

We are IntechOpen, the world's leading publisher of Open Access books Built by scientists, for scientists

6,900

Open access books available

186,000

International authors and editors

200M

Downloads

Our authors are among the

154

Countries delivered to

TOP 1%

most cited scientists

12.2%

Contributors from top 500 universities



WEB OF SCIENCE™

Selection of our books indexed in the Book Citation Index
in Web of Science™ Core Collection (BKCI)

Interested in publishing with us?
Contact book.department@intechopen.com

Numbers displayed above are based on latest data collected.
For more information visit www.intechopen.com



Virtual Assistants and Ethical Implications

Abhishek Kaul

Abstract

Virtual assistants are becoming a part of our daily life, from our homes to our work. Sometime we may not even know that the customer service agent that we were speaking with is a virtual assistant. These assistants continuously collect information from our interactions and learn many things about us. The information they gather over time is enormous. This chapter introduces the concept of ethics, discusses the ethical principles of virtual assistants, (Transparency, Justice & fairness, Non-maleficence, Responsibility and Privacy). Although there is limited regulation governing these virtual assistants, practical guidelines and recommendations are provided for designers and developers to understand the ethical implications when building a virtual assistant. In this chapter, we also discuss technology and learning techniques for virtual assistants and present examples on how to ensure ethical virtual assistant.

Keywords: virtual assistants, artificial intelligence, ethics, deep learning algorithms, natural language processing, natural language understanding, fair AI, transparent AI

1. Introduction

Organizations are rapidly deploying Virtual Assistants aka bot technology [1] for automating communication, customer service, conversational commerce, product recommendation, education support, financial services, medical services, entertainment, social outreach and self-service tasks. They offer 24/7 service and fulfill the need of millennials [2] for real time responses. Virtual assistants enable organizations to reduce costs, increase brand loyalty and better serve customers. However, virtual assistants are built by humans using artificial intelligence (AI) technologies and have wide ranging ethical implications which are important for organizations and consumers to understand.

Many countries have published AI policy guidelines [3, 4]. These guidelines provide a broad level objective for the use of AI – to ensure human-centric, safe, and trustworthy AI. One of the most important aspect in all guidelines is ethics, “AI should be ethical, ensuring adherence to ethical principles and values”. Although, AI by its very nature is a form of statistical discrimination (finding patterns in data), the discrimination becomes objectional when it places certain privileged groups at a systematic advantage and certain unprivileged groups at a systematic disadvantage. For example, the loan application algorithm gives higher credit scores to older males due to training bias. Objectional discriminations can arise due to multiple reasons like wrongly defining the business objective [5] of machine learning model, using unrepresentative data or data with existing prejudice [6] for training or by selecting wrong attributes or features of the AI model.

Significant work has been done in the area of Ethically aligned design for Autonomous and Intelligent systems by IEEE [7]; and in the area of Facial recognition technologies [8]; but the area of virtual assistants has seen limited guidelines or regulation. California “Bot Bill [9]” provides only limited protection for consumers in terms of bot self-declaration.

In subsequent sections of this chapter, we first define “What is ethics?” and then discuss on ethical principles for virtual assistants. These principles provide key ethical considerations that designers and consumers should understand. Next we discuss different types of virtual assistants deployed today deep, dive into the technology and learning techniques that make them ethical. Further, we analyze the guidelines and legislations that companies and governments have published. In the last section, we look into what is the probable future of super intelligent virtual assistants.

2. Defining ethics

What is ethics?

As per the Oxford dictionary ethics means “the moral principles that govern a person’s behaviour or the conducting of an activity”.

Ethics is based on well-founded standards of right and wrong that prescribe what humans ought to do, usually in terms of rights, obligations, benefits to society, fairness, or specific virtues [10].

