Ethics of the marketing communication of AI-based services

Name:	Anne Wind	Supervisor:	Dr. Sjoerd de Vries
Student number:	s1985310	2nd. supervisor:	Dr. Efthymios Constantinides
E-mail:	t.j.a.wind@student.utwente.nl	Date:	23 - 02 - 2020
Master specialization:	Marketing Communication	Total number of words:	<14350

Abstract

As the current technological proliferations continues to incorporate Artificial Intelligence (AI) in business to consumer (B2C) products and services, the safety and autonomy of consumers is at stake. While guidelines for the development of AI has already been set in motion, the marketing communication of these services is currently overlooked. Yet marketing communication is fundamental in a society's understanding of reality. Hence, a society's understanding of AI. Consumers need to be aware of the possible dangers these products are holding. Accordingly, this work proposes an initial set of 6 appropriate guidelines for the marketing communication of AI-based services.

First, a set of preliminary guidelines were drawn from an ethical framework which was described using principlism. Second, the development and the preliminary guidelines were evaluated among a panel of experts. Feedback was used to finalize the preliminary guidelines in a final set. Third, the final set was evaluated for its appropriateness among marketing communication professionals through an online questionnaire. The guidelines proved to be appropriate in describing the elements critical for ethical marketing communication of AI-based services. However, the guidelines are not yet applicable. This work opens the dialogue about ethical marketing communication of AI-based services and facilitates future research in to developing applicable guidelines.

Keywords: Principlism, TARES, eCTA, Artificial Intelligence, e-Delphi, communication ethics, technology ethics, ethics of Artificial Intelligence.

Table of Contents

1. Introduction

The effects of advertising of unhealthy products such as fast foods or cigarettes are well known. Diseases such as obesity or cancer resulting indirectly from these advertisements, motivated governments around the world to regulate the marketing communication of these products. Unfortunately, regulations came much later than when these products first appeared on the market. Similarly, we see a lot of new technological products of which we question if these are actually good for our (mental)health. Should we trust autonomous vehicles to take over the wheel? Can we trust big tech companies with recordings of our conversations? Or should we get involved with transhumanistic ideas such as enhancing our brains with a computer chip? The most prominent technological development is Artificial Intelligence (AI)[1], significant in most of other emerging disruptive technologies (Urban, 2015a) and already disrupting the job market (International Telecommunication Union, 2017). Accordingly, companies and organizations are proposing guidelines for the development of an ethically aligned AI (e.g., Future of Life Institute, 2018; Institute of Electrical and Electronics Engineers, 2019; OpenAI, n.d.; Partnership on AI, n.d.). While it is generally agreed the effects of AI on our society will be drastic, (Urban, 2015a; Urban, 2015b; ITU,2017), the marketing communication of AI is currently overlooked. Hence, the IEEE (2019) emphasize companies should "create roles for senior-level marketers, engineers, and lawyers who can collectively and pragmatically implement ethically aligned design" (p. 131). Especially, considering the current a gap between how AI-based services are marketed and their actual performance (IEEE, 2019).

Considering the high paced development of AI and its prodigious potential (Urban, 2015a), the safety and autonomy of consumers is at stake. Accordingly, the marketing communication of AI-based services should be aligned before these services become common good in the market. To wit, as technology is advancing it seems consumers lose understanding of the exact workings of their products. While the novelty of new technologies already has an effect on perceived risk and technology adoption (Foster & Rosenzweig, 2010), ambiguity can have an effect on technology adoption as well (Barham, Chavas, Fitz, Salas & Schechter, 2014). Potentially problematic for poorer countries as technology diffusion helps these countries to catch up with richer countries (Nelson & Phelps, 1966). Illustrative to this, AI is expected to further increase income inequalities within and between countries (ITU, 2017). Currently, the technology is considered ambiguous and consumers have trust and risk issues regarding the technology (Pega, 2018), making it not amenable to everyone (European Union, 2018a; Harari, 2018). Moreover, marketing communication can be very influential in the acceptance and adoption of technology (Kardes, Cronley & Cline, 2014) and constituting our reality (Harris & Sanborn, 2014). Accordingly, this work researches the question:

RQ 1. "What are appropriate guidelines for ethical marketing communication of AI-based services?"

Prior to developing guidelines one ought to describe the critical dimensions and subsequent principles forming the basis of the guidelines. As such, a method should be sought in order to develop an ethical framework of which the guidelines can be drawn:

RQ 1.1 "What is an appropriate method for developing an ethical framework from which marketing communication guidelines for AI-based services can be drawn?"

In order to answer RQ 1.1, literature regarding methods for developing ethical frameworks are scrutinized (§2.1). Secondly, the method found will be used to develop an ethical framework (§2.2). Third, preliminary guidelines are drawn from this framework in §2.3. Fourth, the used method and the guidelines are evaluated among a panel of experts in an e-Delphi. The experts' evaluation of the methods' appropriateness is described in §4.1. Adjustments to the preliminary guidelines proposed by the expert panel can be applied as well and

1 See §2.2 For a more in-depth discussion of what AI is and why it is ethically sensitive

the guidelines are then finalized (see final document in Appendix B). Finally, RQ 1 is answered by scrutinizing the appropriateness of the final guidelines among marketing communication professionals through an online questionnaire (see §3.2 & §4.2). As such, an expert study is conducted to evaluate the appropriateness of the method and to finalize the guidelines. Subsequently, a professional study is conducted to evaluate the appropriateness of the guidelines.
Since there are various perspectives of what is right and wrong behaviour within ethics (Fieser, n.d.), it is important to describe the ethical scope of this work. As such, the ethical scope and some key concepts are described first in §2.

This research offers three main contributions. First, researching appropriate guidelines will stipulate the elements critical in marketing communication of AI-based services. Enabling the ethical debate about the urgency and importance of ethically aligned marketing communication of AI-based services. Secondly, guidelines will help enable marketing communication professionals to implement ethically aligned marketing communication. Third, methodological and conceptual research directions for developing marketing communication for the guidelines are provided. Potentially translatable for other disruptive technologies as well.

2. Description of guideline formulations

In this section the methodology of the literature review is described. Second, key concepts foundational to this work are clarified. Third, the ethical scope is described. Fourth, methods for developing the ethical frameworks are scrutinized (§2.1), an ethical framework for developing the guidelines is described (§2.2) and the preliminary guidelines are presented (§2.3). See figure 1 for an overview of the literature review and formulation of the guidelines.

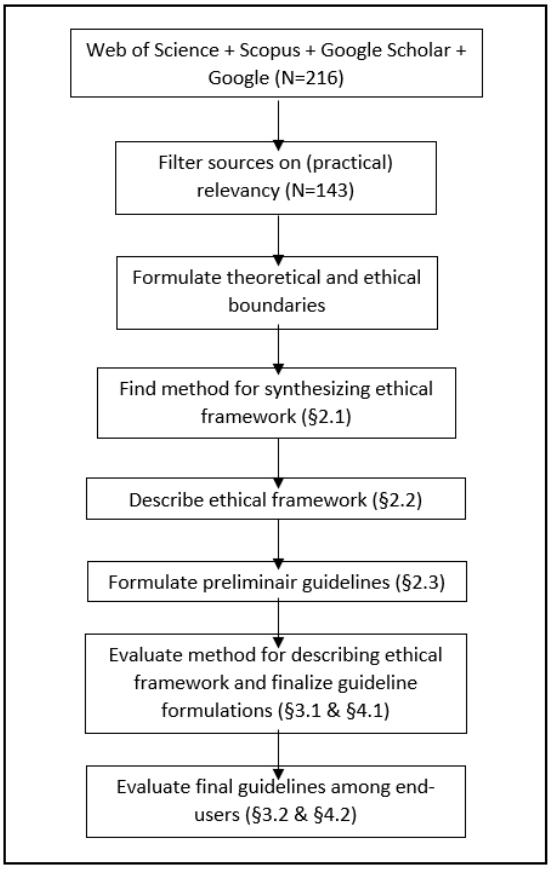
Methodology of the literature review

Communication, technology philosophy and ethical theories are conventionally researched. Accordingly, English peer-reviewed journals, books and conference proceedings were scoped. However, AI ethics in this context of marketing communication is a fairly novel topic. As such, a broader scope for this topic was used, utilizing governmental and organizational reports as well as journalistic articles. Some concern for the publication dates of literature on technical definitions on AI was required, as some definitions can become obsolete quickly (Urban, 2015b).

Tranfield, Denver and Smart (2003) propose to use pre-defined search queries to ensure re-feasible, systematic literature research. However, considering the current state of research within this topic, a wider variety of search results was preferred. Articles were selected whether they described one of the following topics: AI, technology (ethics), communication (ethics) or the development of ethical guidelines. The selection process resulted in 216 sources, which were re-evaluated considering the development of guidelines or describing ethics of the discussed topics. Finally, 143 sources were used, of which six reports and one conference video presentation.

Figure 1

Flowchart literature review and formulation of the guidelines



Key concepts

AI-based services, marketing communication and end-users can be broadly interpreted. Accordingly, these concepts are described in more detail.

AI-based services

AI-based services are defined as products and services designed to be used by consumers. Consider for instance, autonomous vehicles, voice assistants like Google Home or Brain Machine Interfaces (BMI) such as Neuralink's brain interface chips.

The scope of this research' is limited to B2C AI-based services, as media is especially influential in shaping a world that becomes a consumers' reality (Harris & Sanborn, 2014). In that respect, the disruptive aptitudes of (the use of) AI regarding consumers are twofold: interpersonal and intrapersonal. While B2B communication, as part of media, can definitely be influential in shaping epistemic perspectives, it is outside the scope of this study.

Marketing communication

Within this research marketing communication is conceptualized in threefold: broad, non-evaluative (Dance, 1970) and receiver oriented (Watzlawick, Bavelas & Jackson, 1967). As such, anything someone says or does, even if it is misunderstood, is considered communication; as long as one mind is affected by another.

End-users

The guidelines are meant to be used by marketing communication professionals assigned with the task of developing communicative campaigns to market AI-based services and persuade consumers to buy these services. The guidelines ought to be used at the start of the development of AI-based services. Hence, research has shown that the development of communicative campaigns is most effective when conducted at the start of (new) product development (NPD) (Fain, Kline & Duhovnik, 2011; Paiva, Gavronksi & d'Avila, 2011; Swink & Song, 2007).

Ethical scope

This section describes the ethical perspective from which this study is conducted. Since emphasis is placed on Harris and Sanborn (2014) their perspective of mass media influencing our perceived realities, the point of view is relativistic. Where mass media is, in part, responsible for shaping our moral epistemology on AI. After all, ethics of (communications on) AI is not something that exists independent of humans, but is created by ourselves.

For instance, one can assume companies will have an egoistic perspective in developing and distributing these AI-based services. Whilst probably overlooking the potential hazardous consequences of distributing these services for consumers and society. Concurrently, an altruistic perspective of communication professionals can be assumed. Since they should better understand the consequences of their communications and feel responsible for distributing these services ethically acquiescent.

Accordingly, this work is concerned with the role of reasoning our moral actions. As the guidelines prescribe specific moral behaviour of marketeers which needs to come from somewhere. While some philosophers argue moral assessments are fundamentally emotional assessments. This research' conviction is in line with Kantian philosophy. Which argues our moral choices can at least be substantiated by some form of reason or justification (Fieser, n.d.). However, the point is to develop guidelines for a large group of people. Accordingly, the justification should apply to the largest group, employing utilitarian perspectives as well (Fieser, n.d.).

Justice

Guidelines inherently imply a prescription of what is just in a given situation. So, what is 'just' within the scope of this research?

The American company Amazon utilizes AI technology to personalize web-experiences for their users all over the world (Wiggers, 2019). Companies like Amazon make increasingly use of AI robots to manage their warehouses. Which not only fosters job loss in the home country, but overseas as well (Lewis, 2014). Neuralink might succeed in the future to create symbiosis between our brains and AI (Lopatto, 2019), enabling humans who have access to this technology ascend human capabilities. As such, AI is able to mediate our lives on intrapersonal and interpersonal levels, overarching even national borders. An appropriate conceptualization of justice should take in to account the transcendence AI has over multiple levels. Fraser (2008) describes such a conceptualization of justice by describing the 'what', 'who' and the 'how' of justice.

What

First, the 'what' of justice is conceptualized in the normative principle of parity of participation. Where institutionalized obstacles that are preventing people from participating on par with others, need to be dismantled. This dismantlement is considered within three dimensions. (a) People can be restricted by economic structures which deny them full and/or equal participation; invoking the need for redistribution. (b) They can also be restricted by institutionalized cultural structures, denying them necessary standing resulting in misrecognition; invoking the need for recognition. (c) Finally, people can be restricted from decision-making structures that deny them democratic participation in public deliberations, resulting in misrepresentation; invoking the need for representation.

Who

Second, the 'who' of justice or the frame to whom justice is applied, is conceptualized in the all-subjected principle. This principle encompasses everyone who is subject to the same governing structure. Turning a collection of people into fellow subjects of justice not on the basis of nationality, abstract personhood or causal interdependence, but their mutual subjection to a structure mediating their lives.

How

Third, the 'how' of justice should encompass both dialogical and institutional features. Justice should not be determined authoritatively by powerful states or by technocrats employing scientistic presumptions. These approaches are blind for the impeding claims of the disadvantaged. Instead, the framing of justice should be disputed dialogically, seeking resolution in unrestricted inclusive public discussion. Additionally, the dialogue should be supported by fair procedures and representative structures in order to pursue democratically legitimate deliberation. Of which the representatives need to be capable of recognizing the 'who' of justice as discussed above.

2.1 Methods for developing ethical frameworks

Communication guidelines are usually informed from ethical theory, often developed by ethical experts (e.g., Baker and Martinson, 2001; Tilley, 2005). However, it is preferable such methods are also usable by non-academic professionals. Enabling professionals to develop ethical guidelines more efficiently. Explicit methods for describing an ethical framework regarding a specific topic are especially prominent in the field of biomedical ethics. Beever and Brightman (2016) used methods from biomedical ethics to develop their own ethical reasoning within engineering ethics. They informed their work from Pinkus, Schuman, Hummon and Wolfe (1997), who argued biomedical ethics is an interdisciplinary endeavour and places emphasis on the need of theoretical incorporation, together with contextual information and principles. Accordingly, the field can be well translated to other domains than the medical only. There are a variety of bioethical theories (Khushf, 2004), four groups of theories are discussed below.

Narrative approaches

First, we can define narrative approaches to bioethics. Here it is believed that morality can only be drawn from a culture's story. Where, through narrative, one is able to connect contingencies and also describe the more complex interpersonal relationships (Burrel & Hauerwas, 1997; Nelson, 2004). For instance, feminist approaches emphasize one's empathic perspective for the good of others and their community for an adequate response to the need of others. In their narrations they challenge the inadequate traits society is labelling (Tong, 2004). Casuistry proposes a more structural way of narration and tries to describe the specific (often controversial) case at hand and classifies the problem by making analogies, stories or comparable cases. This process usually yields normative considerations which can then be ranked (Boyle, 2004). Another narrative-like approach has a phenomenological perspective. In bioethics it is oriented around the concrete experiences of the doctor patient relationship. The methodology makes use of heuristics and from there narrate how the life-world of doctor and patient contribute to these ethics (Pellegrino, 2004).

These approaches are especially useful for describing and solving a specific problem. Narrating ethicists emphasize that their approach is adequate since it enables one to go down every aspect of a problem. According to them, structural approaches of principle-based theories, for instance, are too rigid to deliver an adequate response (i.e., Stocker, 1987; Walker, 1998; Williams & Bernard, 1981). However, within this work a structural framework is preferred since it will allow non-academics or non-ethical experts to develop ethical principles in similar ways. Useful since the overall structure (marketing AI-products) will be the same. Moreover, casuistry needs similar cases to describe a problem, which are not available considering the novelty of marketing AI-based services. In addition, casuistry might be vulnerable to misuse in the persuasive world of marketing communication as casuistry is often used to persuade in legal contexts (Boyle, 2004).

Common morality approaches

Secondly, the common morality approach can be defined. It is an informal public system which applies to all rational individuals. It governs the behaviour that affects others commonly known as moral rules, ideals and virtues that reduces harm in the world. While this approach uses a framework in its application. It needs similar cases to describe a problem, make analogies and find a solution (Clouser & Gert, 2004). Again, such an approach is inadequate considering the novelty of marketing AI.

