



Sberbank
Gamification
Laboratory

ГОЛОСОВЫЕ ИНТЕРФЕЙСЫ В БАНКОВСКОЙ ДЕЯТЕЛЬНОСТИ

Отчет Лаборатории геймификации

Исполнитель: Андрей Корбут

Москва 2020

Содержание

Почему банки используют голосовые интерфейсы?	3
Краткая история голосовых интерфейсов	5
Особенности голосовых интерфейсов	6
Ограничения применения голосовых интерфейсов в банковской деятельности	8
Сравнение голосовых и текстовых способов коммуникации с банком	9
Способы разрешения проблем при взаимодействии с банковским чат-ботом.	11
Персонализация голосовых банковских интерфейсов	13
Резюме	15

Почему банки используют голосовые интерфейсы?

Банки начали массово применять голосовые интерфейсы лишь несколько лет назад, когда появились достаточно эффективные технологии распознавания и синтеза речи и соответствующие алгоритмы. Многие исследователи возлагают большие надежды на использование голосовых интерфейсов в банках¹. Некоторые даже называют их будущим бизнеса в целом².

Банки и другие финансовые учреждения активно применяют технологии, предполагающие разговорные форматы предоставления услуг. Именно банки сегодня являются лидерами на рынке разработки и использования чат-ботов. Все крупнейшие банки и финансовые учреждения имеют либо отдельного цифрового ассистента, с которым пользователи могут взаимодействовать через мобильное приложение или веб, либо чат-бота, работающего на разных платформах (на Западе наиболее популярны платформы IBM Watson Assistant и Facebook Messenger).

Все эти процессы принято обозначать термином «разговорный банкинг» (conversational banking), которые в последние годы приходит на смену мобильному банкингу, сменившему, в свою очередь, онлайн-банкинг, телефонный банкинг и обслуживание через банкоматы. Временные рамки этих парадигм банкинга приблизительно таковы³:



Помимо чисто экономических (возможность экономии за счет сокращения физических работников) и технических (доступ 24/7) причин, привлекательность разговорного банкинга обуславливается несколькими факторами.

1. Главный фактор — *привычность* формата взаимодействия: разговор (как устный, так и письменный) хорошо

¹ **Bassano C.** The Future of Banking: It's Time to Get Conversational // Business 2 Community. 2019. November 19. URL: <https://www.business2community.com/finance/the-future-of-banking-its-time-to-get-conversational-02259542>

² **Newman D.** Voice Interface Technology: The Future of Business // Forbes. 2018. August 22. URL: <https://www.forbes.com/sites/danielnewman/2018/08/22/voice-interface-technology-the-future-of-business/>

³ **Andreoli A., Billé G., Meleagri N.** Ready to Talk! Banks are Embracing the Power of Conversational Banking. Accenture, 2017. URL: <https://www.accenture.com/us-en/insight-conversational-banking>

знаком пользователям, поэтому разговорные интерфейсы создают впечатление естественности происходящего и приближают взаимодействие с банковскими приложениями к взаимодействию с реальным сотрудником. В банковской сфере это означает повышения доверия к оказываемым услугам и большую лояльность к бренду.

2. Голосовое взаимодействие позволяет получать банковскую информацию и совершать банковские операции, не прерывая других форм деятельности. Голосовой интерфейс допускает *многозадачность* в гораздо большей степени, чем графический. Голосовой интерфейс позволяет иметь доступ к определенным услугам в тех ситуациях, в которых пользователи раньше не могли его иметь, например при вождении автомобиля.
3. Голосовой формат взаимодействия не требует *специальных* навыков, в частности, компьютерной грамотности. Взаимодействовать с голосовым интерфейсом могут даже люди с низким общим уровнем грамотности, что особенно актуально для развивающихся стран и регионов.
4. Разговорный банкинг позволяет сильнее *вовлекать* пользователей. Пользователи охотнее используют банковские сервисы и больше времени взаимодействуют с банком в силу привычности и удобства разговорного формата взаимодействия.
5. Разговорный банкинг также позволяет значительно повышать *персонализацию* предоставляемых банком услуг. Пользователи могут получать советы от цифровых помощников или чат-ботов относительно того, как управлять своими деньгами, исходя из прошлой истории расходов. Пользователи больше доверяют советам банков, когда эти советы предоставляются в разговорной форме.

