Sandel, 2005; Taylor, 1992), as focusing only on an individualistic dimension of autonomy, where the individuals are viewed as independent and self-isolated. Such criticism is particularly strong in an applied ethics field such as medical ethics, in which self-sufficiency and independence sound to be at odds with conditions of assistance, care, support, vulnerability, and dependency that characterize the domains of health and care. Specifically, such scholars emphasize the importance of investigating some further conditions of autonomy in social relations, which they argue are crucial to enable people’s autonomy (in general) and patients’ autonomy (in particular) (Entwistle et al., 2010).

Those who advocate for recognizing such relational dimension of autonomy are broadly recognized under the umbrella term of “relational autonomy” (Mackenzie & Stoljar, 2000a, b). Beyond some conceptual differences, the core idea underpinning this view is that authenticity can only occur in social conditions, as it is through the socio-relational (interpersonal) and cultural dimensions that people develop what is meaningful for them and motivates their agency.

Some authors investigated such relational dimension (hereinafter: *relatedness*) of autonomy and found in social recognition and social support some specific **socio-relational conditions** that *ought to* be enabled to respect and promote autonomy (Anderson & Honneth, 2005; Benson, 2005; Westlund, 2009; Govier, 1993). This is because social recognition and support are crucial to motivate one’s own agency, especially when conditions of oppression of various kinds (e.g., related to pathology and aging) can threaten one’s autonomy by undermining one’s own sense of self-confidence, self-trust, and self-respect required for effective agency. Looking at the current literature on AI and RS, such systems can undermine these conditions when RS are designed in ways that lead the individual to self-isolate. If we think about HRS, recommending strict health regimes (e.g., nutritional plans) raises the risk of hampering participation in social and recreational activities. Such conditions can also be undermined when the options available to the subject are filtered or recommended in ways that create homogenization and a lack of socio-cultural diversity (Pariser, 2011). For instance, there are many stereotypical or poorly informed views about older adults, their needs and resources, and their desired way of life (Rubeis et al., 2022). Suppose HRS are designed according to such ageist stereotypes or biases. In that case, HRS can wrongly treat older people as a homogeneous group/category, reducing exposure to socio-relational aged heterogeneity (Fang et al., 2018; Taipale & Hänninen, 2018), hampering the possibility of meeting social support truly felt as meaningful.

Summarizing, through our inquiry into the concept of autonomy, we have shown its multidimensionality (i.e., autonomy is not a black-and-white scenario) and, specifically, the multiple conditions autonomy requires to meet (i.e., protect and promote). In doing so, we have systematically framed the multiple definitions of the principle of autonomy emerging in AI ethics and have clarified that autonomy requires (i) competency, (ii) authenticity, and (iii) relatedness. Additionally, we have shown how each condition can be undermined, clarifying how the ethical challenges raised by HRS can be understood as challenges to the principle of autonomy. Table 2 organizes the sub-conditions required to enable autonomy in relation to the ethical challenges raised by HRS and the key areas of AA to consider in designing HRS. The next section shows how to use this autonomy-based ethical framework for the design of HRS for AA.

## From Theory to Practice: Enacting Autonomy in Next-Gen HRS

The previous

**Table 2** Ethical framework to enable autonomy in the design of HRS for AA

| The AI ethics principle | Dimensions               | Sub-conditions              | Ethical challenges         | AA's aspects in HRS |
|-------------------------|--------------------------|-----------------------------|----------------------------|---------------------|
| Autonomy                | Procedural independence  | Competency conditions       | EC1, EC2, EC3<br>EC4, EC5  | PH, CO              |
|                         | Substantial independence | Authenticity conditions     | EC1, EC2, EC3,<br>EC4, EC5 | PE                  |
|                         | Relatedness              | Socio-relational conditions | EC3, EC6                   | SO                  |

section showed that the ethical challenges raised by HRS to individuals' well-being and AA can be understood as challenges to individuals' autonomy. In turn, the importance of understanding what autonomy as an ethics principle demands was emphasized, clarifying its key dimensions and sub-conditions, and how they relate to AA's key aspects. Although HRS are still in their infancy, their applications for general health and medical purposes are growing, along with the sectors of AI-empowered eHealth and telemedicine. Hybrid HRS will soon move from the literature to widespread reality, and ethical analyses and frameworks for their trustworthy development will become pivotal (Etemadi et al., 2022). The protection and promotion of autonomy is not just an extra benefit or an after-the-fact consideration but rather the main goal when designing HRS-based systems to support AA. Therefore, we deem the time is mature to start thinking about next-gen autonomy-enacting HRS. In this section, we briefly propose a few practical strategies to enact autonomy in HRS for AA via the respect and promotion of its key dimensions and sub-conditions.