An ethical virtual assistant should be designed with the ethical standards of the society it affects. These standards should extend to virtual assistants and their creators who should design, build and maintain virtual assistants to ensure that their interactions with consumers foster honesty, loyalty, refrain from doing harm, or fraud and provide the right to privacy.

In the subsequent section, we discuss in detail the ethical principles of virtual assistants.

3. Ethical principles for virtual assistants

AI based virtual assistants ability to act intelligently has long been evaluated by the Turing Test [11] and Loebner Prize [12]. The focus is on intelligence of the system to respond to human questions. Looking through the lens of ethical principles, other questions arise, beyond “What can the virtual assistant answer?” For example, “Does the answer promote consumers interests or business interests like recommending the most profitable product which does not best suited”.

In a recent paper published by Jobin, A., Ienca, M. & Vayena, E. on the global landscape of AI ethics guidelines [13], they found globally five emerging ethical principles that are deemed important – Transparency, Justice and fairness, Non-maleficence, Responsibly and Privacy. In this section we interpret these principles with a view on virtual assistants and what considerations should designers, developers and consumers understand when developing and interacting with virtual assistants.

3.1 Transparency

AI transparency refers to the explainability [14], interpretability, disclosure [15] of the algorithmic models including their training data, accuracy, performance, bias and other metrics.

When dealing with virtual assistants, transparency [16] often refers to informing the consumers who they are chatting with ie virtual assistant, not actual human, sharing details on what information can the consumer search, and how his data will be used, stored, analyzed for improving experience.

Brands build trust with consumers by being transparent and honest in communication. Virtual assistants are an extension of brands consumer experience. If virtual assistants impersonate a human, it can lead to poor experience and lack of trust with the brand. This can also be harmful when interacting with consumers on sensitive areas like healthcare or banking.

Designer and developers of virtual assistants should be transparent in disclosing information to consumers in terms of what they can search and disclose how their data would be shared and analyzed. When the consumers know what they can search they will be able to ask questions on the topics that virtual assistant has been trained on and get desirable answers. This will create a delightful experience. Further, consumer should have the choice to opt in their interaction data for other purposes like development of the AI models or for advertisements and more. This will help to gain consumer confidence in virtual assistants and increase adoption. Lastly, consumers should also have the option to connect to a real person, request callback or send an email if they are uncomfortable in interacting with a virtual assistant.

3.2 Justice, fairness and equity

Justice means that AI algorithms are fair and do not discriminate against particular groups intentionally or unintentionally [17]. There have been numerous publications on fairness and how to identify, mitigate bias in Algorithms [18–20]. In case of virtual assistants justice, fairness and equity refer primarily to prioritizing the consumer interests and providing impartial recommendations [21].

AI models on recommendation generally use techniques of collaborative filtering, ie filtering for consumer preferences based on information gathered from many similar consumers. The models constantly learn from consumer feedback ie likes or dislikes and adjust accordingly.

However these models can be biased based on the consumer training data or based on overarching business rules like recommend the most profitable product. For example, will the virtual assistant recommend the meat which is most expensive and near expiry date or the meat which is cheaper and fresh?

Virtual assistants being viewed favorable towards certain recommendations raises the question on fairness especially for consumers. When virtual assistants are used within an organization, then sometimes recommendation may rule driven, which is as per the employee policy.

Designers and developers should regularly test the virtual assistants against the fairness metrics, publish them to consumers and also give consumers the option to provide feedback on recommendation. The more virtual assistant adapts to consumers interest and provides fair recommendation, the more popular the virtual assistant will become with consumers.

3.3 Non-maleficence

This term is used to define consumers safety, security and the commitment that AI model will not cause harm for example, by spamming, hacking, discrimination, violation of privacy or abuse.

In case of virtual assistants, we focus on abuse and sexual harassment for this principle. Abuse refers to both receiving abuse from consumers and giving back abuse to consumers.