Банки могут как разрабатывать собственные продукты, так и использовать технологии, продающиеся фирмами-разработчиками. Сегодня существует множество компаний, представляющих банкам услуги по разработке разговорных, в том числе голосовых, интерфейсов. Лидеры отрасли — CBOT⁴, Finn AI⁵, Kasisto⁶, Clinc⁷.

Наибольшее распространение голосовые интерфейсы получили в области обслуживания клиентов, но они также используются в области продаж и маркетинга. Удобство голосовых интерфейсов в случае работы с клиентами обуславливается тем, что существует множество стандартных запросов, которые могут отправляться и обрабатываться с помощью технологий распознавания и синтеза речи и соответствующих алгоритмов с очень высокой степенью эффективности, даже несмотря на ошибки, возникающие при взаимодействии с этими системами. Голосовые интерфейсы могут использоваться как в контакт-центрах, так и в банкоматах, на сайтах банков, на страницах банков в социальных сетях, в мобильных банковских приложениях и в умных колонках.

В целом использование чат-ботов повышает *лояльность* к банку⁸, но только в том случае, если качество компьютерной системы высоко, т.е. пользователи могут легко понять, как ей пользоваться, и системы отвечает правильно и быстро. Соображения, связанные с рисками передачи финансовой информации во время взаимодействия, играют важную, но второстепенную роль.

⁴ <https://www.cbot.ai>

⁵ <https://www.finn.ai>

⁶ <https://kasisto.com>

⁷ <https://clinc.com>

⁸ **Trivedi J.** Examining the Customer Experience of Using Banking Chatbots and Its Impact on Brand Love: The Moderating Role of Perceived Risk // Journal of Internet Commerce. 2019. Vol. 8. № 1. P. 91–111. URL: <https://doi.org/10.1080/15332861.2019.1567188>

Краткая история голосовых интерфейсов

Историю голосовых интерфейсов принято отсчитывать от 1952 года, когда Bell Labs, знаменитое исследовательское подразделение компании Bell, представило автоматический цифровой распознаватель Одри. Одри могла распознавать цифры от 0 до 9, произносимые пользователем, с точностью 97–99%, однако сложность и дороговизна оборудования не позволили превратить Одри в коммерческий продукт.

В 1962 году компания IBM представила Shoebox — машину, способную понимать 16 устных слов на английском языке.

В 1971 году DAPRA, исследовательское подразделение министерства обороны США, запустило пятилетний проект по разработке системы, способной полноценно распознавать человеческую речь (Speech Understanding Research). В проекте участвовали несколько исследовательских центров и компаний. Проведенные ими исследования заложили основы современных технологий распознавания речи.

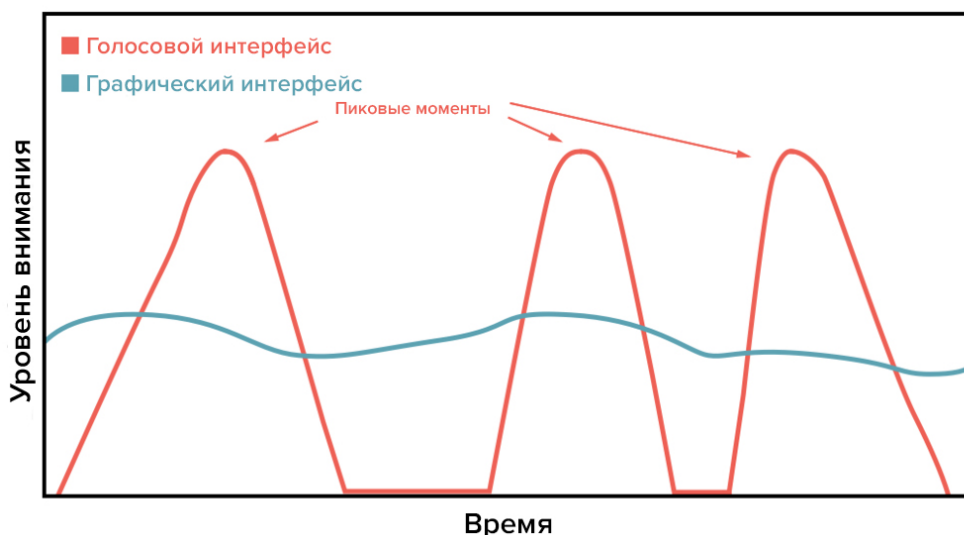
В 1982 году была основана компания Dragon — один из ведущих разработчиков голосовых технологий (другая компания-пионер в этой области — SpeechWorks, была основана в 1984 году). В 1990 году Dragon выпустила Dragon Dictate, первую доступную потребителям программу для перевода устных высказываний в текст.