Respecting and promoting competency conditions (e.g., cognitive and physical capabilities) are crucial when designing HRS to support aging individuals (EC1–EC4). From a practical standpoint (i.e., HRS design), the system must be user-centered and able to understand the user's conditions and tune its interactions, graphical interfaces, and actual contents accordingly. In particular, functionalities and interactions should be carefully modeled on the "specific" user (Tiribelli et al., 2023a) and not be excessively generalized, over-simplified (risking to bore the user and cause an early system abandon), over-complicated (instigating rejection and refusal), or over-stimulating, endangering the user psyche. Similarly, the informed consent must be adapted to the user's capabilities so it can be properly processed and understood.

The respect and promotion of authenticity (i.e., availability of actual meaningful options) are particularly important for users' well-being during aging. The insurgence of pathologies can hinder individuals' possibilities to enjoy activities or pursue personal values, beliefs, and personal/social commitments they feel meaningful. To mitigate this issue, HRS might include (i) a space for users to share meaningful information about *how* they want to live their life. Due to the sensitivity of such information and its vulnerability to third-party manipulation, such space should ensure (ii) confidentiality (Tiribelli, 2023), that is, a zone of



informational privacy (EC2). This means that what is shared by the users cannot be used for nudging/manipulating them for consumeristic gains. Near-future HRS could elaborate on such meaningful elements, seeking to develop personalized health-centered paths and plans, while avoiding behaviors/side effects harming their meaningful values and preferences (EC1, EC3). When a user's preferences are at odds with the health goal, HRS should provide users-tailored, intelligible explanations with educational advice to empower the user (EC4). Moreover, such explanations should rely on verifiable mechanisms and sources, leveraging honest/sincere (truth-telling) assumptions, data, and rules. However, sometimes, HRS harbor commercial interests, possibly/inadvertently entailing questionable uses of health-related information. In such cases, as they operate in the health domain, HRS should be subject to duties of transparency and disclosure. Recommendations might also be vulnerable to bias (EC5), and, especially when based on macro-generalizing correlations, they can end up being tailored on stereotypes that neglect formalized categories' inner heterogeneity. On the one hand, to counter such risk, HRS should include sufficiently articulated user profiling and proper variability of the short-term plans, which are functional to the achievement of long-term objectives. On the other hand, HRS should avoid over-constraining—risking finding no solutions. Therefore, reaching a granularity balance between specification and generalization is imperative.

Ensuring relatedness conditions is pivotal to meeting autonomy, especially for aging individuals, who tend to self-isolate due to the insurgence of pathologies and debilitation. Despite the positive aspects discussed above, connecting people in similar conditions might also lead to epistemic bubbles and exacerbate negative feelings, potentially leading to more isolation and social impoverishment (Nguyen, 2020). Moreover, support offered through health technologies often involves affiliates, caregivers, or HCP and might imply forms of health surveillance or external control, mainly perceived as social pressure. To mitigate such risks, HRS might include mechanisms allowing the users to (i) select their most meaningful inner circles, (ii) provide the possibility to opt out of a given social environment/situation, and (iii) provide feedback, reports, and asking for psychological/social support.

The HRS applications can be considerably heterogeneous and far aside. Therefore, the considerations highlighted above do not presume to be exhaustive. Nevertheless, applying an autonomy-based framework to specific HRS (e.g., for medical purposes) would enable further tailoring (i.e., cognitive-dependent) of the given strategies to increase their context sensitivity and thus expected efficacy.

## Conclusions

This study has investigated the potential of HRS to support AA and their exposure to ethical threats. By analyzing the off-the-shelf technical literature on HRS, we have highlighted that HRS can promote AA if they are functional in four AA's aspects (i.e., Physical, Personal, Social, and Cognitive). Moreover, we have shown that near-future PT-based HRS deploy a series of techniques. On the

one hand, PT-based HRS enable the prevention and personalization of health-care. On the other hand, they can raise a series of ethical challenges and risks undermining users' well-being and AA's aspects (spanning over users' privacy and freedom of choice to HRS' accuracy, fairness, and opacity). Elaborating on that, we have suggested addressing such challenges through the AI ethics principle of autonomy (a fundamental tenet of AA, AI, and RS in health). We have then pursued an ethical investigation into the concept of autonomy and showed how it requires conditions of (i) competency, (ii) authenticity, and (iii) relatedness. We have argued how the risks highlighted might hamper such conditions—henceforth, users' autonomy in aging. By unpacking such an ethical principle, we have proposed an ethical taxonomy or framework of autonomy — filling a current gap in the AI and ethics literature. Finally, we have shown how to operationalize such a framework (autonomy's dimensions and sub-conditions) by proposing possible practical strategies to design next-gen autonomy-enabling HRS for AA.