Many times, virtual assistants are at the beginning of a consumers journey, and if the responses are not helpful it leads to frustration and abuse from consumers. Although, virtual assistants are AI models and do not have feelings (like humans), as consumers, we should refrain from abusing since it impacts the way we behave in society and transcends similar behavior towards even our fellow humans.

Designers and developers need to design the conversation experience with consideration that virtual assistants will receive abuse. They should design the conversation flow empathically so that the consumers are provided a positive response and transferred to a more helpful channel like voice or email on request [22].

Another consideration is gender stereo-typing ie the gender of virtual assistant. In many cases, virtual assistants have a default female voice or persona. Designers and developers can provide options to consumers to select the virtual assistant persona and alter language, voice, tone of responses specific to chosen persona.

In a related study on sexual harassment of virtual assistants, “#MeToo: How Conversational Systems Respond to Sexual Harassment [23]” points different behaviors in commercial, supervised and unsupervised learning based virtual assistants. The unsupervised learning based assistants have more freedom in learning from user conversation and responding similarly. In these cases language correction models should also be deployed to protect users from chatbot abuse. For example, Microsoft’s Tay chatbot was corrupted in less than 24 hours by self-learning through user conversation [24].

3.4 Responsibility and accountability

Responsibility and accountability refer to the AI acting with integrity, clarifying the attribution of responsibility and data ownership. In case of virtual assistants this refers to being transparent, fair, disclosing information on responsibility, legal liability and data ownership to consumers.

There has been much debate on who is ultimately responsible – is it the AI based virtual assistants or the humans who built it. Generally, terms of service agreement which consumers have to agree before using virtual assistants, define the limitations on responsibilities and liabilities in line with regulations.

Data ownership requires special mention here. Questions typically arise on who owns the data when it is captured and generated during conversation with virtual assistant. For example, new data is generated when a virtual assistant interacts with consumers using voice. It will over time develop data related to consumers preferences (preference in music), personality [25] (words and tone of language), family (number of different voices in family or type of requests made ~ nursery rhymes) and more. Sometimes, organization may have built the business model on leveraging this derived data for profit. For example, Virtual assistant derives data on the age of your children and serves you advertisements on children toothbrush.

Designer and developers should be transparent on data ownership and have an opt in feature, if the consumers want to share this new data generated or want to keep it private. If the business model of the virtual assistant is based on offering free services and leverage consumer data for advertisements, then that should also be transparent to the consumer.

3.5 Privacy

Privacy means that your personal information is kept confidential and only shared with consent. Many countries have passed laws and regulations to protect the privacy of their citizens like General Data Protection Regulation [26]. In relation to Virtual assistants, privacy is often referred in relation to data protection and security.

Deeper questions on privacy for Virtual assistants arise from

- who has access to the conversation transcripts?
- are the transcripts being used to profile the consumers?
- are the transcripts being shared with advertisers?
- are the consumer details anonymized before sharing?
- are the transcripts being used for improving the AI model?
- where are the transcripts stored?
- for how much time are the transcripts stored?
- can the consumer delete the transcripts?
- is the communication channel encrypted?

And so on.

Designer and developers should be transparent on privacy policy and publish it online so that consumers can be informed on how their information is stored and protected. This will also help to develop trust in the virtual assistant and consumers will be more willing to share information if they know that they will be served better.

4. Learning techniques of virtual assistants and ethical considerations

In self-service technology, virtual assistants are on the higher maturity curve and are expected to understand and interact with consumers as “humans” to provide information or take action. If we look under the hood of virtual assistants, then we uncover three basic technology building blocks.

1. Channel of communication – Physical device (Amazon Echo, iPhone Siri), Messaging Platform (Slack, Facebook messenger), Website or App. The channel of communication generally includes voice interaction capability if available.
2. Conversational platform – Brain of the virtual assistant which has the rules and AI technology to understand consumer information and context.
3. Backend Database or Automation/APIs – This is the backend system from where information is retrieved or a specific task is executed. For example calling an API to retrieve weather information for location or setting up an alarm.