Virtue theory

Third, virtue theory underlines the role of character and virtue. A virtue ethicist questions how a virtuous individual should act in a given situation (Athanassoulis, n.d.). In comparison with other ethical theories, it is viewed as a theory of balance. On one hand, it refrains from relying too much on abstractions and tries

to emphasize concrete interactions in an ethical Situation. On the other hand, it criticizes over-reliance on the concrete and avoids relativistic point of views. Accordingly, it is based on common social and personal structures of human existence (Thomasma, 2004). However, some believe the theories of virtue lack critical reflection and sound moral conviction (i.e., Boyle, 2004). Thomasma (2004) describes that is precisely why virtue theory can offer a solution, a middle way is needed since ethical theory is too abstract to contribute to much discussion. Still though, the practice has been associated with contributing to the blurring of norms and standards in medicine (Veatch, 185, pp. 338-340). Which might not be ideal for developing novel guidelines for marketing communication of AI.

Principlism

Finally, principlism offers a structural consequentialist decision-making approach (Bulger, 2007). It has some aspects of the common morality approach, emphasizing that all persons serious about morality (regardless of origin) will judge human conduct by a shared set of norms (Gordon, Rauprich & Vollman, 2011). The set of norms entails a framework of principles grouped under four categories: (1) The principle of autonomy (supporting and respect for autonomy), (2) The principle of beneficence (work towards beneficence), (3) the principle of nonmaleficence (averting harm) and (4) the principle of justice (democratically distribute benefits, risks and costs) (Beauchamp & Childress, 2009; Beauchamp & DeGrazia, 2004). The principle of autonomy has five additional conditions: (1) Competence: the complexity or difficulty of the task or judgement; (2) Informed consent: The agents' decision must be autonomous and institutionally authorized; (3) Intention: The agent must intentionally make a choice; (4) Understanding: The agent must choose with substantial understanding; (5) Freedom: The agent must choose without substantial controlling influences (Bulger, 2007).

Principlism can be structurally applied in order to evaluate whether your ethical theory or principles are adequate. This can be done in three phases, specifying, balancing and justifying. First, ethical theories should be specified, by describing them and clarify what they are (Beauchamp, 2003). Next, the principles need to be balanced, since principles often are mutually conflicting for the specific situations they are used in. Accordingly, balancing uncovers what principle has more weight (Beauchamp & Childress, 2013). This can be done by comparing the specified ethical theories or principles with the four principles of principlism. Finally, the ethical principles or theory need to be justified by describing to what extend it considers each of the four principles of principlism (Beever and Brightman, 2016). As such, the principles or ethical theory of choice are not adjusted but rather evaluated, weighted and justified to clarify to what extent and how the principles could and should be used.

Method for developing ethical framework in this work

Principlism is chosen as a method for describing the ethical framework in this work. The reflectivity core to this approach enables users to evaluate statements through inductivist and deductivist methods and make adjustments among abstract theories to reach to the most common viewpoint. Enabling users to describe and reflect on critical facets of a problem. Beever and Brightman (2016) believe this iterative process is profoundly supportive for developing nuanced responses when considering novel, emerging technologies. While enabling one to think conscientiously and deliberately about what one is doing.

2.2 Description of ethical framework

In order to apply principlism in this work and to describe an ethical framework, the relevant dimensions underlying of guidelines for the marketing communication of AI-based services need to be described. Three dimensions can be identified: first, one ought to know what is ethical in marketing communication. Second, the ethicality of technology in general needs to be described. Finally, the ethicality of communicating AI-based services needs to be scrutinized. Together, these three dimensions make up an ethical framework from which the guidelines can be drawn.

To describe the dimensions, principlism is used to reflect on each of the dimensions individually. First, each dimension is specified by describing relevant principles or ethical theories and choosing the most relevant set of principles or theories for the dimension. Second, the chosen set of principles or ethical theories are balanced to uncover which have more weight. Finally, the principles are justified for their accordance with principlism.

Specification of principles for marketing communication

In this research specific emphasis is placed on humans' responsibility in their communicating abilities. A responsible communicator should reflectively analyse their claims, deliberate likely consequences of their communication and conscientiously consider relevant principles in their communication (Johannesen, Valde & Whedbee, 2008). In the scope of this research two distinct ethical implications of communication are reflected upon. First, within marketing communication communicators should double-check the soundness of their message before communicating it to others (Johannesen et al., 2008). As it could be morally culpable when one tries to deliberately use dubious reasoning in persuasive communication (Rescher, 1977). Even when the intention is not to deceive others, communication could be morally questionable. For instance, the use of jargon-laden language could cloud accurate, clear representation of ideas (Johannesen et al., 2008). As such, marketing communication can be viewed as inherently ethical implicative. Second, publicized communicative messages – as a part of mass media – are not only reflecting our worldviews, they are also constructing a world that becomes our reality (Harris & Sanborn, 2014, pp. 69-70). To wit, Agenda Setting Theory (McCombs & Reynolds, 2009) tells us what is important to think about, Social Cognitive Theory (Bandura, 2009) tells us how we should behave in our reality, Cultivation Theory (Morgan, et al., 2009) tells us how a worldview could be constructed, and the Schema/script theory and the limited capacity model (Lang, 2000) informs us how knowledge structures are created from exposure to media. In this manner, humans are mutually constituting their reality (Harris & Sanborn, 2014, pp. 69-70). In sum, marketeers, professional communicators or more general, 'media-makers' should be more cognizant in how they are complicit in creating our (perceived) reality, through (un)intentional dubious reasoning.

Accordingly, principles for marketing communication should be meaningful and flexible for our communication behaviour and for the evaluation of communication of others, encompassing both individual and social ethics (Johannesen et al., 2008). Within communication there are frameworks, ethical codes, models and principles available for ethical evaluation and decision-making regarding marketing communication and public relations (PR), some of these will be discussed next.

Frameworks as introduced by Johannesen et al. (2008, pp. 15 - 16) or Kidder (1995) are statements on the ethical foundations of a particular communication. Which can be used systematically to make informed judgements of communication ethics. In addition, there are ethical frameworks for journalism and mass media such as McElreath (1997) his Potter Box Model of Reasoning. An effective strategy for how media outlets can decide on an ethical framework in a specific case. A three-step strategy where a case leads to values, leading to certain loyalties, which determines the ethical action a media outlet needs to take. Similarly, Bivins (2003) developed a "checklist for ethical decision making" (p. 174-179). However, these

types of frameworks oversee to take in to account the specific design of messages and, thus, overlooking constitutive aptitudes of (un)intentional communication. Additionally, the frameworks of Johannesen et al. (2008), Kidder (1995) and Bivins (2003) employ a checklist approach. This evaluation of ethics in retrospect is undesirable when considering the assessment of technology (see specification of technology).
Next to this, there are also ethical codes in the realm of advertising. Such as the American Association of Advertising Agencies (AAAA) code of ethics (1990), the International Association of Business Communicators (IABC) (n.d.), the Public Relations Society of America (PRSA) Code (n.d.) or the International Chamber of Commerce Communications Code (ICC) (2018). In contrast with the frameworks, these codes do stipulate some of the design choices of marketing communicative messages. Additionally, with these codes an organization can exhibit their ethical integrity by its membership to one these associations. However, a membership is not a testimony to ethical acquiescent marketing communication. While some of these codes are clear short statements (e.g., AAAA and IABC), they can still be misunderstood due to abstract, vague or static language. The codes could also foster a detrimental passive attitude among users regarding ethical considerations (Johannesen et al., 2008). Additionally, while the ICC code is more comprehensive than its peers it is a long document which is not practical. Moreover, none of the codes are specifically aimed at persuasive communication, but entail more generic statements mostly to cover both marketing as PR.
Some have researched models or theories of public relations ethics specifically. Bowen (2004) developed a model which allows practitioners to systematically analyse ethical aspects and make an informed decision from multiple perspectives. However, this model too oversees specific design choices able to constitute certain knowledge structures. Other researchers such as Marsh (2001) emphasize what an adequate line of thought should be when considering public relations. Which is helpful for other researcher to comprise models of their own, but not very practical for professionals. In contrast, Tilley (2005) theorized an ethical tool professionals could use to find ethical approaches that work for them in their specific cases. Focused on enabling a proactive attitude of the practitioners and on aligning their ethical approaches to their campaign design, implementation and evaluation. However, while the ethics pyramid of Tilley (2005) does seem to provide an adequate framework for any marketeer in any context, it is rather an “organizing strategy” (p. 313). Which might be fruitful for future research in ethics of AI communication.
Finally, Baker and Martinson (2001) developed five principles which comprise the TARES test. As a consequentialist test, it takes in to account the constitutive aptitudes of communication, while also specifically guiding persuasive practices to morally accepted ends. The five principles “are prima facie duties that generally hold true, all other things being equal” (Johannesen et al., 2008, p. 14). To be said, the principles are evaluative of ethics rather than aligning and can be viewed as a checklist approach. Baker and Martinson (2001) literally provide checklists for moral reflection in relation to the principles (i.e., pp. 161, 164, 165, 167 & 170). However, these principles are meant to determine “the boundaries of persuasive communications” (p. 172). Also taking in to account constitutive variables such as content and execution of appeal.

Principles for marketing communication

Accordingly, the TARES principles lend themselves nicely for formulating the marketing communication aspect of this research its guidelines. The principles are:

- Truthfulness – of the message (honesty, trustworthiness, non-deceptiveness)
- Authenticity – of the persuader (genuineness, integrity, ethical character, appropriate loyalty)
- Respect – for the persuade (regard for dignity, rights, well-being)
- Equity – of the content and execution of the appeal (fairness, justice, nonexploitation of vulnerability).
- Social responsibility – for the common good (concern for the broad public interest and welfare more than simply selfish self-interest).

Balancing TARES principles of marketing communication

The TARES principles explicitly focus on beneficent persuasion. As the message needs to be truthful, the persuader needs to be honest, sincere, loyal and independent, have explicit respect for the persuadee and genuinely believe the product will benefit the persuadees (Baker & Martinson, 2001). However, under justice, non-maleficence and autonomy and its conditions, the following TARES principles do conflict: Truthfulness (of the message); Respect (for the persuadee); Equity (of the Persuasive Appeal) and Social Responsibility (for the Common Good). Baker and Martinson (2001) prescribe persuaders to "disseminate truthful messages through equitable appeals" (p. 163) to "all other who will be affected by the persuasion" (p. 163). Moreover, there should be "parity between the persuader and persuadee in terms of information, understanding (…) and to the level of playing field (the lack of parity must be fairly accounted for and not unfairly exploited)" (pp. 165-166). Additionally, persuaders should "be sensitive to and concerned about the wider public" (p. 167). The complexity of AI-based services might restrict persuaders to utilize 'equitable appeals' for 'all affected by the persuasion'. For instance, a banner advertisement on a metro station for Neuralink its AI brain chip might be understandable for a middle-aged high educated individual. The person is competent enough to intentionally and voluntarily make the choice to contact Neuralink and register himself for their product. However, an older lower educated person on that same metro station might not understand the banner advertisement because he/she never have heard of AI, let alone, brain enhancement chips. How should persuaders make sure there is 'parity between the persuader and persuadee in terms of information,' account fairly for the lack of parity without exploitation, fairly distribute benefits and risks and make sure the persuasion is non-maleficent in these situations?

Looking back to this research its conceptualization of justice, we are talking about 'misrecognition' through the danger of people being restricted by institutionalized cultural structures. On the one hand persuaders want to be truthful about their message to all who are subjected to the advertisement. But the autonomy principle shows how some people subjected to the advertisement might not be competent enough to understand the message. Concurrently, the justice principle dictates fair distribution of benefits and risks and thus, the persuasion can be at odds with the principle of non-maleficence. However, the persuaders should direct the information of their ads to the largest competent group. As described earlier, this research is written from a utilitarian, consequentialist perspective. Additionally, considering the 'how' of justice, an attempt to account for the non-competent recipients should be made. Since an altruistic demeanour of the persuaders is assumed, the persuaders should incite a dialogue between competent and non-competent recipients of the advertisements. Which requires the persuaders to know exactly who are subjected to their advertisements. As such, they could help raise awareness and educate non-competent consumers through dialogue.

Justifying TARES principles of marketing communication

The consequentialist TARES test is coherent with principlism and thus, can be used to aggregate guidelines for marketing communication of AI-based services (§2.3). The following sections describe TARES for its accordance with the four principles of principlism.

Accordance with principle of autonomy

TARES take in to account of competence by the principle of autonomy since it demands persuaders to examine the loyalty of their practice (Baker & Martinson, 2001, pp. 162). Second, the principle of respect demands persuaders i.a., to regard other humans worthy of dignity and not act for pure self-interest. Third, equity requires that "persuasive claims should not be made beyond the persuadees' ability to understand both the context and underlying motivations and claims of the persuader," (Baker & Martinson, 2001, pp. 166). Similarly, equity but also truthfulness fairly accounts for informed consent, intention and understanding

since equity demands that the persuadee should fully understand the persuaders claim in order to make a good decision. Finally, the principle of truthfulness takes into account the principle of freedom in the way that people should be free from controlling influences (Bulger, 2007, pp. 91). To wit: "The Principle of Truthfulness requires the persuader's intention not to deceive, the intention to provide others with the truthful information they legitimately need to make good decisions about their lives" (Baker & Martinson, 2007, pp. 160). As such, TARES complies with the principle of autonomy.

Accordance with principles of Beneficence and Nonmaleficence

The TARES principle of respect demands regard for dignity; rights; well-being. Secondly, equity demands fairness; justice; nonexploitation of vulnerability and finally social responsibility requires concern for the broad public. As such, beneficence "the principle of contributing to the welfare of others" (Bulger, 2007, pp. 92) and nonmaleficence "the principle of not harming others" (Bulger, 2007, pp. 92) are fairly accounted for with TARES.

Accordance with principle of Justice

In TARES, the principle of equity fairly accounts for the principle of justice. As Baker and Martinson (2001) state that:

The Equity Principle requires either that there be parity between the persuader and persuadee in terms of information, understanding, insight, capacity, and experience, or that accommodations be made to adjust equitably for the disparities and to level the playing field (the lack of parity must be accounted for and not unfairly exploited), (Baker & Martinson, 2001, pp. 165-166).

Specification of ethical principles of technology

There are two branches of dogmas in order to acquire and establish orientational knowledge of (new) technology. First, there is an ethics of technology approach, a kind of philosophical ethics emphasizing the normative implications of decisions on technology. Second, there is technology assessment (TA) which relies on sociological or economic research (Grunwald, 1999). Of which TA is more concerned with managing technology in a society, applicable to the scope of this work.

Palm and Hansson (2006) argued TA needs to be expanded to include the ethical implications of technology adequately and constructed ethical technology assessment (eTA). They proposed to undertake the evaluation of technology in the form of a continuous dialogue with technology developers. However, Kiran, Oudshoorn & Verbeek (2015) critiqued eTA for its checklist approach. Instead, the researchers proposed a set of principles for an ethical-constructive technology assessment (eCTA). Kiran et al. (2015) emphasize that their eCTA approach accounts for external processes of technology development. It is a framework which can be used as a tool for identifying unwanted effects of new technologies early on in their development process while also accounting for changing variables over time. The fluidity of the eCTA framework suits the polymorphic applicability and high pace but opaque development of AI. Moreover, Kiran et al. (2015) approached the eCTA framework to see "how a technology could get a desirable role in society" (Kiran et al., 2015, para. 3.2). As such, eCTA goes beyond a 'checklist-approach' of ethics and leaves room for change of perspectives. In addition, the framework provides four principles to which it is harnessed.

First, technologies have implications at the 'embodiment relation', where humans are given a sensory relation 'through' objects. The researchers state that this embodiment principle requires systematic thinking about the global (macro and micro) impact technologies have (Kiran et al., 2015). To think systematically about ethics, one could use a framework as introduced by Verbeek in his 2013 study. Where three elements of mediation theory help to anticipate mediations more completely (see figure 2). The first element is the locus of the technology, which could be physical, cognitive or contextual. In the second element the form of the technology should be considered. As such, technologies could be coercive, persuasive, seductive or decisive. The last element considers the domain of impact of the technology. Explaining what the technology means for individual and social experiences and their consequential actions.