В 1996 году компания Charles Schwab при содействии компании Nuance предложили Voice Broker — первого голосового брокера, которые давал информацию о котировках акций. Его точность оценивалась в 95%. В этом же году компания BellSouth запустила первый в мире голосовой портал VAL (позже — Info By Voice), однако система была довольно неточной.

В 2010 году вышел Siri — голосовой помощник от компании Apple (концепт был представлен в 2006 году), после чего начали появляться голосовые помощники от других ведущих компаний.

Особенности голосовых интерфейсов

Голосовые интерфейсы, в отличие от графических, не предполагают более или менее постоянное внимание к самому интерфейсу. Д. Вестелунд представляет эту характеристику следующими образом⁹:



⁹ **Westerlund D.** How To Deal With Cognitive Load In UX and Voice Design // CareerFoundry. 2017. August 17. URL: <https://careerfoundry.com/en/blog/ux-design/voice-ui-design-and-cognitive-load/>

Поскольку взаимодействие посредством голоса предполагает чередование позиции говорящего и слушающего, — причем так, чтобы это чередование происходило без пауз между сменой говорящий и без наложения высказываний, — в процессе взаимодействия с голосовым интерфейсом пользователи должны полностью концентрироваться на действиях системы, когда они слушают, и почти не уделять ей внимание, когда они говорят. Внимание в процессе взаимодействия с голосовым интерфейсом *дискретно*, в то время как внимание к графическому интерфейсу *континуально*: оно более или менее постоянно. В случае голосового интерфейса пользователи не могут отвлечься и вернуться к какому-либо его элементу позже, как в случае графического интерфейса.

Поэтому голосовое взаимодействие представляет собой серию *пиков внимания*, как их называет Вестерланд. Во время этих пиков внимания, которые наблюдаются, когда пользователи слушают, когнитивная нагрузка на пользователей значительно увеличивается, поскольку пользователям необходимо не только понимать произносимые системой высказывания, но и связывать то, что они слышат, с предыдущим ходом разговора, который они удерживают в памяти. В отличие от взаимодействия с графическим интерфейсом, во время взаимодействия с голосовым интерфейсом пользователи не имеют возможности вернуться назад и, например, перечитать текст. Взаимодействие разворачивается строго темпорально: пользователи движутся вперед по мере разворачивания диалога. Им приходится ориентироваться на текущую коммуникативную ситуацию и фокусировать все внимание на ней.

В силу этого при взаимодействии с голосовыми интерфейсами на первое место выходит проблема *секвенциальности*: того, в какой последовательности пользователям предоставляется определенная информация и какова последовательность их собственных действий.

Поскольку во время пиков внимания когнитивная нагрузка на пользователя очень высока, при разработке голосовых интерфейсов необходимо делать так, чтобы, во-первых, не перегружать пользователей информацией, сообщаемой, пока они слушают, и, во-вторых, организовывать взаимодействие таким образом, чтобы пики внимания не были слишком частыми. Пользователям сложно справиться с разговором, если им приходится быстро и часто переключаться из позиции говорящего и позицию слушающего.