In this section, focus will be on the conversational platform which has to be designed with ethical consideration. There are many types of technologies deployed for virtual assistants ranging from simple click based predefined options, to pattern matching, natural language understanding and natural language generation. In the section below, different types conversational platforms are discussed with a view on ethical considerations.

4.1 Commercial virtual assistant platforms

Most commercial virtual assistants use pattern matching and natural language understanding AI models. The primary task of the AI model in this case is to classify intent of the question for pre-defined set of answers. The assistants can also understand specific details in the text like country name or time and more. For example if asked “What is the weather in Singapore?” assistant will classify this as the request to find weather information and also extract Singapore country name. This information will be passed to backend API to retrieve the temperature and presented back as the answer. Example of these virtual assistants used by business are IBM Watson Assistant, Microsoft Bot framework, Amazon Lex, Google Dialog flow and more. Learning on these platform is generally supervised and the knowledge corpus is limited to the business use case. Sometimes, extension of these platforms is done where a large document corpus is ingested and most relevant document is brought forward to the user based on search and retrieval techniques.

In these platforms, it is the role of conversation designer and developer to ensure that the virtual assistant adheres with the ethical principles of Transparency, Justice & fairness, Non-maleficence, Responsibly and Privacy. Further, it is a good practice that document corpus is screened before being ingested into these virtual assistants to ensure relevant and proper responses.

4.2 Mass market virtual assistants

Siri, Alexa and Hey Google are examples of mass market, virtual assistants. These virtual assistants are pre-trained from a large language corpus and have the ability to retrieve personal information from calendar, phonebook, music, credit card and more. The organization developing these Virtual assistants publish their terms of service, privacy policy [27] publicly and it is consumers decision to understand and then interact with them.

The ubiquitous nature of these Virtual assistants poses a bigger question to society on how they should respond to different types of talk ranging from Rude talk, Abusive talk, Romantic talk or Suicidal talk. We discuss below two cases in detail, rude and romantic talk.

Rude Talk – the virtual assistants tend to respond back positively with information without prompts for polite or rude requests. This has an influence on manners especially in case of younger consumers [28]. For example “Alexa can you please tell me the weather forecast for today” or “Alexa weather forecast today” – the answer would be the same. These assistants should try to add nicer words like “Thank you” when consumers say “please”.

Romantic Talk and Gender – when asked on gender, the virtual assistants tend to respond on gender neutrality. However, by default they respond in a female voice. In the article by Jessi Hempel [29] in Wired she explains that people tend to perceive female voices as helping us to solve our problems. This also opens the door to romantic talk [30] for female persona based virtual assistants. Most of assistant are trained to handle this conversations by evading, or positively responding to consumers, but they rarely respond negatively [31]. This does extend in some cases to general acceptance of sexual harassment for assistants.

4.3 Niche virtual assistants [open domain]

A special mention here to Virtual assistants who can talk about anything in the open domain. These assistants are trained using sophisticated deep learning AI models (un-supervised learning), have billions of parameters and are closest to how

a human would sensibly and specifically answer questions. Many gigabytes of training data (dialog response) is ingested in these AI models and it generates the answers (natural language generation) based on learning. Example of these assistants are

- Meena [32] - trained on 341 GB of text, filtered from public domain social media conversations.
- DialoGPT [33] - large pre-trained **dialog** response model trained on 147 M multi-turn **dialogs** extracted from Reddit discussion threads.
- Mitsuku [34] - although this virtual assistant uses pattern matching technique, it has won many competitions.
- Cleverbot [35] - this searches through its saved conversations, and responds to the input by finding how a human responded to that input when it was asked, in part or in full

Other than Mitsuku which uses supervised learning, for other virtual assistants, it is difficult predict responses since they are learning from dialog-response corpus. In these cases, it would be beneficial to have a language filter that checks for ethical considerations like abuse words and more before presenting the answers to consumers.