Notwithstanding, the HRS currently in use are still in their early stages, characterized by a limited usage of ML and DL. This is particularly the case within the European Union, where regulations for health tech are stricter. However, HRS tend to evolve rapidly. We believe ethically steering HRS' design to harness their opportunity while counteracting their risks, especially for users' autonomy in aging, is impelling. Therefore, developing robust ethics for HRS is time-critical, and we hope that the proposed autonomy-based ethical framework helps to pave the way for the upcoming efforts.

**Funding** This work has been funded by the European Union - NextGenerationEU under the Italian Ministry of University and Research (MUR) National Innovation Ecosystem grant ECS00000041 - VITALITY - CUP E13C22001060006, and partially supported by the Chist-Era grant CHIST-ERA-19-XAI-005, and by the Swiss National Science Foundation (G.A. 20CH21\_195530).

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Adomavicius, G., Bockstedt, J. C., Curley, S. P., & Zhang, J. (2013). Do recommender systems manipulate consumer preferences? A study of anchoring effects. *Information Systems Research*, 24(4), 956–975.
- Aggarwal, C. C., et al. (2016). *Recommender systems*. Springer.
- Anderson, J. H., & Honneth, A. (2005). Autonomy, vulnerability, recognition, and justice. In J. Christman & J. Anderson (Eds.), *Autonomy and the challenges to liberalism: New essays* (pp. 127–149).
- Anjomshoae, S., Najjar, A., Calvaresi, D., & Främling, K. (2019). Explainable agents and robots: Results from a systematic literature review. In *18th international conference on autonomous agents and*

- multiagent systems (AAMAS 2019)* (pp. 1078–1088). Montreal, Canada, May 13–17, 2019. International Foundation for Autonomous Agents and Multiagent Systems.
- Beauchamp, T. L., & Childress, J. F. (2019). *Principles of biomedical ethics*. Oxford University Press.
- Benson, P. (2005). Feminist intuitions and the normative substance of autonomy. In J. Stacey Taylor (Ed.), *Personal autonomy: New essays on personal autonomy and its role in contemporary moral philosophy* (pp. 124–142). Cambridge University Press.
- Bietti, E. (2020). From ethics washing to ethics bashing: A view on tech ethics from within moral philosophy. In *Proceedings of the 2020 conference on fairness, accountability, and transparency* (pp. 210–219).
- Bozdag, E., & Van Den Hoven, J. (2015). Breaking the filter bubble: Democracy and design. *Ethics and Information Technology*, 17(4), 249–265.
- Bratman, M. E. (2017). Rational planning agency. *Royal Institute of Philosophy Supplement*, 80, 25–48. <https://doi.org/10.1017/s1358246117000042>
- Buonocunto, P., Giantomassi, A., Marinoni, M., Calvaresi, D., & Buttazzo, G. (2018). A limb tracking platform for tele-rehabilitation. *ACM Transactions on Cyber-Physical Systems*, 2(4), 1–23.
- Calandrino, J. A., Kilzer, A., Narayanan, A., Felten, E. W., & Shmatikov, V. (2011). “you might also like:” Privacy risks of collaborative filtering. In *2011 IEEE symposium on security and privacy* (pp. 231–246). IEEE.