Figure 2

Verbeek's (2013) framework for anticipating technology

Locus	**Form**	**Domain**
Physical	Coercive	Individual: experience
Cognitive	Persuasive	Individual: actions
Contextual	Seductive	Social: frameworks of interpretation
	Decisive	Social: social practices

Second, there is a hermeneutic relation with technology, where humans have to read a technology. For instance, the safety catch on a gun indicates how one should handle firearms.
Third, there is an alterity relation with technology, wherein humans interact with a technology. Accordingly, the design of the technology should be done in such a way that they are open to situatedness, cultural pluriformity and fluctuating moral views. For instance, consider a public bathroom designed for all types of people or only for non-handicapped men. Nudge theory by Thaler and Sunstein (2008) as proposed by Verbeek (2013) has a way of designing ethics in technology. It entails designing in a way that elicits i.e., positive behaviour without taking away control.
Fourth, there is a background relation with technology wherein technologies have a contextual role. As such, eCTA considers humans' moral responsibility in actively shaping their lives in accompaniment of technologies. Specifically, eCTA should make visible how this responsibility is enacted in daily life considering use, non-use and selective use of technologies.

Principles for technology

Accordingly, eCTA provides ethical principles in technology:

- The embodiment principle
- The hermeneutic principle
- The alterity principle
- The background relation principle

Balancing eCTA principles of technology

Like TARES, eCTA conflicts with autonomy, justice and in turn non-maleficence. As design principles, they advocate a systematic assessment of how the design of the technology is impacting individual and social experiences. Considering AI's complexity, design choices might be maleficent to people who are not competent enough to understand these design choices. These people could therefore mis out on the technology or don't see how the technology could hurt or benefit them. Yet again there seems to be a danger of misrecognition. Like with TARES, persuaders should first focus on the largest group of competent consumers. Secondly, non-competent consumers should be taken in to account by invoking a dialogue between competent and non-competent consumers.

Justifying eCTA principles of technology

Below the consequentialist eCTA framework is substantiated for its accordance with principlism.

Accordance with principle of autonomy

When the embodiment principle is applied systematically by e.g. Verbeek's (2013) framework, it accounts fairly for two components of autonomy: competence and understanding. By emphasizing the importance of the locus, form and domain of a technology, the full scope of skill needed to be able access the technology is described. Second, the background principle describes the moral significance of use or non-use of a technology which implies cognizance of intention. Third, the hermeneutic principle informs how the user experience its interaction with the technology and how that mediates users' decisions. As such, informed consent can be accounted for (Kiran et al., 2015). The final underlying component of autonomy: freedom, is accounted for by both the embodiment principle by taking into account 'controlling influences' and the hermeneutic principle since it implies being cognizant of choice.

Accordance with principles of Beneficence and Nonmaleficence

The alterity principle of eCTA typically emphasizes the importance of a technology being beneficent for all sorts of people. Nonmaleficence is not explicitly mentioned in eCTA, but humans' responsibility in designing technology and technology's coerciveness as potential threats for users are implied in the background, hermeneutic and embodiment principle.

Accordance with principle of Justice

The alterity principle and more pertinent, the background principle of eCTA takes into account democratically distributing benefits and risks. Or, at least, implies conscientiousness of how technology has contextual implications when i.e., a technology is used versus when it is not-used.

Specification of ethical principles in communicating AI-based services

While the previous dimensions are individually well researched and have established principles, ethics of communicating AI-based services are a novel topic. As such, this section studies AI "at the micro-level, where technologies help to shape engagement, interaction, power, and social awareness" (Verbeek, 2017, p. 301). Secondly for matters of triangulation, a few major institutional organizations' their guidelines and principles are described to derive AI's ethical implications. First, social-economic implications of AI's workings are described. Second, the ethical implications of AI's technical workings are scrutinized. Finally, institutional views are scrutinized and aggregated to describe principles of communicating AI-based services. As such, this section tries to 'lift the veil' on AI-based services' ethical implications within the scope of marketing communication.

While AI is a catch-all term for intelligence demonstrated by machines, the current controversy mostly revolves around machine learning algorithms. Which enables software to learn from data in order to make predictions or decisions without being explicitly programmed to perform the task (Koza, Bennet, Andre & Keane, 1996). As such, it is able to discover patterns in vast datasets and from there generate insights. The vaster the dataset, the more complex AI is needed and more computing power is required, the more computing power used the more powerful AI becomes (Jesus, 2017). Beneficially for AI development, computing power is doubling every two years (Moore, 1965) and recent developments in chip design are making the available compute even ten times larger every year (Seabrook, 2019). Meaning that by 2020 the computing power will match the human brain (Urban, 2015a). Currently, we make use of Artificial Narrow Intelligence (ANI), AI capable of specific tasks such as distinguishing humans from cars or recommend where you should invest in. Since computing power keeps increasing, research indicates we could achieve greater forms of AI such as Artificial General Intelligence (AGI), similar to human level intelligence (HMLI), or even Artificial Super Intelligence (ASI) inconceivably greater than HMLI. While AGI and ASI are still under dispute (Armstrong & Sotala, 2012; Baum, Goertzel & Goertzel, 2011; Brundage, 2017; Dietterich and Horvitz, 2015; Müller & Bostrom, 2016; Plebe & Perconti, 2012; Russel & Norvig, 2010), ANI is already disrupting the job market (ITU, 2017) and AI's future impact is considered to be prodigious (Urban, 2015b)[2]. To wit, governments are anticipating the potential impact of AI on societal economic, political and ethical levels (China's State Council, 2017; EU, 2018a; House of Commons Science and Technology Committee, 2016; ITU, 2017; Office of Science and Technology Policy, National Science and Technology Council Committee on Technology, 2016; US Senate Subcommittee on Space, Science and Competitiveness, 2016). UN organizations come together every year to contemplate on how to constitute an AI for good (i.e., ITU, 2017). Alongside, tech companies and organizations are developing principles for ethically aligning their AI research and services (e.g., Future of Life Institute, 2018; IEEE, 2019; OpenAI, n.d.; Partnership on AI, n.d.).

Democratization

First, there are arguments to democratize AI and education on AI, these go hand in hand. To wit, AI is in itself a complex technology and concurrently applicable to various domains such as: reasoning; knowledge; planning; communication and perception (Corea, 2018). For each of these domains various AI systems can be developed each with their own capabilities. On the contrary of what ANI implicates, a lot of these AI-technologies are overlapping. Some parts of one AI technology is used in the other, but we cannot speak of AGI. As these technologies are not applicable to completely different tasks and none of these technologies are self-aware. Instead, there are 'baskets' of AI technologies developed for specific tasks. In order to solve a problem, you might need one or more ANI technologies which are not mutually exclusive per se, but rather complementary.

The current lack of an unequivocally explanation of AI might be the reason we often do not know we

2 See Urban, 2018a and 2018b for an extensive discussion AI and its potential.

already make use of AI. "As soon as it works, no one calls it AI anymore," said John McCarthy the computer scientist who invented the term AI (Vardi, 2012). In fact, what was considered AI 40 years ago are common functionalities now (Antonov, 2018). AI's ambiguity comes to be a real problem as AI becomes more complicated and more indispensable for our world. Especially when consumers ignorance on AI (Pega, 2018) sustains. To wit, the study of Pega (2018) showed that consumers were more comfortable using AI when they had better understanding of it. Moreover, knowledgeable people on AI could use the data-processing tool to their advantage and grasp an inconceivable greater advantage on others. As such, AI should not be concentrated in too few hands (Harari, 2018) and education on AI should be democratically distributed as well (ITU, 2017).

Transparency

Second, the processes on which AI makes its decisions are harder to uncover as the technology gets more sophisticated. One could ask themselves if we should let AI make decisions for us if we have a hard time knowing how the technology draws its conclusions. Bostrom and Yudkowsky (2011) emphasize how AI based on machine learning is "nearly impossible" (p. 1) to understand for why and how it draws its conclusions. While AI is also playing an increasingly large role in our society, sometimes without being labelled as AI. Therefore, it is "increasingly important to develop AI algorithms that are not just powerful and scalable, but also transparent to inspection" (p. 2). In this section AI's implications considering data, society and the technical challenges arising when considering AI ethically are described.

Transparency: data implication

Gourarie (2016) and Hardt (2014) explain that algorithms are prone to bias because of two reasons. First, as humans put in the data their biases are also encoded with. For instance, consider historical datasets where households are labelled by race or sexual preference. Secondly, algorithms look for patterns, as such minorities are disadvantaged by definition since there is always less data available about minorities. Not to mention the danger of hacking, since we become increasingly more dependent on data, hacking can become a larger problem (Harari, 2018). While blockchain technology (Schmelzer, 2018; Sun, Yan & Zhang, 2016) or quantum communication (Giles, 2019) might solve these problems in the future, these solutions are not (yet) scalable. As such, data used by or with AI should be 'clean' and susceptible to scrutinization.

Transparency: social implications

Bostrom and Yudkowsky (2011) argue that when AI is used for work with social dimensions, the AI should have social requirements. Typically, not a topic present in machine learning journals per se (Bostrom & Yudkowsky, 2011). However, one would want to know how and when an AI decides how they should live their lives. Consider, for instance, iBorderCtrl an AI which might be assessing refugees whether or not they can pass EU borders in the future (EU, 2018b). Bostrom and Yudkowsky (2011) emphasize that it is important for a legal system to be predictable, "so that, e.g., contracts can be written knowing how they will be executed" (p. 2). To wit, when e.g., iBorderCtrl fails and let malicious immigrants in the EU and rejects innocent immigrants, who is to blame? The EU? The developers of iBorderCtrl? As such, a number of studies emphasize the importance of transparency and some have been introducing the idea of a 'black box' system similar to black boxes in airplanes to be able to make AI more transparent for scrutinizing accountability (Calo, R. 2018; Desai & Kroll, 2018; Diakopoulos, 2014, 2016; Miller, 2018; Otterlo, 2018; Tene & Polonetsky, 2014; United Nations Institute for Disarmament Research, 2016). Similarly, Gunning (2017) introduced the idea of explainable AI (XAI) and Hao (2019) emphasized how Generative Adversary Networks (GAN) algorithms could enable AI with techniques to explain 'themselves'. Finally, Pasquale (2016) argues that humans still control what kind of data we fed into the algorithmic processes. Implying there is still some control, even when we do not know the specific workings of the algorithms. Although, these researchers and others (e.g., Bench-Capon,

2003;Bench-Capon & Atkinson, 2009; Broersen, Dastani, Hulstijn, Huang & Torre, 2001; Conitzer, Sinnot-Armstrong, Schaich Borg, Deng & Kramer, 2017; Hollander & Wu, 2011;Lopez-Sanchez, Rodriguez-Aguilar, Morales & Woolrdige, 2017; Noothigattu et al., 2018; Rossi, 2016) have indeed studied 'ethical algorithms', it appears to be a hard task to design algorithms which are (culturally) ethical in the eyes of all humans. Representing social responsibilities in algorithms could be a very exhaustive task (Bostrom & Yudkowsky, 2011), since the AI has to account for a wide variety of events under an even wider variety of contexts.

Transparency: technical challenges

Considering the post-phenomenologist perspective of this article on technology (see 2.4.4), some technical challenges arise when accompanying ethics in the design, implementation and use of AI. Bostrom and Yudkowsky (2011) give two examples. First, the design of a toaster is represented within the designer's mind, not intrinsically within the toaster. As such, accidentally covering the toaster with a piece of cloth will still cause an unwanted side effect. Here, the designer is not able to cover the products' safety over all contexts. Secondly, designing AI of which the designer accounts for a large array of possible outcomes is almost impossible. To wit, the AI Deep Blue (chess algorithm) needed to make its own decisions in order to beat grandmaster chess player Kasparov. First, the vast number of possible chess positions is impracticable to encode for humans. Second, if the designers would encode moves what they considered to be good moves, Deep Blue would not be any better at chess than their designers. As such, Bostrom and Yudkowsky (2011) argue that specific behaviour of AI might not be predictable, even if the designers do their absolute best. As such, assessing AI's safety becomes challenging. Instead, we must verify what the AI is trying to do, since predicting AI's behaviour in all operating contexts is unfeasible. Others, such as Maas (2018) propose to regulate AI according to other high-risk technologies as elaborated by Perrow (1984). Who suggests accounting for high-risk technologies using normal accident theory, which in short means that a technology is so complex and tightly-coupled one should expect accidents to happen.

Institutional views

In addition to scrutinizing AI at the micro-level, this section describes the views on AI's implications of various institutional organizations involved with AI. Since they develop and work with AI, these organizations might have a deeper understanding of AI's implications going beyond their innerworkings.

First, an earlier version of the Institute of Electrical and Electronics Engineers (2017) report was widely used over a variety of western governmental reports (i.e., EU, 2018a; ITU,2017; House of Commons Science and Technology Committee, 2016). To remain concrete, only the five general principles of AI for ethical design, development and implementation of AI were considered (IEEE, 2017, p. 20-32).

Second, the Future of Life Institute formed a set of 23 principles who were signed by thousands of researchers. This institute conducts research on AI, advises, consults and creates awareness about AI while also providing educational material to help understand AI (Future of Life Institute, n.d.).

Third, the Partnership on AI consist of a variety of companies such as Accenture, Apple, Amazon, Amnesty International and Deepmind (Partnership on AI, n.d.). The final company considered is OpenAI, although the company has not made listed statements, they have declared their goals. As such, they strive for discovering and enacting a safe path to AGI. To ensure their mission they call for widely and evenly distribution of A(G)I. In their journey they will publish at conferences, make software (tools) open-source and communicate their research (OpenAI, n.d.).

In general, all the statements and reports discuss the importance of transparency of data and/or the importance of collective accessibility of AI and education on AI. As such, the above elaborated is coupled under either 'transparency' or 'democratization' (see table 1).

Table 1
Aggregated ethical considerations

	Transparency	Democratization
IEEE (2017)[a]	1.Human Rights; 3.Accountability; 4.Transparency.	1.Human Rights; 2.Prioritizing Well-Being; 3.Accountability; 5.Awareness of misuse.
OpenAI (n.d.)[b]	Be transparent in research to AI	Widely and even distribution of A(G)I & Open-source of data and algorithms
Future of Life Institute (n.d.)[c]	2.Research Funding; 3.Science-Policy Link; 4.Research Culture; 6.Safety; 7.Failure Transparency; 8.Judicial Transparency; 9.Responsibility; 12.Personal Privacy; 16.Human Control; 21.Risks; 22.Recursive Self-Improvement;	1.Research Goal; 2.Research funding; 5.Race Avoidance; 10.Value Alignment; 11.Human Values; 13.Liberty and Privacy; 14.Shared Benefit; 15.Shared Prosperity; 17.Non-subversion; 18.AI Arms Race; 19.Capability Caution; 20.Importance; 21.Risks; 22.Recursive Self-Improvements; 23.Common Good.
Partnership on AI (n.d.)[d]	4;5;6a;6c;6d;8	1;2;3;4;5;6a;6b;6c;6e;7; 8

Note. [a]The five general principles as summarized in the IEEE report (2017).[b]Extracted from OpenAI's mission (OpenAI, n.d.).[c]Future of Life Institute's 23 principles (Future of Life, 2018).[d]Partnership on AI's tenets (Partnership On AI, n.d.).

Drawing from the previous, principles for communicating AI-based services should include these products' need for democratization and transparency. First, education on AI and the technology itself should be distributed democratically, in order to facilitate equal access to the technology. Second, AI-based services should be transparent since we might come across situations where we need to scrutinize accountability. Humans' accountability needs to be considered since they put in the data and make the design choices. Which determine the AI-based services' grounds for conclusions and mediations respectively. The AI-based service' accountability needs to be considered as well, as we might come across situations where we need to derive how the technology made its conclusions.

Similarly, the IEEE report makes statements on both the subject of transparency and democratization, while OpenAI seems to emphasize democratization in general and in order to scrutinize accountability. Both the IEEE report as OpenAI's statement are mainly issued on AI's development and design. More extensively, the Future of Life Institute also discusses the e.g., transparency of AI's funding, or the risk of a global AI arms race. While also concerning the alignment of all humans' values in the design of AI. Finally, the Partnership on AI's tenets are all geared towards the collective in terms of (among other things) transparency of their research, understandability and trust of AI.

Here, democratization refers to the cultural amenability and accessibility of AI as a product or service, but also to the amenability to information, education and equal rights of humans in relation to AI. Transparency refers to transparency of AI services' development, research and implementation. In order to scrutinize the AI services' accountability, risks and safety. These points should be considered as critical information for members of the public, when communicating AI-based services.

Principles of communicating AI-based services

As such, ethical principles for communicating AI-based services are:

- The principle of transparency – of the technology's workings, development and implementation.
- The principle of democratization – of the technology's amenability and accessibility.