5. Guidelines and legislations

Many countries have published AI policy guidelines. These guidelines provide a broad level objective for the use of AI – to ensure human-centric, safe, and trustworthy AI. Most guidelines make the organization using AI responsible and accountable for their decisions and ask for the same ethical standards in AI-driven decisions as in human-driven decisions.

General key points achieved from global guidelines are:

- it should be lawful, complying with all applicable laws and regulations;
- it should be ethical, ensuring adherence to ethical principles and values; and
- it should be robust, both from a technical and social perspective since, even with good intentions, AI systems can cause unintentional harm.

Specially for virtual assistants, as defined above, five emerging ethical principles that are deemed important – Transparency, Justice & fairness, Non-maleficence, Responsibly and Privacy.

Legislation has been passed in California [36] to ensure Transparency of Virtual assistants. This law makes it mandatory for Virtual assistants (Bots) to disclose that they are not a real person and are virtual. Many other countries are passing laws and issuing guidelines to make it mandatory for Designers and developers to develop ethical Virtual assistants.

Many commercial organization have also issued their ethical guidelines. IBM has established an Ethics Board [37] which provides governance, review and decision making processes for IBM on ethics policies, practices, communication, research, products and services. They have also published open source toolkits which designers and developers can use to test whether there machine learning, AI models are transparent, fair and explainable.

Google Deepmind [38] has established a focus group which focuses on ethical standards and safety. They look at it from the lens of Privacy, transparency and fairness; AI morality and values; Governance and accountability; AI & world's complex challenges, Misuse and unintended consequences; Economic impact and inclusion.

Microsoft [39] has issued guidelines for responsible bots, which are aimed at helping designers and developers to design a bot that builds trust in the company and service that the bot represents.

Many other companies have issued guidelines to ensure that the Virtual assistants developed on their platform maintain high ethical standards [40] like use supervised learning, divert issues on serious topics, do not spam users, keep user privacy, no advertisements and so on.

6. What comes next

Virtual Assistant AI technology is growing at exponential pace. In the next few years we will have virtual assistants that surpass an average human's ability to respond sensibly and specially to a consumer's question. Nick Bostrom [41] presents an interesting perspective on super intelligent moral thinking. In the distant future, as AI capabilities surpass human intelligence, it could do better than human thinkers and have the correct answers on ethics by weighing up evidence. We have already started seeing the initial versions of these intelligent machines.

IBM Debater [42] is an example of super intelligent system. This AI system can independently debate a human and provide persuasive arguments on complex topics. The system is able to listen and understand a long spontaneous speech, model human dilemmas to form an argument and generate a whole speech of an opinion and deliver it persuasively. The system has participated live and won many debate competitions.

Another example is from Soul Machines [43]. It provides Digital people i.e. animation of life like people on the screen. These screen animations of people is similar to an actual human who speak with expressions (eye, lips and facial movements). This provides a comfort feeling when interacting with virtual assistant.

As virtual assistants become a part of our daily life, ethical issues regarding virtual assistants will continue to grow. It is important for the society at large to discuss and agree on the ethical principles of Transparency, Justice & fairness, Non-maleficence, Responsibility and Privacy for virtual assistants.

Conflict of interest

The views expressed in this chapter are my own and are not representative of my employer.

Notes/thanks/other declarations

I thank Ali Soofastaei, who has been my mentor and guide for the initiative of publishing this chapter on Virtual Assistants and Ethical Considerations.

IntechOpen

IntechOpen

Author details

Abhishek Kaul
IBM, Singapore

*Address all correspondence to: abhishekkaul@gmail.com

IntechOpen

© 2021 The Author(s). Licensee IntechOpen. This chapter is distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/3.0>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited. 