- Calvaresi, D., Calbimonte, J.-P., Dubosson, F., Najjar, A., & Schumacher, M. (2019). Social network chatbots for smoking cessation: Agent and multi-agent frameworks. In *2019 IEEE/WIC/ACM international conference on web intelligence (WI)* (pp. 286–292). IEEE.
- Calvaresi, D., Calbimonte, J.-P., Siboni, E., Eggenschwiler, S., Manzo, G., Hilfiker, R., & Schumacher, M. (2021). Erebots: Privacy-compliant agent-based platform for multi-scenario personalized health-assistant chatbots. *Electronics*, 10(6), 666.
- Calvaresi, D., Carli, R., Piguët, J.-G., Contreras, V. H., Luzzani, G., Najjar, A., Calbimonte, J.-P., & Schumacher, M. (2022). Ethical and legal considerations for nutrition virtual coaches. *AI and Ethics*, 3, 1–28.
- Calvaresi, D., Cesarini, D., Sernani, P., Marinoni, M., Dragoni, A. F., & Sturm, A. (2017). Exploring the ambient assisted living domain: A systematic review. *Journal of Ambient Intelligence and Humanized Computing*, 8(2), 239–257.
- Calvaresi, D., Eggenschwiler, S., Calbimonte, J.-P., Manzo, G., & Schumacher, M. (2021b). *A personalized agent-based chatbot for nutritional coaching* (pp. 682–687). WI-IAT '21, New York, NY, USA. Association for Computing Machinery. Retrieved from <https://doi.org/10.1145/3486622.3493992>.
- Carli, R., Najjar, A., & Calvaresi, D. (2022). Risk and exposure of xai in persuasion and argumentation: The case of manipulation. In *International workshop on explainable, transparent autonomous agents and multi-agent systems* (pp. 204–220). Springer.
- Chen, R., Hua, Q., Chang, Y.-S., Wang, B., Zhang, L., & Kong, X. (2018). A survey of collaborative filtering-based recommender systems: From traditional methods to hybrid methods based on social networks. *IEEE Access*, 6, 64301–64320.
- Christman, J., & Anderson, J. (2005). *Autonomy and the challenges to liberalism: New essays*. Cambridge University Press.
- Cowls, J., King, T., Taddeo, M., & Floridi, L. (2019). Designing AI for social good: Seven essential factors. Available at SSRN 3388669.
- Creecy, R. F., Berg, W. E., & Wright, R. (1985). Loneliness among the elderly: A causal approach. *Journal of Gerontology*, 40(4), 487–493.
- Cvecka, J., Tirpakova, V., Sedliak, M., Kern, H., Mayr, W., & Hamar, D. (2015). Physical activity in elderly. *European Journal of Translational Myology*, 25(4), 249.
- De Croon, R., Van Houdt, L., Htun, N. N., Štiglic, G., Abeele, V. V., Verbert, K., et al. (2021). Health recommender systems: Systematic review. *Journal of Medical Internet Research*, 23(6), e18035.
- Dimitrov, Y., Gospodinova, Z., Wheeler, R., Žnidaršič, M., Ženko, B., Veleva, V., & Miteva, N. (2019). Social activity modelling and multimodal coaching for active aging. In *Proceedings of the 12th ACM international conference on pervasive technologies related to assistive environments* (pp. 608–615).
- Dong, X., Yu, L., Wu, Z., Sun, Y., Yuan, L., & Zhang, F. (2017). A hybrid collaborative filtering model with deep structure for recommender systems. In *Proceedings of the AAAI conference on artificial intelligence* (Vol. 31).