Balancing Principles of communicating AI-based services

As was the case with TARES and eCTA, again these principles conflict with autonomy, justice and in turn non-maleficence. How should communication practitioners be transparent about their products and services and democratize i.e. information on their products and services while a big chunk of the consumers is currently unknowledgeable about AI (Pega, 2018)? Again, the key here is to make sure to be transparent and democratize i.e. information for the largest group of competent consumers. Additionally, make sure messages include incentives for competent consumers to educate non-competent consumers through dialogue.

Justifying Principles of communicating AI-based services

This section describes the derived principles of AI-based services for their accordance with the four principles of principlism.

Accordance with principle of autonomy

Competency is fairly accounted for by the democratization principle since it demands AI to be accessible and amendable to everyone. Second, the principle of transparency demands the technology to be transparent so that e.g., conscious informed consent and intention can be executed. Finally, understanding and freedom can be accounted for through the principle of democratization. Since it demands democratic education and understandability of the technology so that users know for themselves how to evaluate an AI-based service and decide when it is i.e., too coercive.

Accordance with principles of Beneficence and Nonmaleficence

Transparency of AI enabled services fosters the ability to hold someone or something accountable when something goes wrong. Which is needed because the ability to use AI maliciously is unavoidable, yet it should never be one's intention. Democratization is i.a. about democratic education on AI so that users know when AI is best used in a nonmaleficence manner. Transparency of AI-based services also enables users to know when a product is beneficial for them. Which, in turn, should foster democratization of the product. As such democratized, beneficial AI products contribute to the welfare of others by definition and as such beneficence and nonmaleficence can be accounted for.

Accordance with principle of Justice

Here, the principle of democratization is inherent to justice, as democratization is among other things formulated since the moral implications of AI demand that AI products should 'fairly distribute benefits, risks, and costs.'

The ethical framework

In the previous sections principlism has enabled us to formulate the principles critical for formulating ethical guidelines in marketing communicating AI-based services. Together, these principles form an ethical framework covering the dimensions: ethical principles of marketing communication, technology and communicating AI-based services.

Principles regarding ethics of marketing communication are Baker and Martinson's (2001) TARES test:

- Truthfulness – of the message (honesty, trustworthiness, non-deceptiveness)
- Authenticity – of the persuader (genuineness, integrity, ethical character, appropriate loyalty)
- Respect – for the persuade (regard for dignity, rights, well-being)
- Equity – of the content and execution of the appeal (fairness, justice, nonexploitation of vulnerability)
- Social responsibility – for the common good (concern for the broad public interest and welfare more than simply selfish self-interest)

Second, principles for ethics of technology development can be drawn from the eCTA framework:

- The embodiment principle
- The hermeneutic principle
- The alterity principle
- The background relation principle

Third, principles for ethics of communicating AI-based services are:

- The principle of transparency – of the technology's workings, development and implementation.
- The principle of democratization – of the technology's amenability and accessibility.

Additionally, the balancing of the above guidelines yielded an additional principle of dialogue. To inform how one should act in conflicting situations:

- In a situation when there is a risk of misrecognition, the principles should be aimed at the largest competent group while motivating them to start a dialogue with non-competent group members.

2.3 Preliminary guidelines for marketing communication of AI-based services

In order to find appropriate guidelines, a first draft of the guidelines should be developed. First, the principles from the ethical framework (§ 2.2) will be coupled in table 2 in order to guide the formulation process. Second, these formulations could be formulated with less terminology and collectively to view them as a set of guidelines[3].

Table 2

Aggregation of ethical framework and formulation of preliminary guidelines

		TARES principles[b]:	
		Respect	
		Equity	Authenticity
		Social responsibility	Truthfulness
		Principles of AI enabled services[c]:	
eCTA based principles[a]:		Democratization	Transparency
The embodiment principle	Locus	1. "The locus of marketing activities (physical, cognitive or contextual) should be democratically amenable and interpretable for the democratic amenability of the AI service, while being insusceptible to exploitation of vulnerability."	2. "The locus of marketing activities (physical, cognitive or contextual) should be interpretable for the AI service' transparency and integrity, while being insusceptible to exploitation of vulnerability."
	Form	3. "The form of the marketing activity (coercive, persuasive, seductive or decisive) should be democratically amenable and interpretable for the democratic amenability of the AI service, while being insusceptible to exploitation of vulnerability."	4. "The form of the marketing activity (coercive, persuasive, seductive or decisive) should interpretable for the AI service' transparency and integrity, while being insusceptible to exploitation of vulnerability.".
	Domain	5. "The marketing activity should be democratically amenable and interpretable for how the AI service' informs its democratic amenability from individual -and social perceptions, while being insusceptible to exploitation of vulnerability. "	6. "The marketing activity should be interpretable for how the AI service' informs its integrity for individual -and social perceptions, while being insusceptible to exploitation of vulnerability."
The hermeneutic principle		7. "The framework wherein the marketing activity is presented should be democratically amenable, interpretable for the democratic amenability of the AI service, while being insusceptible to exploitation of vulnerability."	8. "The framework wherein the marketing activity is presented should be interpretable to the transparency and integrity of the AI service, while being insusceptible to exploitation of vulnerability."
The alterity principle		9. "The marketing activity should be accompanied in its design to be democratically amenable and be interpretable for the AI service' democratic amenability, while being insusceptible to exploitation of vulnerability."	10. "The marketing activity should accompanied in its design to be interpretable about the AI service' integrity, while being insusceptible to exploitation of vulnerability."
The background principle		11. "The marketing activity should democratically amenable on how the AI based service shapes our daily lives, without being susceptible to exploitation of vulnerability."	12. "The marketing activity should interpretable and transparant for how the AI based service shapes our daily lives, without being susceptible to exploitation of vulnerability."

Note. [a]Kiran et al., 2015;[b]Baker & Martinson, 2001[c]See §2.2

The eCTA framework guides designers through a thinking process of accompanying their design with ethics (Kiran et al., 2015). Similarly, the eCTA frameworks' aptitude is used for guiding the guidelines' focal points on the left in table 2. Additionally, Verbeek's framework for anticipating technology (2013) is incorporated in the embodiment principle to help anticipate systematically the embodied impact of AI services. Secondly, the action-guiding TARES principles (Baker & Martinson, 2001) are situated on top of the table in order to determine the contents. The principles of communicating AI-based services are used as well to aid the

3 Please note that this formulation was substantiated by an older version of this report. Which was evaluated as well in de e-Delphi. The current report is adjusted after the e-Delphi. You can view the original report upon request.

formulation of the guidelines. They are coupled under the TARES principles since they can be synthesized on fairly the same premises. Finally, the four TARES principles and the AI-based services principles can be coupled under respect since this principle "is at the heart of TARES Test, and is the underlying foundation (…) for all of its other principles" (Baker & Martinson, 2001, pp. 163).

First, guideline 1 through 4 and 7 and 8 concern more or less the democratic and transparent design choices of the marketing communication activity. These guidelines could be aggregated accordingly:

1. Democratic and transparent amenability

The locus, form and framework of the activity should be democratically amenable, while being insusceptible to exploitation of vulnerability. Moreover, the activity should be accompanied in its design to be constituted as such.

Second, guidelines 9 and 10 concern the level of interpretability of the design choices, these guidelines could be aggregated as follows:

2. Interpretability

The activity should be interpretable for the AI service' transparency, integrity and democratic amenability and accompanied in its design to be constituted as such.

Third, guideline 5 and 6 concern the effects the activity might have on individuals and/or social life:

3. Social effects

The activity should be interpretable for how the AI service' informs its democratic amenability and integrity from individual -and social perceptions.

Finally, guideline 11 and 12 concern the constitutive aptitudes of the activity:

4. Constitutive effects

The activity should be interpretable on to what extent the AI service can constitute its democratic amenability and integrity in daily life.

In order develop appropriate guidelines, these guidelines need to be further examined and evaluated by a panel of experts. Moreover, their applicability needs to be assessed among the intended end-users. Accordingly, in the next chapter the methods used to develop these guidelines further and to evaluate them are described[4].

4 View appendix B, for the final guideline document, which was substantiated on the current report.

3. Methodology

In the previous sections principlism was applied as method for developing an ethical framework from which preliminary guidelines were drawn. In order to evaluate the used method and to further develop the guidelines, an expert study was conducted. Secondly, the final guidelines need to be evaluated for their appropriateness as practical guidelines for professionals. Accordingly, a professional study is conducted as well.

3.1 Expert study

In order to obtain an expert point of view, an Electronic Delphi (e-Delphi) was used. This method was originally developed to assess the future implications of new technologies (Linstune & Turoff, 2002). Additionally, this method is especially appraised for cases with wider social implications such as technological influences on nature and society (Beachamp & Childress, 2001). Finally, the structural approach of the e-Delphi makes it eligible for repetition, useful for future research (Linstone & Turoff, 2002).

Applications of the Delphi methods vary, an Electronic Delphi (e-Delphi) was used to reduce costs and time investments, but more importantly, to obtain greater control over anonymized responses (Hall, Smith, Hefferman & Fackrell, 2018; Thangaratinam & Redman, 2005). The questionnaires sent to the participants via e-mail or social media were anonymously taken. As such, potential bias when participants are influenced by i.e. social status from other participants is eliminated (Thangaratinam & Redman, 2005). The method (usually) involves three rounds of questionnaires among a panel of participants. After each round the questions were adjusted to the answers of the previous round and the panel was provided with feedback from each other on the whole set of responses. Accordingly, participants may change their views weighing the information from other participants and eventually reach consensus (Thangaratinam & Redman, 2005).

Crucial to e-Delphi are the questionnaires, here the AGREE collaboration instrument for assessing guidelines is used (AGREE Next Steps Consortium, 2017), applicable to e-Delphi (e.g., Msibi, Mogale, de Waal & Ngcobo, 2018). Moreover, the instrument has been adopted by various organizations such as the World Health Organization (Cluzeau, 2003). Studies have shown that the AGREE II instrument evaluates key elements crucial to evaluating guidelines (Harbour & Miller, 2001; Semlitsch, Blank, Kopp, Siering & Siebnhofer, 2015). The instrument evaluates multiple domains, of which the scope & purpose, stakeholder involvement, rigour of development, clarity of presentation, applicability and editorial independence.

Relevant for the goal of this study, since ‘Scope and purpose’ tells us something about the rationale of the guidelines, describing the effectivity of the descriptive function of principlism. Second, ‘Rigour of development’ tells something about the method of principlism itself. The other domains are relevant for the guidelines. While the AGREE II can also be used to develop guidelines, the instrument within this study is only used for evaluating. Questionnaires were developed using Qualtrics (Qualtrics, Provo, UT).

The experts were asked to rate 23 items on a 7-point Likert scale (1 is strongly disagree, 7 is strongly agree) within the six domains. When experts agreed with the statements, the item is positively rated. In agreement with the AGREE user manual, some items were altered or removed when necessary (AGREE Next Steps Consortium, 2017), see table 3.

Table 3
Modifications and/or removals of items from the AGREE II instrument

Original items	Modifications	Removed
The health question(s) covered by the guideline is (are) specifically described.	The ethical question(s) covered by the guideline is (are) specifically described.[a]	
The guideline development group includes individuals from all relevant professional groups.		Removed since the development was done utilizing principlism and literature research
The methods for formulating the recommendations are clearly described.	The methods for formulating the guidelines are clearly described.[b]	
The health benefits, side effects, and risks have been considered in formulating the recommendations.	The health benefits, side effects, and risks have been considered in formulating the guidelines.[c]	
There is an explicit link between the recommendations and the supporting evidence.	There is an explicit link between the guidelines and the supporting evidence.[d]	
The guideline has been externally reviewed by experts prior to its publication.		Removed since the guidelines are not yet meant to be publicized.
A procedure for updating the guideline is provided.		Removed since this item considers publicized guidelines.
The recommendations are specific and unambiguous.	The guidelines are specific and unambiguous.[e]	
The different options for management of the condition or health issue are clearly presented.		Removed since this item considers explicit health issues rather than ethical.
Key recommendations are easily identifiable.	Key points of the guidelines are easily identifiable.[f]	
The guideline presents monitoring and/or auditing criteria.		Removed since this these criteria are out of the scope of the guidelines.

Note. [a]The guidelines consider ethical questions rather than health questions.[b]The guidelines are the product of agregating the reflections and principles, rather than recommendations.[c]Idem.[d]Idem.[e]Idem.[f].Idem.

During the survey, the experts were supplemented with guideline documents to inform their answers. Of which, a summary of the development of the guidelines (appendix A) and a document with an (extended) copy of the first two chapters of this study[5], enabling the experts to review the method of development. Upon recommendation of the experts after the first round, they were also provided with a whitepaper including an introductory text and the guidelines (appendix B) in the second round.

After the first-round, 7 rated questions were not asked again for the second round. Ratings below 7 will be scrutinized for their respective comments. These comments were incorporated in the reformulation of the guidelines and their rationale (first two chapters of this study). The motives for adjustments (on account of these comments) were open to review for the experts. The adjusted guidelines and guideline documents were assessed again by the experts and after two rounds the e-Delphi ended and the results were analysed. The whitepaper was not adjusted again, this version was used for evaluation among the professionals (§3.2).

Expert panel

An e-Delphi method does not require a fixed number of participants necessarily (Thangaratinam & Redman, 2005) as "representativeness (…) is assessed on the qualities of the expert panel rather than its number" (Powell, 2002, pp. 378). Quality expert panels can be defined by heterogeneity, varying personalities, differing perspectives and backgrounds (Delbecq, Van de Ven & Gustafson, 1975; Murphy et al., 1998; Rowe, 1994). As such, a larger number of groups could, however, constitute larger variety among the expert panels. However, considering the novelty of this study, it would be a waste of time and effort to gather a large group of participants if the elemental premises on which the guidelines are build, are drafted incorrectly. As such, the panel was comprised of 5 experts.

5 The experts were supplemented with an older version of this report. Which was evaluated as well in de e-Delphi. The current report is adjusted after the e-Delphi. You can view the original report upon request.

Expert panel criteria

Considering the three dimensions mentioned in section 2, also three different areas of expertise are needed: AI, marketing communication and technology ethics experts. Since these knowledge domains are somewhat overlapping it is difficult to find experts exactly matching these domains. As such, the relevant knowledge criteria needed are defined below of which the expert should at least have one. While criterium 2 through 4 are based on the three knowledge domains, criterium 1 and 5 consider some of the stakeholder involvement and guideline development.

1. The expert should at least be knowledgeable of AI's development and implementation process.
2. The expert should at least be knowledgeable of AI's ethical implications such as:
    - Transparency;
    - Democratization.
3. The expert should at least be knowledgeable of the ethical implications of marketing communication such as:
    - Communication as constitutive of our perceived reality and epistemic perspectives;
    - The role of persuasive communication (marketing) in constituting our perceived reality and our epistemic perspectives.
4. The expert should at least be knowledgeable of the ethical implications of technology such as:
    - Technologies' mediating competences;
    - Technologies' constituting competences.
5. The expert should at least be knowledgeable of how (ethical) guidelines ought to be constructed.

In order to prevent biased assessing, only experts from western cultures are selected. The critical guidelines are predominantly conceptualized from a western perspective, using western ethical beliefs. Although the principles of principlism ought to be applicable to all cultures (Beauchamp & Childress, 2009), the developed critical guidelines might not.
The selected experts and the selection rationale should be described (Thangaratinam & Redman, 2005). Accordingly, table 4 displays their profession and corresponding knowledge criteria. Which was drawn from their work and/or publications.

Table 4
Selected experts

Experts	Profession	Knowledge criteria
Technical leader & advisor IBM	Consumer Industry Architect & Data and AI ethics lead for IBM Benelux	1, 2 & 4
Prof. Dr. of Law	Full Professor of Private Law & Vice Dean for Research at Tilburg Law School	3 & 5
Dr. of Communication Ethics	Associate Professor in English (Expressive Arts) Communication and Media Studies	3
Dr. of Technology Ethics	Assistant professor at the Philosophy and Ethics group at the department of Industrial Engineering and Innovation Sciences	3 , 4 & 5
Dr. of Communication	Associate Professor Communication Science	3

3.2 Professional study

In order to assess the guidelines' appropriateness, it is critical to evaluate the guidelines attributes among end-users since they should feel motivated to apply these guidelines. To wit, within clinical settings the careful, academic construction of guidelines have not led to consistent application of guidelines (Brouwers, Graham, Hanna, Cameron & Browman, 2004). Brouwers et al. (2004) refer to well-known studies (Fishbein & Ajzen, 1975; Fishbein & Ajzen, 1980; Rogers, 2003) describing how users' beliefs towards an instrument's attributes influences one's intention to use it and one's actual behaviour towards using it. Accordingly, Brouwers et al. (2004) developed an instrument measuring guideline users' belief. Which will be applied within this work.