References

- [1] Building better bots with Watson Conversation. Available from: <https://www.ibm.com/blogs/watson/2016/07/building-better-bots-watson-conversation/> [Accessed: 2020-12-10]
- [2] Why Millennials Have Higher Expectations for Customer Experience Than Older Generations. Available from: <https://www.forbes.com/sites/nicolemartin1/2019/03/26/why-millennials-have-higher-expectations-for-customer-experience-than-older-generations/> [Accessed: 2020-12-10]
- [3] Ethics guidelines for trustworthy AI. Available from: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> [Accessed: 2020-12-10]
- [4] Guidance Data Ethics Framework. Available from <https://www.gov.uk/government/publications/data-ethics-framework> [Accessed: 2020-12-10]
- [5] The case for fairer algorithms. Available from: https://medium.com/@Ethics_Society/the-case-for-fairer-algorithms-c008a12126f8 [Accessed: 2020-12-10]
- [6] InCoding — In The Beginning Was The Coded Gaze. Available from: <https://medium.com/mit-media-lab/incoding-in-the-beginning-4e2a5c51a45d> [Accessed: 2020-12-10]
- [7] Ethics In Action | Ethically Aligned Design. Available from: <https://ethicsinaction.ieee.org/> [Accessed: 2020-12-10]
- [8] The ethical questions that haunt facial-recognition research. Available from: <https://www.nature.com/articles/d41586-020-03187-3> [Accessed: 2020-12-10]
- [9] An act to add Chapter 6 (commencing with Section 17940) to Part 3 of Division 7 of the Business and Professions Code, relating to bots. Available from: https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001 [Accessed: 2020-12-10]
- [10] What is Ethics? Available from: <https://www.scu.edu/ethics/ethics-resources/ethical-decision-making/what-is-ethics/> [Accessed: 2020-12-10]
- [11] What is Ethics? Available from: <https://plato.stanford.edu/entries/turing-test/> [Accessed: 2020-12-10]
- [12] What is Ethics? Available from: https://en.wikipedia.org/wiki/Loebner_Prize [Accessed: 2020-12-10]
- [13] Jobin, A., Ienca, M. & Vayena, E. The global landscape of AI ethics guidelines. Nat Mach Intell 1, 389-399 (2019). <https://doi.org/10.1038/s42256-019-0088-2>
- [14] AI Explainability 360 – Resources. Available from: <http://aix360-dev.mybluemix.net/resources#guidance> [Accessed: 2020-12-10]
- [15] Introduction to AI FactSheets. Available from: <https://aifs360.mybluemix.net/introduction> [Accessed: 2020-12-10]
- [16] 5 Chatbot Code Of Ethics Every Business Should Follow. Available from: <https://botcore.ai/blog/5-chatbot-code-of-ethics-every-business-should-follow/> [Accessed: 2020-12-10]
- [17] How to Fight Discrimination in AI. Available from: <https://hbr.org/2020/08/how-to-fight-discrimination-in-ai> [Accessed: 2020-12-10]
- [18] AI Fairness 360. Available from: <https://aif360.mybluemix.net/> [Accessed: 2020-12-10]