- Dworkin, G. (1988). The theory and practice of autonomy. *Social Philosophy and Policy*, 20(2), 70–98.
- Ekstrom, L. W. (1993). A coherence theory of autonomy. *Philosophy and Phenomenological Research*, 53(3), 599–616.
- Entwistle, V. A., Carter, S. M., Cribb, A., & McCaffery, K. (2010). Supporting patient autonomy: The importance of clinician–patient relationships. *Journal of General Internal Medicine*, 25(7), 741–745.
- Espín, V., Hurtado, M. V., & Noguera, M. (2016). Nutrition for elder care: A nutritional semantic recommender system for the elderly. *Expert Systems*, 33(2), 201–210.
- Etemadi, M., Abkenar, S. B., Ahmadzadeh, A., Kashani, M. H., Asghari, P., Akbari, M., & Mahdipour, E. (2022). A systematic review of healthcare recommender systems: Open issues, challenges, and techniques. *Expert Systems with Applications* (pp. 118823).
- Falk, K. (2019). *Practical recommender systems*. Manning.
- Fang, M., Siden, E., Korol, A., Demestihis, M.-A., Sixsmith, J., & Sixsmith, A. (2018). A scoping review exploration of the intended and unintended consequences of Ehealth on older people: A health equity impact assessment. *Human Technology*, 14(3), 297.
- Felfernig, A., & Burke, R. (2008). Constraint-based recommender systems: Technologies and research issues. *ACM International Conference Proceeding Series* (p. 3).
- Fernández-Ballesteros, R., Robine, J. M., Walker, A., & Kalache, A. (2013). Active aging: A global goal. *Current Gerontology and Geriatrics Research*, 2013(3), 298012.
- Fleder, D., & Hosanagar, K. (2009). Blockbuster culture's next rise or fall: The impact of recommender systems on sales diversity. *Management Science*, 55(5), 697–712.
- Floridi, L. (2011). The informational nature of personal identity. *Minds and Machines*, 21(4), 549–566.
- Frankfurt, H. (1971). Freedom of the will and the concept of a person. *Journal of Philosophy*, 68(1), 5–20. <https://doi.org/10.2307/2024717>
- Fricker, M. (2007). *Epistemic injustice: Power and the ethics of knowing*. Oxford University Press.
- Giovanola, B., & Tiribelli, S. (2022a). Beyond bias and discrimination: Redefining the AI ethics principle of fairness in healthcare machine-learning algorithms. *AI & Society*, 1–15.
- Giovanola, B., & Tiribelli, S. (2022). Weapons of moral construction? On the value of fairness in algorithmic decision-making. *Ethics and Information Technology*, 24(1), 1–13.
- Govier, T. (1993). Self-trust, autonomy, and self-esteem. *Hypatia*, 8(1), 99–120.
- Graziani, M., Dutkiewicz, L., Calvaresi, D., Amorim, J. P., Yordanova, K., Vered, M., Nair, R., Abreu, P. H., Blanke, T., Pulignano, V., et al. (2022). A global taxonomy of interpretable AI: Unifying the terminology for the technical and social sciences. *Artificial Intelligence Review*, 1–32.
- Guidance, W. (2021). Ethics and governance of artificial intelligence for health. *World Health Organization*.
- Guo, G., Zhang, J., Thalmann, D., Basu, A., & Yorke-Smith, N. (2014). From ratings to trust: An empirical study of implicit trust in recommender systems. In *Proceedings of the 29th annual ACM symposium on applied computing* (pp. 248–253).
- Gutmann, A. (1985). Communitarian critics of liberalism. *Philosophy & Public Affairs*, 14(3), 308–322.
- Herpich, M., Rist, T., Seiderer, A., & André, E. (2017). Towards a gamified recommender system for the elderly. In *Proceedings of the 2017 international conference on digital health* (pp. 211–215).
- Ho, A. (2020). Are we ready for artificial intelligence health monitoring in elder care? *BMC Geriatrics*, 20(1), 1–7.
- Holwerda, T. J., Beekman, A. T., Deeg, D. J., Stek, M. L., van Tilburg, T. G., Visser, P. J., Schmand, B., Jonker, C., & Schoevers, R. A. (2012). Increased risk of mortality associated with social isolation in older men: Only when feeling lonely? Results from the Amsterdam study of the elderly (Amstel). *Psychological Medicine*, 42(4), 843–853.
- Hunter, A. (2018). Towards a framework for computational persuasion with applications in behaviour change. *Argument & Computation*, 9(1), 15–40.
- Iacobucci, G. (2020). Babylon health holds talks with “significant” number of NHS trusts. *BMJ*.
- Iwata, T., Saito, K., & Yamada, T. (2007). Modeling user behavior in recommender systems based on maximum entropy. In *Proceedings of the 16th international conference on world wide web* (pp. 1281–1282).
- Jamshidi, S., Torkamani, M. A., Mellen, J., Jhaveri, M., Pan, P., Chung, J., & Kardes, H. (2018). A hybrid health journey recommender system using electronic medical records. In *HealthRecSys@RecSys* (pp. 57–62).