Within the context of this study, a variation of end-users is possible (see §3.3.2). Considering this, an instrument developed by Shoemaker, Wolf & Brach (2014) is used as well for the questionnaire. Which takes in to account "consumers of diverse background" (p. 3). The instrument measures understandability and actionability of print and audio-visual materials. With the help of an expert panel, Shoemaker et al. (2014) defined understandability and actionability as following:

> Understandability: Patient education materials are understandable when consumers of diverse backgrounds and varying levels of health literacy can process and explain key messages.
>
> Actionability: Patient education materials are actionable when consumers of diverse backgrounds and varying levels of health literacy can identify what they can do based on the information presented (p. 3).

Accordingly, the instruments of Brouwers et al. (2004) and Shoemaker et al. (2014) are combined for this study. Not all items of the instruments were relevant, table 5 displays the aggregated instrument.
For measuring understandability, actionability, acceptance and applicability a five-point Likert-scale was used (1 is strongly disagree – 5 is strongly agree). For the comparative value and outcome variables (overall rating), both five-point Likert-scale as multiple choice questions were asked. Aside from the instrument, the respondents were also questioned demographic details and were able to give a final comment. The questionnaire was developed with Qualtrics as well (Qualtrics, Provo, UT).

Table 5
Compounded instrument

Components	Items
Understandability[a]	The Guideline Whitepaper makes its purpose completely evident
	The Guideline Whitepaper does not include information or content that distracts from its purpose.
	The Guideline Whitepaper uses common, everyday language.
	The Guideline Whitepaper uses the active voice.
	The Guideline Whitepaper presents informations in a logical sequence.
	The Guideline Whitepaper uses visual cues (e.g., arrows, boxes, bullets, bold, larget font, highlighting) to draw attention to key points.
	The Guideline Whitepaper visual aids reinforce rather than distract from the content.
Actionability[a]	The Guideline Whitepaper clearly identifies at least one action the user can take.
	The Guideline Whitepaper addresses the user directly when describing actions.
	The Guideline Whitepaper breaks down any action into manageable, explicit steps.
	The Guideline Whitepaper provides a tangible tool (e.g. menu planners, checklists) whenever it could help the user take action.
	The Guideline Whitepaper uses visual aids whenever they could make it easier to act on the instrutions.
Acceptance of guidelines[b]	I agree with the guidelines as stated.
	The guidelines are suitable for the consumers for whom they are intended.
	When applied, the guidelines will produce more benefits for my consumers than harms.
	The guidelines are likely to be supported by a majority of my collegues.
	If I follow the guidelines, the expected effects on consumers outcomes will be obvious.
Applicability of guidelines[b]	The guidelines are too rigid to apply to individual consumers.
	To apply the guidelines will require reorganization of service/care in my organization.
	To apply the guidelines will be technically challenging.
	The guidelines are too expensive to apply.
Comparative Value[b]	The guidelines reflect a more effective approach for improving patient outcomes than current usual practice.
	When applied, the guidelines will result in better use of resources than current usual practice.
Outcome Variables[b]	This Guideline Whitepaper should be approved as practice guidelines.
	If this Guideline Whitepaper were to become practice guidelines, how likely would you be to make use of it in your own organization?
	If I follow the guidelines, the expected effects on consumers outcomes will be obvious.

Note.Questions measuren on a 5-point Likert-scale (1=strongly disagree, 5=strongly agree). [a]Components and their respective items are selected and/or adapted from Shoemaker et al. (2014). These questions are negative (1=stronlgy agree, 5=strongly disagree). [b]Components and their respective items are selected and/or adapted from Brouwers et al. (2004).

Professionals

The focus in this work is a first step in to developing guidelines for AI-based services. Accordingly, it is acceptable to work with a smaller sample size and a broader margin of error. A search on Linkedin for marketeers around the world yielded 10,7 million results. Considering not everyone will be working with AI, an estimate of a population of 10 million communication professionals was assumed. With a confidence interval of 95% and a margin of error of 15%, 43 respondents were required.

Three types of end-users among the respondents can be defined. First, marketing professionals of third-party marketing bureaus were considered. Second, marketing professionals within companies developing AI-based services were considered. Finally, student marketeers were considered as well to include a wide variety of marketeers ranging from junior to senior over a variety of experience fields. Of the 43 participants 32 were male and 11 female. The participants spanned ages <18 – 44, one was under 18, 17 were in the range of 18 – 24, 20 in the range of 25-34, and five were ranged 25 – 44. The education spanned high school degree to doctorate. Of which three had a high school or similar degree, 25 had or were doing a bachelor's degree, 14 had or were doing their master's degree and one participant had or was doing his doctorates degree. Two participants were retired or unemployed, 16 were students, 13 were working as a marketeer within a marketing agency, four were marketeers within a technology related company and eight were a marketeer in another type of company.

Multiple ways of finding respondents was used. Suitable respondents were recruited via social media, personal networks, approach of relevant companies, as well as other Universities and student associations. Respondents were also reached through pamphlets around University campuses. All methods were supplied with an introductory text and the anonymous link to the survey. Since these methods could reach students or professionals who are not (student) communication professionals, the respondents in the survey were asked whether they are student marketeer "or otherwise identify as a communication marketing professional", if no was selected the survey ended.

4. Results

First, the results of the expert study are discussed. Second, the results of the professional study are discussed. Finally, the results are summarized and the answers to the research questions are stated.

4.1 Expert study

Considering the small size of the expert panel, conclusions are drawn from the comments the experts made. The results from the Likert-scales are merely used as indications about the expert's beliefs of the guidelines. First, the average Likert-scale score per appraiser over the two e-Delphi rounds is shown, as well as their overall ratings and whether or not they would recommend these guidelines (table 6). Second, the average and percentage scores per domain is shown (table 7), while discussing the comments to which the guidelines and their rationale were altered over the consecutive rounds.

Overall

After the first round two of five experts agreed with statements evaluating the guidelines positively. However, overall the ratings outcomes were not positive (M overall score = 3.8). Drawing from table 6, all the appraisers indicated the guidelines improved in the second round. Finally, the appraisers recommended the guidelines for use but also advised to further improve the guidelines.

Table 6
AGREE II Scores per appraiser

		Round 1			Round 2		
Appraiser	Expertise[a]	Overal rating[b]	M	Recommendation	Overal rating[b]	M	Recommendation
Technical leader & advisor IBM	Big data & AI ethics advisor IBM Benelux	4	3.6	Yes, with modifications	4	4.1	Yes, with modifications
Prof. Dr. of Law	Private Law	5	4.2	Yes, with modifications	5	4.8	Yes
Dr. of Communication Ethics	Media and media ethics	2	1.2	No	5	3.7[c]	Yes, with modifications
Dr. of Technology Ethics	Technology and philosophy	3	3.6	No	5	5.3	Yes, with modifications
Dr. of Communication	Communication	5	4.2	Yes, with modifications	5	5.1	Yes
Overall score	n/a	3.8	3.4	n/a	n/a	4.6	n/a

Note. A 7-point Liker scale (1= strongly disagree, 7=strongly agree) was used for all items, except for the final two items which include recommendation for use and a final oppurtunity to comment.[a]See section 3.1 for a specific description of the expert panel expertises. [b]These are the single answers for the final item, asking the participants to give an overall rating. [c]Some items were missing, these were counted as 0.

Domain scores round 1

In the first round, applicability and clarity of presentation were rated lowest. The experts found the guidelines too abstract and therefore, not applicable for use. Stakeholder involvement was rated below the median score (M = 3.3) as well. The experts noted that the research consisted solely of literature research. And advised additional research should be done among marketeers.

Highest scores were among editorial independence, scope and purpose and rigour of development. The experts noted that there was not much information about the editorial independence, but did not suspected conflict of interest per se. Considering scope and purpose, the experts indicated that it was clear the guidelines are meant to improve marketing of AI-based services particularly to safeguarding moral values. However, they also indicated that the language was often hard to follow. Considering the use of principlism, the Prof. Dr. of Law suggested that the process of balancing should be added to the reflection process to

improve the scope and purpose. Moreover, a theory of justice was not consulted. The experts did find the development "comprehensive". The Dr. of Technology Ethics denoted the combination of ethical fields "plausible". However, the Dr. of Communication Ethics suggested the ethical scope and ethical model for communication could be improved.

Table 7
Average domain scores

Domain	Round 1		Round 2[a]	
	M	% of total score[b]	M	% of total score[b]
Scope & Purpose	3.6	48%	5.0	67%
Stakeholder involvement	3.3	38%	4.1	52%
Rigour of development	3.8	47%	4.6	59%
Clarity of presentation	2.7	28%	5.2	70%
Applicability	2.4	23%	4.7	61%
Editorial independence	4.0	50%	3.9	48%
Overall	3.4	40%	4.6	60%

Note. [a]Some items were missing, these were calculated as 0.[b]Rounded to whole percentages, calculated according to the AGREE II user manual (2017). Measured on a 7-point Likert-scale(1=strongly disagree, 7=srongly agree).

Specifically, the feedback was divided in confirming, questioning, repudiating, advising and suggesting (table 8). Most of the comments were repudiating and gave direction to what was fundamentally wrong in the eyes of the experts. The second largest group of comments entailed advisements, which were mostly directions to certain literature or subjects which needed to be considered.
Most of the comments were about the formulation of the critical guidelines. Which was evaluated dependent on "technical vocabulary" and should be "more accessible for practical use", while the language should be "simplified". Moreover, the panel recommended to refer to "ethical values" and aim "for the same degree of clarity as the TARES principles." Moreover, the term "interpretable" should be avoided and "activity" should be clarified to reduce vagueness.

Table 8
E-Delphi comments frequency

Appraisers	Round 1					
	Confirming	Questioning	Repudiating	Advising	Suggesting	n Comments
Technical leader & advisor IBM	1	0	2	2	1	6
Prof. Dr. of Law	8	4	7	5	0	24
Dr. of Communication Ethics	1	3	15	6	4	29
Dr. of Technology Ethics	4	3	9	4	2	22
Dr. of Communication	0	0	4	2	0	6
Total	14	10	37	19	7	87
	Round 2					
	Confirming	Questioning	Repudiating	Advising	Suggesting	n Comments
Technical leader & advisor IBM	10	3	1	1	0	15
Prof. Dr. of Law	5	0	1	0	2	8
Dr. of Communication Ethics	2	9	4	4	14	33
Dr. of Technology Ethics	0	0	0	0	0	0
Dr. of Communication	0	0	0	0	0	0
Total	17	12	7	5	16	56

Adjustments

On behalf of the described comments the following adjustments were made. The rationale (first two chapters) was supplemented with the discussion of justice at recommendation of the Dr. of Communication Ethics. The substantiation on the use of principlism was improved at recommendation of Prof. Dr. of Law, as well as a description of balancing. The guidelines were now written within a whitepaper to clarify how one should use them to improve the applicability. Moreover, the guideline formulations were improved at recommendation of all experts.

Domain scores round 2

After applying the above described adjustments, the guidelines were most improved among the domains clarity of presentation and applicability. The other domains were improved as well. Except for editorial independence which decreased slightly. Overall, the guidelines scored 60% of the highest possible score among all the domains, with a mean answer improving from 3.4 to 4.6.

The Prof. Dr. of Law found it: “Very important to address ethics in AI marketing. These guidelines are a useful first step.” “(…) a lot more pragmatic compared to earlier version. But I think, next to your paper, in order for marketeers to really use it in daily activities, it should go to even more pragmatic to apply on certain texts/ campaigns/ etc.” Which is in line with Dr. of Communication Ethics, who also suggested end-users should

be helped thinking about specific applied scenarios. Experts indicated the rigour of development improved as well and were happy with the added literature of justice. As well as with the substantiation of methods for development. Suggestions were given as well. For instance, the guidelines were considered specific and unambiguous, "except for the last principle, which is ambiguous". The technical advisor for IBM indicated that they were not specific, but "leave room for interpretation. Which is ok in my view". Referring to the IBM principles of Data, Trust and Transparency[6] (personal communication, January 10, 2019).
Interestingly, the Dr. of Communication Ethics graded a 1 in both versions for the item asking whether potential resource implications were considered. Explaining her answer in the second round with:

> "Marketers will absolutely decry the extra cost involved in dialogue, extra research to ensure truthfulness, etc. And who pays their budget? The manufacturers. So, there's a whole other ethical quandary right there. If they want to keep their marketing company afloat, they need to be competitive. But to be ethical costs money. Will manufacturers choose the more ethical marketing company, or the cheaper one?"

The experts believed some ethical theories or guidelines were too easily dismissed or not adequately substantiated. And still had questions regarding to whom the guidelines exactly apply, e.g. consumers or patients as well. Moreover, the Dr. of Communication Ethics held that the views and preferences of the target population were not adequately sought in the research. However, the dialogue principle of the guidelines is adequately taking in account of all stakeholders if this would be a ubiquitous requirement, not only when the consumers are vulnerable.

Looking at RQ 1.1, the experts believed the methods used for formulating the guidelines were appropriate. They found it comprehensive, but lacking some substantiation. Already after the first round a lot of substantiation was needed to be added to the ethical scope and the application of principlism. Accordingly, the method is open for further improvement.
Considering RQ 1, they found the guidelines appropriate. They would advise them for use, but with modifications and indicated the guidelines should be treated as a first step in to development of actual, applicable guidelines.

6 See for instance https://www.ibm.com/blogs/policy/trust-principles/

4.2 Professional study

To measure the appropriateness of the guidelines, the understandability, actionability, acceptance, applicability, outcome variables (table 9) and comparative value (table 10) were measured. The scores for comparative value are displayed separately, since this component was measured differently than the others.

Table 9

Descriptive statistics for Understandability, Actionability, Acceptance, Applicability and Outcome variables

	N	*N of items*	*M*	*SD*
Understandability	43	7	3.71	0.43
Actionability	43	5	3.34	0.55
Acceptance	42[a]	5	3.94	0.53
Applicability[b]	43	4	2.90	0.67
Outcome variables	43	3	3.47	0.67

Note. [a]One answer for Q22 was missing. [b]Scores are negative, lower is more applicable (see table 5). Measured on a 5-point Likert-scale(1=strongly disagree, 5=strongly agree).

Drawing from table 9, the respondents evaluated the guidelines understandable, actionable, agreed with the guidelines, thought they were applicable and the outcomes of applying the guidelines appeared to be clear to them. However, the mean answers for these components are not compelling and just higher than the median score, especially for applicability.

Table 10

Statistics for comparative value

	Questions [a]	*N*	*Yes*	*Unsure*	*No*
Frequencies	Q30 More effective approach for improving	43	34	9	0
		N	*M*	*SD*	
Descriptives	Q31 Better use of resources than current usual practice[b]	43	3.67	0.94	

Note. [a]Questions are shortened for displaying purposes. [b]Measured on a 5-point Likert-scale (1=strongly disagree, 5=strongly agree).

Looking at table 10, the respondents found the guidelines a more effective approach for improving than what is standard. Respondents believed the guidelines are a better use of resources than current practices (M = 3.67). However, the standard deviation among the answers for question 31 were considerable. Since there is a lot of variation among all the answers, this could be an indication that the questions were not valid or too much variation among individuals within the sample size. For example, in background knowledge or proficiency in English.

To illustrate, all respondents were asked to give some final remarks. There were 12 respondents who commented. The comments translate somewhat the variations among the answers discussed above. Comments included that the guidelines were nicely composed and to the point, addressing all issues and clear about the implications of AI-based products. Others said they were too theoretic, vague in implementation and how to apply in a practical manner.

One respondent summarized his answer by “the author has put great care and time in crafting a concise message, he might be blind sighted in how hard this information is digested by new readers.”

Drawing from the end-user questionnaire, the guidelines are fairly appropriate but only as an outset of actual guidelines. The respondents found the guidelines understandable, actionable, applicable, agreed with the guidelines and found the guidelines represent a better and more effective approach than current standards. However, there was a lot of variation among these answers. Especially for applicability, the comparative value and the outcome variables.

Review all answers to all the RQ's in table 11 below.