- [19] Fairlearn. Available from: <https://github.com/fairlearn/fairlearn> [Accessed: 2020-12-10]
- [20] Fairness Indicators. Available from: <https://github.com/tensorflow/fairness-indicators> [Accessed: 2020-12-10]
- [21] The code of ethics for AI and chatbots that every brand should follow. Available from: <https://www.ibm.com/blogs/watson/2017/10/the-code-of-ethics-for-ai-and-chatbots-that-every-brand-should-follow/> [Accessed: 2020-12-10]
- [22] Hard questions about bot ethics. Available from: <https://medium.com/slack-developer-blog/hard-questions-about-bot-ethics-4f80797e34f0> [Accessed: 2020-12-10]
- [23] #MeToo Alexa: How Conversational Systems Respond to Sexual Harassment. Available from: <https://www.aclweb.org/anthology/W18-0802/> [Accessed: 2020-12-10]
- [24] Twitter taught Microsoft's AI chatbot to be a racist asshole in less than a day. Available from: <https://www.theverge.com/2016/3/24/11297050/tay-microsoft-chatbot-racist> [Accessed: 2020-12-10]
- [25] Introducing Amazon Halo – Measure body composition, activity, sleep, and tone of voice - Winter + Silver – Medium. Available from: <https://www.amazon.com/Amazon-Halo-Fitness-And-Health-Band/dp/B07QK955LS> [Accessed: 2020-12-10]
- [26] General Data Protection Regulation. Available from: <https://gdpr-info.eu/> [Accessed: 2020-12-10]
- [27] Alexa and Echo devices are designed to protect your privacy. Available from: <https://www.amazon.com/b/?node=19149155011> [Accessed: 2020-12-10]
- [28] Amazon Echo Is Magical. It's Also Turning My Kid Into an Asshole. Available from: <https://hunterwalk.com/2016/04/06/amazon-echo-is-magical-its-also-turning-my-kid-into-an-asshole/> [Accessed: 2020-12-10]
- [29] Siri and Cortana Sound Like Ladies Because of Sexism. Available from: <https://www.wired.com/2015/10/why-siri-cortana-voice-interfaces-sound-female-sexism/> [Accessed: 2020-12-10]
- [30] Designing an Ethical Chatbot. Available from: <https://www.infoq.com/presentations/designing-chatbot-ethics/> [Accessed: 2020-12-10]
- [31] We tested bots like Siri and Alexa to see who would stand up to sexual harassment. Available from: <https://qz.com/911681/we-tested-apples-siri-amazon-echos-alexa-microsofts-cortana-and-googles-google-home-to-see-which-personal-assistant-bots-stand-up-for-themselves-in-the-face-of-sexual-harassment/> [Accessed: 2020-12-10]
- [32] Towards a Conversational Agent that Can Chat About...Anything. Available from: <https://ai.googleblog.com/2020/01/towards-conversational-agent-that-can.html> [Accessed: 2020-12-10]
- [33] DialoGPT: Toward Human-Quality Conversational Response Generation via Large-Scale Pretraining. Available from: <https://www.microsoft.com/en-us/research/project/large-scale-pretraining-for-response-generation/> [Accessed: 2020-12-10]
- [34] Mitsuku. Available from: <https://en.wikipedia.org/wiki/Mitsuku> [Accessed: 2020-12-10]
- [35] Cleverbot. Available from: <https://en.wikipedia.org/wiki/Cleverbot> [Accessed: 2020-12-10]
- [36] An act to add Chapter 6 (commencing with Section 17940) to

Part 3 of Division 7 of the Business and Professions Code, relating to bots. Available from: [https://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001](https://leginfo.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1001) [Accessed: 2020-12-10]

[37] AI Ethics. Available from: <https://www.ibm.com/artificial-intelligence/ethics> [Accessed: 2020-12-10]

[38] Exploring the real-world impacts of AI. Available from: <https://deepmind.com/about/ethics-and-society> [Accessed: 2020-12-10]

[39] Responsible bots: 10 guidelines for developers of conversational AI. Available from: <https://www.microsoft.com/en-us/research/publication/responsible-bots/> [Accessed: 2020-12-10]

[40] Ethics and Chatbots. Available from: <https://medium.com/pandorabots-blog/ethics-and-chatbots-8d4aab75cca> [Accessed: 2020-12-10]

[41] The ethics of artificial intelligence. Available from: <https://www.nickbostrom.com/ethics/artificial-intelligence.pdf> [Accessed: 2020-12-10]

[42] Project Debater. Available from: <https://www.research.ibm.com/artificial-intelligence/project-debater/> [Accessed: 2020-12-10]

[43] Soul Machines. Available from: <https://www.soulmachines.com/> [Accessed: 2020-12-10]