- Jannach, D., Zanker, M., Felfernig, A., & Friedrich, G. (2010). *Recommender systems: An introduction*. Cambridge University Press.
- Jaworska, A. (2009). Caring, minimal autonomy, and the limits of liberalism. In H. Lindemann, M. Verkerk, & M. U. Walker (Eds.), *Naturalized bioethics: Toward responsible knowing and practice*. Cambridge University Press.
- Ji, Y., Sun, A., Zhang, J., & Li, C. (2020). A critical study on data leakage in recommender system offline evaluation. arXiv preprint [arXiv:2010.11060](https://arxiv.org/abs/2010.11060).
- Jobin, A., Ienca, M., & Vayena, E. (2019). The global landscape of AI ethics guidelines. *Nature Machine Intelligence*, 1(9), 389–399.
- Kampik, T., Nieves, J. C., & Lindgren, H. (2018). Coercion and deception in persuasive technologies. In *20th international trust workshop (co-located with AAMAS/IJCAI/ECAI/ICML 2018)* (pp. 38–49). Stockholm, Sweden, 14 July, 2018. CEUR-WS.
- Karabadjji, N. E. I., Beldjoudi, S., Seridi, H., Aridhi, S., & Dhifli, W. (2018). Improving memory-based user collaborative filtering with evolutionary multi-objective optimization. *Expert Systems with Applications*, 98, 153–165.
- Killmister, S. (2017). *Taking the measure of autonomy: A four-dimensional theory of self-governance*. Routledge.
- Knees, P., Neidhardt, J., & Nalis, I. (2024). *Recommender systems: Techniques, effects, and measures toward pluralism and fairness* (pp. 417–434). Springer. Retrieved from <https://doi.org/10.1007/978-3-031-45304-527>
- Knijnenburg, B. P., Willemsen, M. C., Gantner, Z., Soncu, H., & Newell, C. (2012). Explaining the user experience of recommender systems. *User Modeling and User-Adapted Interaction*, 22(4), 441–504.
- Konstan, J. A., & Riedl, J. (2012). Recommender systems: From algorithms to user experience. *User Modeling and User-Adapted Interaction*, 22(1), 101–123.
- Korsgaard, C. M. (2014). The normative constitution of agency. In M. Vargas & G. Yaffe (Eds.), *Rational and social agency: The philosophy of Michael Bratman* (pp. 190–214). Oxford University Press.
- Levy, G., & Razin, R. (2019). Echo chambers and their effects on economic and political outcomes. *Annual Review of Economics*, 11, 303–328.
- Lin, W., Alvarez, S. A., & Ruiz, C. (2002). Efficient adaptive-support association rule mining for recommender systems. *Data Mining and Knowledge Discovery*, 6(1), 83–105.
- Logesh, R., Subramaniaswamy, V., Vijayakumar, V., & Li, X. (2019). Efficient user profiling based intelligent travel recommender system for individual and group of users. *Mobile Networks and Applications*, 24(3), 1018–1033.
- Longo, D. R., Schubert, S. L., Wright, B. A., LeMaster, J., Williams, C. D., & Clore, J. N. (2010). Health information seeking, receipt, and use in diabetes self-management. *The Annals of Family Medicine*, 8(4), 334–340.
- Luo, X., Zhou, M., Xia, Y., & Zhu, Q. (2014). An efficient non-negative matrix-factorization-based approach to collaborative filtering for recommender systems. *IEEE Transactions on Industrial Informatics*, 10(2), 1273–1284.
- Lupton, D. (2013). The digitally engaged patient: Self-monitoring and self-care in the digital health era. *Social Theory & Health*, 11(3), 256–270.
- Mackenzie, C. (2014). Three dimensions of autonomy: A relational analysis. In A. Veltman & M. Piper (Eds.), *Autonomy, oppression and gender pages* (pp. 15–41). Oxford University Press.
- Mackenzie, C. & Stoljar, N. (2000a). Introduction: Autonomy refigured. In *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*. Oxford University Press.
- Mackenzie, C. & Stoljar, N. (2000b). *Relational autonomy: Feminist perspectives on autonomy, agency, and the social self*. Oxford University Press.
- Manzo, G., Calvaresi, D., Jimenez-del Toro, O., Calbimonte, J.-P., & Schumacher, M. (2021). Cohort and trajectory analysis in multi-agent support systems for cancer survivors. *Journal of Medical Systems*, 45(12), 1–10.
- McLean, S., Sheikh, A., Cresswell, K., Nurmatov, U., Mukherjee, M., Hemmi, A., & Pagliari, C. (2013). The impact of tele health care on the quality and safety of care: A systematic overview. *PLoS One*, 8(8), e71238.
- McPhee, S. D., Johnson, T. R., & Dietrich, M. S. (2004). Comparing health status with healthy habits in elderly assisted-living residents. *Family and Community Health*, 27(2), 158–169.