Table 11
Research question outcomes

Research question	Answer
RQ 1. "What are appropriate guidelines for ethical marketing communication of AI-based services?"	Consider below the set of six appropriate guidelines. 1. Fair and understandable distribution of the marketing activity its design, presentation and framing (equal accessibility for all). 2. Truthful distribution of the marketing activity its information on the AI-based services its transparency, integrity and extent of equal accessibility (sincere transparency). 3. Informative of social effects. The marketing activity should be informative for how the AI-based service is equally accessible and honest for individuals, and within social situations (social responsibility). 4. Informative of contextual effects. The marketing activity should be equally accessible and transparent on how the AI-based service will shape our daily lives (contextual transparency). 5. Promote dialogue between consumers on the marketing activity its information, when there is a danger of exclusion of vulnerable consumer due to complex information. 6. Accompaniment of the above ethical values in the design of the marketing activity, without taking advantage of vulnerabilities.
RQ 1.1 "What is an appropriate method for developing an ethical framework from which marketing communication guidelines for AI-based services can be drawn?"	Principlism is an appropriate method to develop an ethical framework (§4.1 & §5).

5. Discussion

The purpose of this work was to find appropriate guidelines for the marketing communication of AI-based services. In this quest, a method was found to develop an ethical framework from which the guidelines could be drawn. The measurement of the appropriateness of the method principlism, and additional refinement of the guidelines were done through an expert study using e-Delphi and the AGREE II instrument (AGREE Next Steps Consortium, 2017). The expert panel consisted of experts in the fields of AI ethics and development, technology ethics, communication ethics, ethics and law and communication. The methods used for describing the ethical framework and drawing the guidelines, was evaluated as appropriate. Experts believed the methods were systematic and comprehensive. Especially looking at the ethical and substantive underpinnings of the framework. However, some substantiations and choices for pieces of literature could be improved. Indicating the need for ethical theoretical knowledge in using this method. While the latter was supplemented through an e-Delphi, the e-Delphi was time consuming. Making the current practice not separable from academic recourses and experts. Accordingly, the methods used in this context could be further refined. For instance, by involving ethicist at the start of a such a research. The final set of guidelines were evaluated by the experts as appropriate considering they would advise these guidelines for use but with modifications. The guidelines could be further improved among the formulation of the guidelines and how these guidelines should be applied.

Next, the guidelines were evaluated through a professional study using an online questionnaire among marketing communication professionals. They evaluated the guidelines as understandable, actionable, applicable, agreed with the guidelines and found the guidelines represent an effective approach for improving. However, there was much variation among the answers. Indicating some problems with either the questions or too much individual variation among the sample size.

Some respondents clarified their answers in the comments by saying the guidelines were ambiguously formulated, but the message and the goal of the guidelines was clear to them. For instance, comments entailed, “too theoretic” or “vague in practical implementation” and hard to digest. While also stating that the guidelines were clear and concise about the implications of AI.

Accordingly, the guidelines can be viewed as appropriately. Meaning the guidelines are relevant in stipulating the elements critical for ethical marketing communication of AI-based services. The guidelines are, however, not yet applicable and should be improved for readability and applicability accordingly.

Since the guidelines are the indirect product of the formulation methods used, one should recognize the cohesion between these methods and the guidelines. The guidelines were drawn from a framework which was build using principlism. Next the framework is reframed in a table to guide the formulation of the guidelines and from there a first version of the guidelines is formulated. Which is in turn evaluated and adjusted through an e-Delphi questionnaire. As such, the process after describing a framework using principlism is pertinent for the development of guidelines and, possibly, for their quality. Future research should definitely re-evaluate this process. To wit, while the AGREE II instrument is appraised for its use of evaluating guidelines by renowned organizations, some experts were not sure they understood the questions adequately. Moreover, some questions needed to be altered to fit the context of the study. In addition, a larger panel size with more than two rounds for the e-Delphi could provide more information of the efficacy of used methods in developing guidelines for marketing communicating AI-based services. Especially considering the Likert-scale scores could not be used statistically in this work. A different method than e-Delphi could be used to evaluate principlism in this context as well. Since e-Delphi’s are known as time consuming and for their low retention rates (Hall et al., 2018; Semlitsch et al., 2015). While all participants were retained, not everyone commented on all questions. Especially in the second round some signs of fatigue among the experts appeared, as the total number of comments declined (table 8). In addition, with e-Delphi there is always the danger of response bias (Leary, 2014).

Some limitations to the quantitative method could be identified as well. Of the 111 reached respondents, 67 did not finish their survey. The topic of this work demanded the respondents to read a considerable amount of background information. This also meant only 43 respondents finished the survey, determining a relatively high margin of error (15%). Future research should look in to more efficient manners of evaluating guidelines for a topic with this level of knowledge required. Finally, when future research finds applicable guidelines, research should look in to concrete examples of how marketing communication should be designed using these guidelines. Next, future research should measure the outcome on consumers their attitude towards the products. By designing a marketing campaign for an AI-based service with guidelines and one without and measure the attitude towards the product among consumers. Accordingly, one could measure if the guidelines actually have the desired effect among consumers.

Other limitations include the obsoletions of literature used. A lot of governmental reports used in this work referred to the IEEE 2017 report. During this work a newer version was publicized, which included eight general principles. The original five general principles of the 2017 version was used in this work. Future research should take note of that. Moreover, some new governmental reports are publicized as well, also referring to the new IEEE report (i.e., EU, 2019).
Considering the conceptualization of justice in chapter 2, a few remarks on the 'what' and 'how' of justice need to be made. First, the guidelines inherently prescribe some sort of behaviour. As such, there is some danger people will be misrepresented as they had no participation in these decision-making structures. Second, this research has some sort of a scientific presumption in that the guidelines are evaluated by experts (see chapter 3). Meaning there is some danger of overseeing foreclosing claims of the disadvantaged.
One specific but irrefutable point was overlooked in this work. As pointed out by the Dr. of Communication Ethics, the guidelines should take in to consideration that making marketing campaigns more ethically acquiescent using guidelines, will definitely take more time and money. If there are no strict global regulations compelling companies to market their AI-based services ethically, how should marketeers and marketing companies design their campaigns ethically and cost effectively? Hence, it is still up for debate if manufacturers will choose the most cost-effective marketeer or the most ethical. Solutions could include making consumers aware of some sort of quality mark AI-based services should have. So that companies are intrinsically motivated to promote the ethicality of their products. Something Dr. van Wynsberghe, Prof. Sharkey and others are busy researching with their Foundation for Responsible Robotics (FRR).

Future research should refine principlism for this purpose and improve both the formulation process of guidelines after applying principlism. Accordingly, the transparency of the method can be improved, which in turn can help improve the guidelines readability an applicability. In addition, larger sample sizes in both qualitative and quantitative will improve the reliability of research outcomes. Allowing future research to steer in more definitive directions, to enable research of more concrete examples of what marketing communication looks like using the guidelines.

6. Conclusion

This work researched the question: "What are appropriate guidelines for ethical marketing communication of AI-based services?". To answer the research question, principlism was used to describe an ethical framework where preliminary guidelines were drawn from. This method was evaluated and the guidelines were further refined through an expert study using e-Delphi. The final set of guidelines was evaluated through an online questionnaire among professional marketeers.

Accordingly, this work proposes six appropriate guidelines for the marketing communication of AI-based services. These guidelines are appropriate in describing the elements critical in ethical marketing communication of AI-based services. However, the guidelines proved not yet to be applicable for use. As the topic of this research is still in its infancy, the guidelines should be further improved.

Nonetheless, I am convinced this work proves the urgency of ethics in marketing communication of AI-based services and is able to facilitate future research. Considering the prominent role of AI and several fields working towards ethical acquiescent AI-based services, I see significant value in expanding the field of communication science into ethically aligning our communications regarding AI and to contribute to an AI for good.

7. Acknowledgements

I want to acknowledge the contributions of: Ms. Kuijt, Prof. Dr. Mak, Dr. Tilley, Dr. Spahn and Prof. Dr. Reijmersdal.

References

AGREE Next Steps Consortium (2017). The AGREE II Instrument [Electronic version]. Retrieved from http://www.agreetrust.org.

American Association of Advertising Agencies. (1990). Standard of Practice of the American Associa¬tion of Advertising agencies. Retrieved from: https://ams.aaaa.org/eweb/upload/inside/standards.pdf

Antonov, A. (2018, March 8). Applying Artificial Intelligence and Machine Learning to Finance and Technology. Medium. Retrieved from https://towardsdatascience.com/applying-artificial-intelli¬gence-and-machine-learning-to-finance-and-technology-378cbd5e5c85

Armstrong, S., & Sotala, K. (2012). How We're Predicting AI--Or Failing To. Beyond AI: Artificial Dreams, 52–75. Pilsen: University of West Bohemia. Retrieved from: https://intel-ligence.org/files/Pre¬dictingAI.pdf.

Athanassoulis, N.(n.d.). Virtue Ethics. The Internet Encyclopedia of Philosophy, ISSN 2161-0003. Retrieved from https://www.iep.utm.edu/virtue/, 22-11-2019

Baker, S & Martinson, D.L. (2001). The TARES Test: Five Principles for Ethical Persuasion. Journal of Mass Media Ethics, 16, 148-175. doi: https://doi.org/10.1080/08900523.2001.9679610

Bandura, A. (2009). Social cognitive theory of mass communication. In Media effects (pp. 110-140). Routledge.

Barham, B. L., Chavas, J. P., Fitz, D., Salas, V. R., & Schechter, L. (2014). The roles of risk and ambigui¬ty in technology adoption. Journal of Economic Behavior & Organization, 97, 204-218.

Baum, S.D., Goertzel, B., & Goertzel, T.G. (2011). How Long until Human-Level AI? Results from an Expert Assessment. Technological Forecasting & Social Change, 78, 185–95. https://doi.org/10.1016/j.tech¬fore.2010.09.006

Beauchamp, T. & Childress, J. (2001). Principles of Biomedical Ethics. Oxford and New York: Oxford University Press.

Beauchamp, T. L. (2003). Methods and principles in biomedical ethics. Journal of Medical Ethics, 29(5), 269–274.

Beauchamp, T.L., & Childress, J. (2009). Principles of Biomedical Ethics. Oxford: Oxford University Press: 3.

Beauchamp, T. L., & Childress., J. F. (2013). Principles of biomedical ethics, seventh edition. New York: Oxford University Press.

Beauchamp, T.L., & DeGrazia, D. (2004). Principlism. In Khushf, G. (Eds.), Handbook of Bioethics: Tak¬ing Stock Of The Field From A Philosophical Perspective. (pp. 55-75). Dordrecht: Springer.

Beever, J., & Brightman, O.A. (2016). Reflexive Principlism as an Effective Approach for Developing Ethical Reasoning in Engineering. Science and Engineering Ethics, 22, 275-291. doi: 10.1007/s11948-015-9633-5

Bench-Capon, T.J.M. (2003). Persuasion in Practical Argument Using Value-based Argumentation Frameworks. Journal of Logic and Computation, 13(3), 429-448. doi:https://doi.org/10.1093/log¬com/13.3.429

Bench-Capon, T.J.M., & Atkinson, K. (2009). Abstract Argumentation and Values. In Rahwan, L. (Eds.), Argumentation in Artificial Intelligence (pp. 45-64). New York: Springer US.

Bivins, T. (2003). Mixed Media: Moral Distinctions in Advertising, Public Relations, and Journalism. Mahwah: Erlbaum.

Bostrom, N. & Yudkowsky, E. (2011). The ethics of Artificial Intelligence, in Frankish, K. (ed.) Cam¬bridge Handbook of Artificial Intelligence, Cambridge: Cambridge University Press

Bowen, S. A. (2004). Expansion of ethics as the tenth generic principle of public relations excellence: A Kantian theory and model for managing ethical issues. Journal of public relations research, 16(1), 65-92.

Boyle, J. (2004). Casuistry. In Khushf, G. (Eds.), Handbook of Bioethics: Taking Stock Of The Field From A Philosophical Perspective. (pp. 75-89). Dordrecht: Springer.

Broersen, J., Dastani, M., Hulstijn, J., Huang, Z., & van der Torre, L. (2001, January). The BOID Archi¬tecture – Conflicts Between Beliefs, Obligations, Intentions and Desires. Paper presented at the Fifth In¬ternational Conference on Autonomous Agents 2001, Montréal, Quebec. Retrieved from https://dspace.library.uu.nl/bitstream/handle/1874/19919/broersen_01_boid.pdf?sequence=1

Brouwers, M. C., Graham, I. D., Hanna, S. E., Cameron, D. A., & Browman, G. P. (2004). Clinicians' as¬sessments of practice guidelines in oncology: the CAPGO survey. International journal of technology assess¬ment in health care, 20(4), 421-426.

Brundage, M. (2017). My AI Forecasts--Past, Present, and Future. Retrieved from: http://www.miles¬brundage.com/1/post/2017/01/my-ai-forecasts-past-present-and-future-main-post.html.

Bulger, J.W. (2007). Principlism. Teaching Ethics, 8(1), 81-100. Retrieved from https://www.pdcnet.org/tej/content/tej_2007_0008_0001_0081_0100

Burrell, D., & Hauerwas, S. (1977). From system to story: An alternative pattern for rationality in Ethics. In The Foundations of Ethics and Its Relationship to Science: Knowledge, Value, and Belief (pp. 111-152).

Calo, R. (2018). AI Policy: A primer and roadmap. Retrieved from: http://dx.doi.org/10.2139/ssrn.3015350

China's State Council. (2017). A Next Generation Artificial Intelligence Development Plan. Translat¬ed by Flora Saplo (FLIA Scholar), Welming Chen (FLIA Research Assistant), and Adrian Lo (FLIA Research Intern) Foundation for Law & International Affairs. Retrieved from: https://flia.org/wp-content/up¬loads/2017/07/A-New-Generation-of-Artificial-Intelligence-Development-Plan-1.pdf

Clark, B. (2018, December 4). How Cambridge Analytics leveraged fashion to sway undecided voters. The Next Web. Retrieved from https://thenextweb.com/artificial-intelligence/2018/12/03/how-cam¬bridge-analytica-leveraged-fashion-to-sway-undecided-voters/

Clouser, K.D., & Gert, B. (2004) Common Morality. In Khushf, G. (Eds.),Handbook of Bioethics: Taking Stock of The Field From a Philosophical Perspective. (pp. 121-142). Dordrecht: Springer.

Cluzeau, F. (2003). Development and validation of an international appraisal instrument for assessing the quality of clinical practice guidelines: the AGREE project. Quality & Safety in Health Care, 12, 18-23. Retrieved from https://www.agreetrust.org/wp-content/uploads/2013/06/AGREE_Collaboration_2003.pdf

Conitzer, V., Sinnot-Armstrong, W., Schaich Borg J., Deng, Y., & Kramer, M. (2017). Moral Decision Making Frameworks for Artificial Intelligence. Paper presented at the thirty-first Conference on Artificial Intelligence (AAAI-17), San Francisco, California USA. Retrieved from https://aaai.org/ocs/index.php/AAAI/AAAI17/paper/view/14651

Corea, F. (2018, August 22). AI Knowledge Map: How To Classify AI Technologies. Forbes. Retrieved from https://www.forbes.com/sites/cognitiveworld/2018/08/22/ai-knowledge-map-how-to-classi¬fy-ai-technologies/#70ab4f5a7773

Dainton, M., & Zelley, E. D. (2017). Applying communication theory for professional life: A practical introduction. Sage publications.

Dance, F.E.X. (1970). The "concept" of communication. Journal of Communication, 20, 201-210.
Delbecq A.L., Van de Ven A.H. & Gustafson D.H. (1975). Group Techniques for Program Planning: A Guide to Nominal and E-Delphi Processes. Scott, Foresman and Co., Glenview, IL.

Desai, D.R., & Kroll, J.A. (2018). Trust but verify: A Guide To Algorithms and The Law. Harvard Journal of Law & Technology, 31, 1, 17-19. Retrieved from: https://ssrn.com/abstract=2959472

Diakopoulos, N. (2014). Algorithmic Accountability Reporting: On the Investigation of Black Boxes. Tow Center For Digital Journalism. doi: https://doi.org/10.7916/D8ZK5TW2

Diakopoulos, N. (2016). Accountability in algorithmic decision making. Communications of the ACM, 59, (2), 56–62. doi:10.1145/2844110

Dietterich, T. G., & Horvitz, E.J. (2015). Rise of Concerns about AI: Reflections and Directions. Commu¬nications of the ACM 58, (10), 38–40. Retrieved from: https://doi.org/10.1145/2770869

European Union. European Group on Ethics in Science and New Technologies. (2018a). Statement on: Artificial Intelligence, Robotics and 'Autonomous' Systems (doi:10.2777/S31856). Retrieved from: https://ec.europa.eu/digital-single-market/en/news/commission-presents-plan-make-most-artificial-intelligence

European Union. European Commission. (2018b). Smart lie-detection system to tighten EU's busy borders. Retrieved from http://ec.europa.eu/research/infocentre/article_en.cfm?artid=49726

European Union. European Commission. (2019). Ethics Guidelines for Trustworthy AI. Retrieved from https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Fain, N., Kline, M., & Duhovnik, J. (2011). Integrating R&D and Marketing in New Product Develop¬ment. Strojniski Vestnik, 7-8, 599-609. doi: 10.5545/sv-jme.2011.004

Fieser, J. (n.d.). Ethics. The Internet Encyclopedia of Philosophy, ISSN 2161-0003. Retrieved from https://www.iep.utm.edu/ethics/ , 23-05-2019.

Fishbein, M., & Ajzen, I. (1975). Belief, attitude, intention and behavior: An introduction to theory and research. Reading, MA: Addison-Wesley Publishing Company.

Fishbein, M., & Ajzen, I. (1980). Understanding attitudes and predicting social behavior.

Foster, A. D., & Rosenzweig, M. R. (2010). Microeconomics of technology adoption. Annu. Rev. Econ., 2(1), 395-424.

Fraser, N. (2008). The Scales of Justice. Cambridge: Polity Press`.

Future of Life Institute. (n.d.). Asilomar AI Principles. Retrieved from: https://futureoflife.org/ai-prin¬ciples/

Giles, M. (2019, February 14). Explainer: What is quantum communication? MIT Technology Review. Retrieved from https://www.technologyreview.com/s/612964/what-is-quantum-communications/?utm_campaign=site_visitor.unpaid.engagement&utm_source=hs_email&utm_medium=email&utm_con¬tent=70748911&_hsenc=p2ANqtz-8UzB1DG7jFNCBRJJrsMQF07Ze_dV_I-yzrFTSND3Jj6gcMwzO9QX¬Q4XrsVHaioH061hRAmMj0ggu7asmb7O8sd5g139g&_hsmi=70748911

Gordon, J., Rauprich, O., & Vollmann, J. (2011). Applying the four-principle approach. Bioethics, 25-6, pp. 293-300. doi:10.1111/j.1467-8519.2009.01757.x

Gourari, C. (2016, April 14). Investigating the algorithms that govern our lives. Columbia Journalism Review. Retrieved from https://www.cjr.org/innovations/investigating_algorithms.php?curator=MediaRE¬DEF

Grunwald, A. (1999). Technology Assessment or Ethics of Technology? Ethical Perspectives, 6(2), 170-182. doi:10.2143/EP.6.2.505355

Gunning, D. (2017). Explainable Artificial Intelligence (XAI). Retrieved from Defence Advanced Research Projects Agency website: https://www.darpa.mil/attachments/XAIProgramUpdate.pdf

Hall, D. A., Smith, H., Heffernan, E., & Fackrell, K. (2018). Recruiting and retaining participants in e-Delphi surveys for core outcome set development: Evaluating the COMiT'ID study. PloS one, 13(7), e0201378.

Hao, K. (2019, January 10). A neural network can learn to organize the world it sees into concepts just like we do. MIT Technology Review. Retrieved from https://www.technologyreview.com/s/612746/a-neural-network-can-learn-to-organize-the-world-it-sees-into-conceptsjust-like-we-do/

Harari, Y.N. (2018, January 24). Will the Future Be Human? Presentation at the World Economic Forum Annual Meeting. Davos-Klosters, Switzerland [Online video]. Available from https://www.weforum.org/events/world-economic-forum-annual-meeting-2018/sessions/will-the-future-be-human

Harbour, R., & Miller, J. (2001). A new system for grading recommendations in evidence based guide¬lines. Bmj, 323(7308), 334-336.

Hardt, M. (2014, September 26). How Big Data is unfair. Medium. Retrieved from https://medium.com/@mrtz/how-big-data-is-unfair-9aa544d739de

Harris, R.J., & Sanborn, F.W. (2014). A Cognitive Psychology of Mass Communication. New York: Rout¬ledge

Hollander, C.D., & Wu, A.S. (2011). The Current State of Normative Agent – Based Systems. Journal of Artificial Societies and Social Simulation, 14(2), 11-11.

House of Commons Science and Technology Committee. (2016). Robotics and Artificial Intelligence: Fifth Report of Session 2016–17. Retrieved from: https://www.publications.parlia-ment.uk/pa/cm201617/cmselect/cmsctech/145/145.pdf

International Association of Business Communicators. (n.d.). IABC Code of Ethics for Professional Communicators. Retrieved from https://www.iabc.com/about-us/purpose/code-of-ethics/

International Chamber of Commerce. (2018). ICC advertising and marketing communications code. Retrieved from https://cms.iccwbo.org/content/uploads/sites/3/2018/09/icc-advertising-and-market¬ing-communications-code-int.pdf

International Telecommunication Union (2017). AI for Good Global Summit Report. Retrieved from: https://www.itu.int/en/ITUT/AI/Documents/Report/AI_for_Good_Global_Summit_Report_2017.pdf

Jesus, C. (2017, January 1). Artificial Intelligence: What It Is and How It Really Works. Futurism. Re¬trieved from: https://futurism.com/1-evergreen-making-sense-of-terms-deep-learning-machine-learning-and-ai

Johannesen, R.L., Valde, K.S., & Whedbee, K.E. (2008). Ethics in Human Communication. Illinois: Waveland Press.

Kardes, F.R., Cronley, M.L., & Cline, T.W. (2014). Chapter 3 Branding Strategy and Consumer Behav¬ior. In F.R. Kardes, M.L. Cronley & T.W. Cline (Eds.), Consumer Behavior (pp. 69-91). Stamford, Connecticut: Cengage Learning.

Khushf, G. (Ed.). (2004). Handbook of Bioethics:: Taking Stock of the Field from a Philosophical Per¬spective (Vol. 78). Springer Science & Business Media. doi: 10.1007/1-4020-2127-5

Kidder, R.M. (1995). How Good People Make Tough Choices. New York: Morrow.

Kiran, A.H., Oudshoorn, N., & Verbeek, P.P. (2015) Beyond checklists: toward an ethical-constructive technology assessment. Journal of Responsible Innovation, 2:1, 5-19. doi: 10.1080/23299460.2014.992769

Koza J.R., Bennett F.H., Andre D., Keane M.A. (1996). Automated Design of Both the Topology and Sizing of Analog Electrical Circuits Using Genetic Programming. Dordrecht: Springer.

Lang, A. (2009). The limited capacity model of motivated mediated message processing. The SAGE handbook of media processes and effects, 193-204.

Leary, M.R. (2014). Introduction to Behavioral Research Methods. Harlow: Pearson Education Limited

Lewis, C. (2014, May 12). Robots Are Starting to Make Offshoring Less Attractive. Harvard Business Review. Retrieved from https://hbr.org/2014/05/robots-are-starting-to-make-offshoring-less-attractive

Linstone, H.A., & Turoff, M. (2002). The Delphi Method Techniques and Applications. Boston: Addi¬son-Wesley.

Lopatto, E. (2019, July 16). Elon Musk unveils Neuralink's plans for brain-reading 'threads' and a robot to insert them. The Verge. Retrieved from https://www.theverge.com/2019/7/16/20697123/elon-musk-neuralink-brain-reading-thread-robot

Lopez-Sanchez, M., Serramia, M., Rodriguez-Aguilar, A.J., Morales, J., & Wooldrige, M. (2017). Auto¬mated Decision Making to Help Establish Norm-Based Regulations. Paper presented at the Sixteenth In¬ternational Conference on Autonomous Agents and Multiagent Systems, Sao Paulo. Retrieved from http://www.ifaamas.org/Proceedings/aamas2017/pdfs/p1613.pdf

Maas, M.M. (2018). Regulating for 'Normal AI Accidents' Operational Lessons for the Responsible Governance of AI deployment. Retrieved from: http://www.aies-conference.com/2018/accepted-papers/

Marsh Jr, C. W. (2001). Public relations ethics: Contrasting models from the rhetorics of Plato, Aristot¬le, and Isocrates. Journal of Mass Media Ethics, 16(2-3), 78-98.

McCombs, M., & Reynolds, A. (2009). How the news shapes our civic agenda. In Media effects (pp. 17-32). Routledge.

McElreath, M. (1997). Managing strategic and ethical public relations campaigns (2nd ed.). Dubuque, IA: Brown & Benchmark.

Miller, T. (2018) Explanation in Artificial Intelligence: Insights from the Social Sciences.arXiv:1706.07269v3

Moore, G. (1965). Moore's law. Electronics Magazine, 38(8), 114.

Morgan, M., Shanahan, J., & Signorielli, N. (2009). Growing up with television: Cultivation processes. Media effects: Advances in theory and research, 3, 34-49.

Msibi, P.N., Mogale, R., de Waal, M., & Ngcobo, N. (2018). Using e-Delphi to formulate and ap¬praise the guidelines for women's health concerns at a coal mine: A case study. Curationis, 41(1), 1-6. doi: 10.4102/curationis.v41i1.1934.

Müller, C.V., & Bostrom, N. (2016). Future Progress in Artificial Intelligence: A Survey of Expert Opin¬ion. In Fundamental Issues of Artificial Intelligence. Springer. pp. 553-571

Murphy, M.K., Black, N.A., Lamping, D.L., McKee, C.M., Sanderson, C.F.B., Askham, J., & Marteau, T. (1998). Consensus development methods, and their use in clinical guideline development. Health Tech¬nology Assessment, 2(3). Retrieved from https://pdfs.semanticscholar.org/ae93/f886bcc02a914fd1669cb¬ba1345ea0748121.pdf

Nelson, H.L. (2004) Four Narrative Approaches to Bioethics. In Khushf, G. (Eds.), Handbook of Bioeth¬ics: Taking Stock of The Field From a Philosophical Perspective. (pp. 163-181). Dordrecht: Springer.

Nelson, R. R., & Phelps, E. S. (1966). Investment in humans, technological diffusion, and economic growth. The American economic review, 56(1/2), 69-75.

Noothigattu, R., Gaikwad, S.N.S., Awad, E., Dsouza, S., Rahwan, L., Ravikumar, P., & Procaccia, A.D. (2018). A Voting-Based System for Ethical Decision Making. Paper presented at the Thirty-Second AAAI Con¬ference, New Orleans USA. Retrieved from https://arxiv.org/abs/1709.06692

Office of Science and Technology Policy, National Science and Technology Council Committee on Technology. (2016). Preparing for the Future of Artificial Intelligence. Retrieved from: https://obamawhite¬house.archives.gov/sites/de-fault/files/whitehouse_files/microsites/ostp/NSTC/prepar-ing_for_the_fu¬ture_of_ai.pdf.

OpenAI (n.d.). Artificial General Intelligence (AGI) will be the most significant technology ever created by humans. Retrieved from https://openai.com/about/#mission

Otterlo, M. (2018). From Algorithmic Black Boxes to Adaptive White Boxes: Declarative Decision-The¬oretic Ethical Programs as Codes of Ethics. Retrieved from: http://www.aies-conference.com/2018/accept¬ed-papers/

Paiva, E.L., Gavronski, I., & Castro D'Avila, L. (2011). The Relationship between Manufac¬turing Integration and Performance from an Activity-Oriented Perspective. Brazilian Administra¬tion Review, 8-4, pp. 376-394. Retrieved from http://www.scielo.br/scielo.php?script=sci_arttex-t&pid=S1807-76922011000400003&lng=en&nrm=iso&tlng=en

Palm, E., & Hansson, S. O. (2006). The case for ethical technology assessment (eTA). Technological forecasting and social change, 73(5), 543-558.

Partnership On AI (n.d.). Tenets. Retrieved from https://www.partnershiponai.org/tenets/

Pasquale, F. (2016, February 5). Bittersweet Mysteries of Machine Learning (A Provocation) The London School of Economics and Political Science. Retrieved from: http://blogs.lse.ac.uk/mediapolicypro¬ject/2016/02/05/bittersweet-mysteries-of-machine-learning-a-provocation/

Pega. (2018). What Consumers Really Think About AI: A Global Study. Retrieved from Pega https://www.pega.com/ai-survey?utm_source=emd&utm_medium=pr&utm_content=AI-Consumer-Study-Part-2

Pellegrino, E.D. (2004) Philosophy of Medicine and Medical Ethics: A Phenomenological Perspective. In Khushf, G. (Eds.), Handbook of Bioethics: Taking Stock Of The Field From A Philosophical Perspective. (pp. 183-205). Dordrecht: Springer.

Perrow, C. (1984). Normal Accidents: Living With High-Risk Technologies, Basic Books.
Pinkus, R. L. B., Shuman, L. J., Hummon, N. P., & Wolfe, H. (1997). Engineering ethics: Balancing cost, schedule, and risk, lessons learned from the space shuttle. New York: Cambridge University Press.

Plebe, A., & Perconti, P. (2012). The Slowdown Hypothesis. Singularity Hypotheses, 349–65. The Fron¬tiers Collection. Springer Berlin Heidelberg. doi: https://doi.org/10.1007/978-3-642-32560-1_17

Powell, C. (2003). The Delpi technique: myths and realities. Journal of Advanced Nursing, 41(4), 376-382. Retrieved from https://www.ncbi.nlm.nih.gov/pubmed/12581103

Public Relations Society of America. (n.d.). PRSA Code of Ethics. Retrieved from https://www.prsa.org/wp-content/uploads/2018/04/PRSACodeofEthics.pdf

Qualtrics, L. L. C. (2019). Qualtrics Survey Software [online computer software]. Retrieved from http://www.qualtrics.com/

Rescher, N. (1977). Dialectics: A Controversy-Oriented Approach to the Theory of Knowledge. Albany: State University of New York Press.

Rossi, F. (2016). Moral preferences. In The 10th Workshop on Advances in Preference Handling (MPREF). Available online at http://www.mpref-2016.preflib.org/wpcontent/uploads/2016/06/paper-15.pdf.

Rogers, E.M. (2003). Diffusion of innovations (5th ed.). New York: Free Press.

Rowe E. (1994). Enhancing Judgement and Decision Making: a critical and empirical investigation of the E-Delphi technique. Unpublished PhD Thesis, University of Western England, Bristol

Russell, S.J., & Norvig, P., (2010). Artificial Intelligence: A Modern Approach. Prentice Hall, third ed. Englewood Cliffs, New Jersey.

Shoemaker, S. J., Wolf, M. S., & Brach, C. (2014). Development of the Patient Education Materials Assessment Tool (PEMAT): a new measure of understandability and actionability for print and audiovisual patient information. Patient education and counseling, 96(3), 395-403.

Stocker, M. (1987). The Schizophrenia of Modern Ethical Theories,[w:] The Virtues. Contemporary Essays on Moral Character, ed. R. Kruschwitz, R. Roberts.

Sun, J., Yan, J., & Zhang, K.Z.K. (2016). Blockchain-based sharing services: What blockchain technology can contribute to smart cities. Financial Innovation, 2(26).doi: https://doi.org/10.1186/s40854-016-0040-y

Schmelzer, R. (2018, may). Combination of blockchain and AI makes models more transparent. Search Enterprise AI. Retrieved from https://searchenterpriseai.techtarget.com/feature/Combination-of-block¬chain-and-AI-makes-models-more-transparent

Seabrook, J. (2019, October 14). The Next Word: Where will predictive text take us? The New Yorker. Retrieved from https://www.newyorker.com/magazine/2019/10/14/can-a-machine-learn-to-write-for-the-new-yorker

Semlitsch, T., Blank, W. A., Kopp, I. B., Siering, U., & Siebenhofer, A. (2015). Evaluating guidelines: a review of key quality criteria. Deutsches Ärzteblatt International, 112(27-28), 471.

Swink, M., & Song, M. (2007). Effects of marketing-manufacturing integration on new product devel¬opment time and competitive advantage. Journal of Operations Management, 25, 203-217. doi:10.1016/j. jom.2006.03.001

Tene, O., & Polonetsky, J. (2014). A Theory of Creepy: Technology, Privacy and Shifting Social Norms. Yale Journal of Law and Technology, 16, 1, 59-100. Retrieved from: https://digitalcommons.law.yale.edu/ cgi/ viewcontent.cgi?referer=https://www.google.nl/&httpsredir=1&article=1098&context=yjolt

Thaler, R., & Sunstein, C. (2008) Nudge: Improving Decisions About Health, Wealth, and Happiness. New Haven: Yale University Press.