- Mhlambi, S., & Tiribelli, S. (2023). Decolonizing AI ethics: Relational autonomy as a means to counter AI harms. *Topoi*, 42(3), 867–880. <https://doi.org/10.1007/s11245-022-09874-2>
- Milano, S., Taddeo, M., & Floridi, L. (2020). Recommender systems and their ethical challenges. *AI & Society*, 35(4), 957–967.
- Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. *Big Data & Society*, 3(2), 2053951716679679.
- Müller, A., Schweizer, J., Helms, T. M., Oeff, M., Sprenger, C., & Zugck, C. (2010). Telemedical support in patients with chronic heart failure: Experience from different projects in Germany. *International Journal of Telemedicine and Applications*, 2010.
- Munn, L. (2022). The uselessness of AI ethics. *AI and Ethics*, 1–9. <https://doi.org/10.1007/s43681-022-00209-w>
- Nassabi, M. H., Vollenbroek-Hutten, M., et al. (2014). An ontology-based recommender system to promote physical activity for pre-frail elderly. *Mensch & Computer 2014-Workshopband*.
- Natale, S. et al. (2021). Deceitful media: Artificial intelligence and social life after the turing test.
- Nguyen, C. T. (2020). Echo chambers and epistemic bubbles. *Episteme*, 17(2), 141–161. <https://doi.org/10.1017/epi.2018.32>
- Oard, D. W., Kim, J., et al. (1998). Implicit feedback for recommender systems. In *Proceedings of the AAAI workshop on recommender systems* (Vol. 83, pp. 81–83). WoUongong.
- Oliva-Felipe, L., Barrué, C., Cortés, A., Wolverson, E., Antomarini, M., Landrin, I., Votis, K., Paliokas, I., & Cortés, U. (2018). Health recommender system design in the context of caregiverspro-mmd project. In *Proceedings of the 11th pervasive technologies related to assistive environments conference* (pp. 462–469).
- Orji, R., & Moffatt, K. (2018). Persuasive technology for health and wellness: State-of-the-art and emerging trends. *Health Informatics Journal*, 24(1), 66–91.
- Oshana, M. (2006). *Personal autonomy in society*. Routledge.
- Pariser, E. (2011). *The filter bubble: What the internet is hiding from you*. Penguin.
- Paúl, C., Ribeiro, O., & Teixeira, L. (2012). Active ageing: An empirical approach to the who model. *Current Gerontology and Geriatrics Research*, 2012.
- Pink, T. (2016). *Self-determination: The ethics of action* (Vol. 1). Oxford University Press.
- Pirlich, M., & Lochs, H. (2001). Nutrition in the elderly. *Best Practice & Research Clinical Gastroenterology*, 15(6), 869–884.
- Prunkl, C. (2022). Human autonomy in the age of artificial intelligence. *Nature Machine Intelligence*, 4(2), 99–101.
- Rességuier, A., & Rodrigues, R. (2020). AI ethics should not remain toothless! A call to bring back the teeth of ethics. *Big Data & Society*, 7(2), 2053951720942541.
- Ricci, F., Rokach, L., Shapira, B., & Kantor, P. (2011). *Introduction to recommender systems handbook*. Springer.
- Rosa, R. L., Schwartz, G. M., Ruggiero, W. V., & Rodríguez, D. Z. (2018). A knowledge-based recommendation system that includes sentiment analysis and deep learning. *IEEE Transactions on Industrial Informatics*, 15(4), 2124–2135.
- Rubeis, G. (2020). The disruptive power of artificial intelligence ethical aspects of gerontechnology in elderly care. *Archives of Gerontology and Geriatrics*, 91, 104186.
- Rubeis, G., Fang, M. L., & Sixsmith, A. (2022). Equity in agetechnology for ageing well in technology-driven places: The role of social determinants in designing AI-based assistive technologies. *Science and Engineering Ethics*, 28(6), 1–15.
- Rubeis, G., Schochow, M., & Steger, F. (2018). Patient autonomy and quality of care in telehealth care. *Science and Engineering Ethics*, 24(1), 93–107.
- Sahoo, A. K., Pradhan, C., Barik, R. K., & Dubey, H. (2019). Deepreco: Deep learning based health recommender system using collaborative filtering. *Computation*, 7(2), 25.
- Sanchez-Bocanegra, C. L., Sanchez-Laguna, F., & Sevillano, J. L. (2015). Introduction on health recommender systems. *Methods in Molecular Biology*, 1246, 131–46.
- Sandel, M. (2005). *Liberalism and the limits of justice*. Routledge.
- Schäfer, H., Hors-Fraile, S., Karumur, R. P., Calero Valdez, A., Said, A., Torkamaan, H., Ulmer, T., & Trattner, C. (2017). Towards health (aware) recommender systems. In *Proceedings of the 2017 international conference on digital health* (pp. 157–161).
- Sezgin, E., & Özkan, S. (2013). A systematic literature review on health recommender systems. In *2013 E-health and bioengineering conference (EHB)* (pp. 1–4). IEEE.

- Sharon, T. (2017). Self-tracking for health and the quantified self: Re-articulating autonomy, solidarity, and authenticity in an age of personalized healthcare. *Philosophy & Technology*, 30(1), 93–121.
- Sixsmith, A., Gutman, G., et al. (2013). *Technologies for active aging*. Springer.
- Spillo, G., Musto, C., De Gemmis, M., Lops, P., & Semeraro, G. (2022). Knowledge-aware recommendations based on neuro-symbolic graph embeddings and first-order logical rules. In *Proceedings of the 16th ACM conference on recommender systems* (pp. 616–621).
- Susser, D., Roessler, B., & Nissenbaum, H. (2019). Online manipulation: Hidden influences in a digital world. *Georgetown Law Technology Review*, 4, 1–45.
- Taipale, S., & Hänninen, R. (2018). More years, more technologies: Aging in the digital era. *Human Technology*, 14(3), 258.
- Tardif, S. & Simard, M. (2011). Cognitive stimulation programs in healthy elderly: A review. *International Journal of Alzheimer's Disease*, 2011.
- Tarus, J. K., Niu, Z., & Mustafa, G. (2018). Knowledge-based recommendation: A review of ontology-based recommender systems for e-learning. *Artificial Intelligence Review*, 50(1), 21–48.
- Taylor, C. (1992). *The ethics of authenticity*.
- Tiribelli, S. (2023). *Moral freedom in the age of artificial intelligence*. Mimesis International.
- Tiribelli, S., et al. (2023a). *Identità personale e algoritmi. una questione di filosofia morale*.
- Tiribelli, S., Monnot, A., Shah, S. F., Arora, A., Toong, P. J., & Kong, S. (2023). Ethics principles for artificial intelligence-based telemedicine for public health. *American Journal of Public Health*, 113(5), 577–584.
- Tomaka, J., Thompson, S., & Palacios, R. (2006). The relation of social isolation, loneliness, and social support to disease outcomes among the elderly. *Journal of Aging and Health*, 18(3), 359–384.
- Topol, E. (2015). *The patient will see you now: The future of medicine is in your hands*. Basic Books.
- Torres, J., Garcia, M., Artola, G., Garcia-Navarro, T., Amaya, I., Larburu, N., & Martin, C. (2023). Well-being recommender system, a user-centered framework for generating a recommender system for healthy aging. In *Proceedings of the 9th international conference on information and communication technologies for ageing well and e-health (ICT4AWE 2023)* (pp. 118–125).
- Trattner, C., & Elsewiler, D. (2017). Investigating the healthiness of internet-sourced recipes: Implications for meal planning and recommender systems. In *Proceedings of the 26th international conference on world wide web* (pp. 489–498).
- Valentine, L., D'Alfonso, S., & Lederman, R. (2023). Recommender systems for mental health apps: Advantages and ethical challenges. *AI & Society*, 38(4), 1627–1638.
- Van Alstyne, M., Brynjolfsson, E., et al. (1996). Electronic communities: Global village or cyberbalkans. In *Proceedings of the 17th international conference on information systems* (p. 32). Wiley.
- Varshney, L. R. (2020). Respect for human autonomy in recommender systems. arXiv preprint [arXiv:2009.02603](https://arxiv.org/abs/2009.02603).
- Vassilev, I., Rowsell, A., Pope, C., Kennedy, A., O'Cathain, A., Salisbury, C., & Rogers, A. (2015). Assessing the implementability of telehealth interventions for self-management support: A realist review. *Implementation Science*, 10(1), 1–25.
- Victor, C., Scambler, S., Bond, J., & Bowling, A. (2000). Being alone in later life: Loneliness, social isolation and living alone. *Reviews in Clinical Gerontology*, 10(4), 407–417.
- Wang, D., Liang, Y., Xu, D., Feng, X., & Guan, R. (2018). A content-based recommender system for computer science publications. *Knowledge-Based Systems*, 157, 1–9.
- Wang, S., Cao, L., Wang, Y., Sheng, Q. Z., Orgun, M. A., & Lian, D. (2021). A survey on session-based recommender systems. *ACM Computing Surveys (CSUR)*, 54(7), 1–38.
- Wei, J., He, J., Chen, K., Zhou, Y., & Tang, Z. (2017). Collaborative filtering and deep learning based recommendation system for cold start items. *Expert Systems with Applications*, 69, 29–39.
- Westlund, A. C. (2009). Rethinking relational autonomy. *Hypatia*, 24(4), 26–49.
- Whittlestone, J., Nyrup, R., Alexandrova, A., & Cave, S. (2019). The role and limits of principles in AI ethics: Towards a focus on tensions. In *Proceedings of the 2019 AAAI/ACM conference on AI, ethics, and society* (pp. 195–200).
- WHO. Active ageing: A policy framework. Retrieved 2023, from <https://apps.who.int/iris/handle/10665/67215>
- Xin, Y., & Jaakkola, T. (2014). Controlling privacy in recommender systems. *Advances in Neural Information Processing Systems*, 27.

- Yıldırım, S., Söyler, Ş. G., & Akarsu, Ö. (2020). Building a non-personalized recommender system by learning product and basket representation. In *2020 IEEE international conference on big data (Big Data)* (pp. 4450–4455). IEEE.
- Zhang, S., Yao, L., & Xu, X. (2017). Autosvd++ an efficient hybrid collaborative filtering model via contractive auto-encoders. In *Proceedings of the 40th international ACM SIGIR conference on research and development in information retrieval* (pp. 957–960).

**Publisher's Note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.