Thangaratinam, S., & Redman, C.W.E. (2005). The Delphi Technique. The Obstetrician & Gynaecolo¬gists, 7(2), 120-5. doi: 10.1576/toag.7.2.120.27071

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, Version 2. IEEE, 2017. http://standards.ieee.org/develop/indconn/ec/autonomous_systems.html.

The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems, First Edition. IEEE, 2019. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html.

Thomasma, D.C. (2004). Virtue Theory in Philosophy of Medicine. In Khushf, G. (Eds.), Handbook of Bioethics: Taking Stock Of The Field From A Philosophical Perspective. (pp. 89-121). Dordrecht: Springer.

Tilley, E. (2005). The ethics pyramid: Making ethics unavoidable in the public relations process. Journal of Mass Media Ethics, 20(4), 305-320.

Tong. R. (2004). Feminist Approaches To Bioethics. In Khushf, G. (Eds.), Handbook of Bioethics: Taking Stock Of The Field From A Philosophical Perspective. (pp. 143-163). Dordrecht: Springer.

Tranfield D, Denyer D and Smart P (2003) Towards a methodology for developing evidence informed management knowledge by means of systematic review. British Journal of Management 14: 3207–3222.

United Nations Institute for Disarmament Research. (2016) Safety, Unintentional Risk and Accidents in the Weaponization of Increasingly Autonomous Technologies. Retrieved from: http://www.unidir.org/files/publications/pdfs/safety-unintentional-risk-and-accidents-en-668.pdf

Urban, T. (2015a, January 22). The AI Revolution: The Road to Superintelligence. Wait But Why. Re¬trieved from https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-1.html

Urban, T. (2015b, January 27). The AI Revolution: Our Immortality or Extinction. Wait But Why. Retrieved from https://waitbutwhy.com/2015/01/artificial-intelligence-revolution-2.html

US Senate Subcommittee on Space, Science and Competitiveness. (2016). The Dawn of Artifi-cial Intelligence. Retrieved from: http://www.commerce.senate.gov/public/index.cfm/2016/11/com¬merce-announces-first-artificial-intelligence-hearing

Vardi, M.Y. (2012). Artificial Intelligence: Past and Future. Magazine Communications of the ACM, 55(1), 5-5.

Veatch, R. M. (1985). Against virtue: a deontological critique of virtue theory in medical ethics. In Vir¬tue and medicine (pp. 329-345). Springer, Dordrecht. doi: https://doi.org/10.1007/978-94-009-5229-4_17

Verbeek, P.P. (2013). Technology Design as Experimental Ethics. In S. van den Burg, & T. Swierstra, Ethics on the Laboratory Floor (pp. 83-100). Retrieved from http://www.ppverbeek.nl/

Verbeek, P.P. (2017). The Struggle for Technology: Towards a Realistic Political Theory of Technology. Foundations of Science, 22, (2), 301-304. https://doi.org/10.1007/s10699-015-9470-7

Walker, M.U. (2007). Moral Understandings. Routledge, New York.

Watzlawick, P., Bavelas, J.B., & Jackson, D.D. (1967). Pragmatics of human communication: A study of interactional patterns, pathologies, and paradoxes. New York: Norton.

Wiggers, K. (2019, June 10). Amazon launches Personalize, a fully managed AI-powered recommenda¬tion service. Venturebeat. Retrieved from https://venturebeat.com/2019/06/10/amazon-launches-person¬alize-a-fully-managed-ai-powered-recommendation-service/

Williams, B., & Bernard, W. (1981). Moral luck: philosophical papers 1973-1980. Cambridge University Press. doi: https://doi.org/10.1017/CBO9781139165860

Appendices

Appendix A. Summaries Development Process

Summary Round 1:

Summary DISC-guideline development process

The marketing communication of AI-based services should be done more ethically cognizant. While ethical guidelines for the development and implementation of AI-based services are currently heavily researched, the marketing communication of these services are often overlooked. Yet, marketing communication can be very influential in technology adoption and constituting our perceived reality of these products. Moreover, as pointed out by the IEEE, AI-based services are currently often inaccurately marketed. Marketing communication practitioners are in need of marketing communication guidelines for AI-based services. As such, the guidelines could help consumers get more involved with the technology and to know for themselves how to be more protected for the potential misuse of AI and when an AI product should be trusted. Additionally, guidelines could make the marketing communication of AI more responsible and therefore help constitute a responsible global epistemic perspective towards AI and help constitute an AI for good. This document briefly describes the development of four critical guidelines for the marketing communication of AI-based services, denoted as the DISC-guidelines. To review this development process in more depth, or to check the literary background, please consult the full theoretical framework.

To develop these guidelines, I have asked myself the following central question:

"What are prima facie guidelines for ethical marketing communication of AI-based services?"

In order to answer this question three critical points of knowledge were identified. First, one should know what are ethical principles in marketing communication. Second, ethical principles in technology development should be considered. Finally, one ought to know what is ethical when communicating AI-based services. Subsequently, a moral framework was developed harnessed to these three critical points of knowledge, the framework was further described by utilizing principlism as a synthesizing method (see section 2.1 in full theoretical framework). Principlism is a system of ethics based on four prima facie principles:

- The principle of autonomy (supporting and respect for autonomy);
- The principle of beneficence (work towards beneficence);
- The principle of nonmaleficence (averting harm);
- The principle of justice (democratically distribute benefits and risks).

Principlism can be used to develop guidelines or principles by specifying, balancing and justifying these four principles for a specific case of interest. Accordingly, the relevant principles or ethical considerations from the critical knowledge points were reflected upon, where after the critical principles considering the knowledge points could be specified. These principles were justified for their coherence with the four principles of principlism. Since no conflicting principles were found, balancing was not needed.

These specified principles are (see section 2.2 in the full theoretical framework to review this process in more depth):

-Principles of marketing communication (section 2.2 in full theoretical framework)

Baker and Martinson's (2001) TARES test consist of five principles for ethical persuasive communication:

- Truthfulness – of the message (honesty, trustworthiness, non-deceptiveness);
- Authenticity – of the persuader (genuineness, integrity, ethical character, appropriate loyalty);
- Respect – for the persuade (regard for dignity, rights, well-being);
- Equity – of the content and execution of the appeal (fairness, justice, nonexploitation of vulnerability);
- Social responsibility – for the common good (concern for the broad public interest and welfare more than simply selfish self-interest).

-Principles of technology development (section 2.2 in full theoretical framework)

Kiran, Oudshoorn and Verbeek (2015) developed an Ethical-Constructive Technology Assessment (eCTA) framework to guide designers/developers through their design process:

- The embodiment principle – considering technologies locus, form and domain;
- The hermeneutic principle – considering how technologies will be interpreted;
- The alterity principle – considering technologies physical mediations;
- The background relation principle – considering technologies constitutive mediations.

- Principles of communicating AI-based services (section 2.2 in full theoretical framework)

These principles were induced from AI-based services' their innerworkings and deduced from scrutinizing AI-based companies and organizations their principles:

- The principle of transparency – of the technology's workings, development and implementation.
- The principle of democratization – of the technology's amenability and accessibility.

The found principles were systematically coupled in a table (see table 2 in section 2.5 of the full theoretical framework) in order to systematically formulate critical guidelines. Twelve guidelines were formulated independent of each other, in order to view a set of principles collectively these critical guidelines were reasoned aggregated. As such, the four critical DISC-guidelines emerged:

1. Democratic and transparent amenability:

The locus, form and framework of the activity should be democratically amenable, while being insusceptible to exploitation of vulnerability. Moreover, the activity should be accompanied in its design to be constituted as such.

2. Interpretability:

The activity should be interpretable for the AI service' transparency, integrity and democratic amenability and accompanied in its design to be constituted as such.

3. Social effects:

The activity should be interpretable for how the AI service' informs its democratic amenability and integrity from individual -and social perceptions.

4. Constitutive effects:

The activity should be interpretable on to what extent the AI service can constitute its democratic amenability and integrity in daily life.

Summary Round 2:

Summary literature review and guideline development process

The marketing communication of AI-based services should be done more ethically cognizant. While ethical guidelines for the development and implementation of AI-based services are currently heavily researched, the marketing communication of these services are often overlooked. Yet, marketing communication can be very influential in technology adoption and constituting our perceived reality of these products. Moreover, as pointed out by the IEEE, AI-based services are currently often inaccurately marketed. Marketing communication practitioners are in need of marketing communication guidelines for AI-based services. As such, the guidelines could help consumers get more involved with the technology and to know for themselves how to be more protected for the potential misuse of AI and when an AI product should be trusted. Additionally, guidelines could make the marketing communication of AI more responsible and therefore help constitute a responsible global epistemic perspective towards AI and help constitute an AI for good. This document briefly describes the development of four critical guidelines for the marketing communication of AI-based services. To review this development process in more depth, or to check the literary background, please consult the full literature review.

To develop these guidelines, I have asked myself the following central question:

"What are prima facie guidelines for ethical marketing communication of AI-based services?"

In order to answer this question three critical points of knowledge were identified. First, one should know what are ethical principles in marketing communication. Second, ethical principles in technology development should be considered. Finally, one ought to know what is ethical when communicating AI-based services. Subsequently, a moral framework was developed harnessed to these three critical points of knowledge, the framework was further described by utilizing principlism as a synthesizing method (see section 2.2.5 in full literature review). Principlism is a system of ethics based on four prima facie principles:

- The principle of autonomy (supporting and respect for autonomy);
- The principle of beneficence (work towards beneficence);
- The principle of nonmaleficence (averting harm);
- The principle of justice (democratically distribute benefits and risks).

Principlism can be used to develop guidelines or principles by specifying, balancing and justifying these four principles for a specific case of interest. Accordingly, the relevant principles or ethical considerations from the critical knowledge points were reflected upon, where after the critical principles considering the knowledge points could be specified. These principles were balanced in order to resolve any conflicts between the principles. Finally, the principles were justified for their coherence with the four principles of principlism.
These specified principles are (see section 2.2 in the full literature review to review this process in more depth):

-Principles of marketing communication (section 2.2 in full theoretical framework)
Baker and Martinson's (2001) TARES test consist of five principles for ethical persuasive communication:

- Truthfulness – of the message (honesty, trustworthiness, non-deceptiveness);
- Authenticity – of the persuader (genuineness, integrity, ethical character, appropriate loyalty);
- Respect – for the persuade (regard for dignity, rights, well-being);
- Equity – of the content and execution of the appeal (fairness, justice, nonexploitation of vulnerability);
- Social responsibility – for the common good (concern for the broad public interest and welfare more than simply selfish self-interest).

-Principles of technology development (section 2.2 in full theoretical framework)
Kiran, Oudshoorn and Verbeek (2015) developed an Ethical-Constructive Technology Assessment (eCTA) framework to guide designers/developers through their design process:

- The embodiment principle – considering technologies locus, form and domain;
- The hermeneutic principle – considering how technologies will be interpreted;
- The alterity principle – considering technologies physical mediations;
- The background relation principle – considering technologies constitutive mediations.

-Principles of communicating AI-based services (section 2.2 in full theoretical framework)
These principles were induced from AI-based services' their innerworkings and deduced from scrutinizing AI-based companies and organizations their principles:

- The principle of transparency – of the technology's workings, development and implementation.
- The principle of democratization – of the technology's amenability and accessibility.

The found principles were systematically coupled in a table (see table 2 in section 2.6 of the full literature review) in order to systematically formulate critical guidelines. Twelve guidelines were formulated independent of each other, in order to view a set of principles collectively these critical guidelines were reasoned aggregated. From the first e-Delhpi round, the guidelines were evaluated as too vague. As such, the guidelines were reformulated accordingly:

1. Fair and understandable distribution of the marketing activity its design, presentation and framing (equal accessibility for all).
2. Truthful distribution of the marketing activity its information on the AI based services its transparency, integrity and extent of equal accessibility (sincere transparency).
3. Informative of social effects the marketing activity should be informative for how the AI based service is equally accessible and honest for individuals, and within social situations (social responsibility).
4. Informative of contextual effects the marketing activity should be equally accessible and transparent on how the AI based service will shape our daily lives (contextual transparency).
5. Promote dialogue between consumers on the marketing activity its information, when there is a danger of exclusion of vulnerable consumers due to complex information.
6. Accompaniment of the above ethical values in the design of the marketing activity, without taking advantage of vulnerabilities.

Context of use, facilitators & barriers
The six guidelines should not be used as a checklist, but rather as handhelds for the thought process when communication marketing practitioners are marketing AI based services. The fundamental purpose of these guidelines is to ethically align the development of marketing activities regarding AI based services. These

marketing activities could include communicative messages ranging from posters, advertorials or billboards to television commercials or even complete campaigns. The activities should be developed alongside the development of the AI based service in order to maximize adequate alignment of ethical values without diminishment of organizational performance. Accordingly, the fundamental purpose of these guidelines is to ensure equal distribution of information on AI based services so that the most people have a fair change of participation in the adoption of this technology. As such, the guidelines help minimize the disruption between social economic structures through AI. Also view the white paper (646 words) for a description of how to apply these guidelines.

The guidelines will be freely available for everyone to download from a website yet to be specified. No additional resources are needed to use and/or implement the guidelines. The guidelines were developed without funding and every participating developer contributed voluntarily. Finally, the researcher conducted the research as part of his master thesis.

Appendix B. Whitepaper Guidelines for marketing AI-based services

Whitepaper Guidelines for marketing AI-based services

The proliferation of AI-based services in the business to consumer markets have increasingly ethical implications. Virtual assistants such as Amazons' Alexa, Google Home or Samsungs' Bixby indicated consumers needed to be weary of how their data is collected and stored. Autonomous vehicles such as Tesla made us question if we should trust AI systems with our lives and who actually is responsible when accidents do happen? The future promises even more remarkable products, such as AI chips to enhance our brains. What will the implications of those products be? While companies and organizations are developing guidelines for ethically aligning the development of AI-based services, guidelines for the marketing of these services are lagging behind. While this is a fundamental process of implementing and adopting new products. As such, we have developed guidelines for the marketing communication of AI-based services. This whitepaper declares what they are and how marketeers should use them.

Context of use

Within these guidelines, AI-based services entail business to consumer products and services. The guidelines guide the development of marketing activities and should be used concurrently with the development process of the AI-based service in question. As such, the alignment of ethical values in the marketing activity is maximized while maintaining an efficient workflow.

How to use

The guidelines below should be viewed as general cascading 'handhelds' determining the thought process of the marketeer and should not be used as a checklist. Instead, use them as a starting point for what to think about when designing your marketing activity or campaign. In general, the guidelines emphasize equally accessible and understandable marketing activities while being transparent about the ethical implications of the AI-based service.

When designing the marketing activity, consider whether all the design facets are understandable and accessibly for everyone.

> -Fair and understandable distribution of the marketing activity its design, presentation and framing (equal accessibility for all).

When developing the content for the marketing activity, consider whether the ethical implications of the AI-based service are truthfully included.

> -Truthful distribution of the marketing activity its information on the AI-based services its transparency, integrity and extent of equal accessibility (sincere transparency).

When developing the content for the marketing activity, consider whether information is included for how the AI-based service is accessible for individuals and for social situations. Consider for instance how everyone is able to address the virtual assistant Alexa and how the assistant is not only confined to its owner.

> -Informative of social effects. The marketing activity should be informative for how the AI-based service is equally accessible and honest for individuals and within social situations (social responsibility).

When developing the content for the marketing activity, consider whether information is included for how

the AI-based service could be shaping our lives. For instance, consider how humans became dependent on smartphones.

-Informative of contextual effects. The marketing activity should be equally accessible and transparent on how the AI-based service will shape our daily lives (contextual transparency).

In situations where complex content not accessible to everyone cannot be avoided. Consider how the target group who can understand the marketing activity, can be motivated to start a dialogue with 'non-competent' consumers on the marketing activity. Consider for instance elements of corroborating awareness campaigns for sexual harassment, suicidal people or smokers.

-Promote dialogue between consumers on the marketing activity its information, when there is a danger of exclusion of vulnerable consumers due to complex information.

Finally, consider how you can align the discussed guidelines within the design of your marketing activity and whether your marketing activity is also beneficent for vulnerable consumers (i.e., blind, handicapped or mentally ill people).

-Accompaniment of the above ethical values in the design of the marketing activity, without taking advantage of vulnerabilities.