Другая особенность голосового интерфейса заключается в том, что для передачи определенного сообщения голосом требуются гораздо больший объем информации. То, что в графическом интерфейсе можно представить в виде кнопки с надписью «ОТПРАВИТЬ», в голосовом интерфейсе должно быть представлено в виде фразы, например «Хотите ли вы отправить эти сведения?». И сама понятность этой фразы будет обеспечиваться не только *функциональным* контекстом действия (решаемой задачей), но и *коммуникативным* контекстом, т.е. правилами организации разговора и его историей.

Этот коммуникативный контекст оказывает прямо влияние на *продолжительность* взаимодействия с голосовым интерфейсом. Поскольку это взаимодействие напоминает «нормальную» коммуникацию, пользователи применяют к нему те же правила смены говорящих в разговоре, которые применяются в привычных коммуникативных ситуациях. Например, они стараются не прерывать систему, пока та произносит фразу, до того места, где прерывание будет уместно, что приводит к увеличению количества времени, затрачиваемого на решение задачи.

Ограничения применения голосовых интерфейсов в банковской деятельности

Для голосового взаимодействия важно содержание тех задач, которые в ходе его решаются. В случае банковской информации, которая во многом представляет собой цифры, разработка голосовых интерфейсов сталкивается с той проблемой, что пользователям гораздо проще воспринимать цифры графически, чем на слух. Это касается как собственно распознавания цифр, так и оценки отношений между ними. Довольно легко установить, что 3000 меньше 5000, если видеть эти цифры, но, если они произносятся («три тысячи», «пять тысяч»), это становится сложнее. Поскольку манипулировать цифрами удобнее в графической, а не голосовой форме, это как накладывает ограничения на характеристики содержания голосовых интерфейсов, так и снижает степень доверия пользователей к голосовым интерфейсам.

Другое ограничение, обусловленное спецификой банковской деятельности, связано с тем, что банковская информация воспринимается пользователями как *приватная* информация и, соответственно, разглашение этой информации регулируется гораздо строже, чем многих других сведений. При этом голосовой интерфейс предполагает такой формат взаимодействия, при которой любая сообщаемая информация оказывается потенциально публичной, поскольку проговаривается вслух. Эта особенность приводит к тому, что пользователи начинают гораздо внимательнее относиться к обстановке, в которой они осуществляют голосовое взаимодействие. В частности, пользователи начинают избегать голосового взаимодействия с банком в публичных ситуациях¹⁰ или ограничивают это взаимодействие теми задачами, которые не требуют сообщения конфиденциальной, с их точки зрения, информации. В целом пользователи не хотят, чтобы им приходилось сообщать банковскую информацию. Они хотят, чтобы это делал голосовой банковский помощник (поскольку они могут управлять тем, какая информация от голосового помощника будет публично доступна, используя для этого, например, наушники), а им нужно было только кратко отвечать.

¹⁰ **Lundqvist J.** Designing a Trustworthy Voice User Interface for Payments and Transactions: A Study in User Experience Design. Umeå University, 2019. URL: <https://www.diva-portal.org/smash/get/diva2:1360094/FULLTEXT01.pdf>

Сравнение голосовых и текстовых способов коммуникации с банком

Ожидания от разговорного банковского помощника выше, чем от текстового¹¹. Пользователи чаще реагируют на ошибки системы, пытаясь их исправить, и прилагают больше усилий для исправления (дают более многословные объяснения, больше перефразируют, меняют интонацию). Такое приписывание более продвинутых коммуникативных способностей ведет к тому, что пользователи воспринимают голосовых помощников как более помогающих и приписывают ошибки во взаимодействии скорее себе, чем системе.

В одном из исследований¹² было изучено взаимодействие пользователей (начинающие пользователи из сельских регионов Индии) с тремя прототипами банковских интерфейсов: аудио (IVR), аудиовизуальным и графическим. Пользователи выполняли две базовые задачи: поиск ближайшего банкомата и перевод денег другому лицу. Было выявлено, что по показателю успешности выполнения задачи аудиовизуальный интерфейс превосходит чисто аудиальный и графический (при этом IVR позволяет лучше справляться с задачей, чем чисто графический интерфейс). По показателю скорости выполнения задачи на первом месте был графический интерфейс, а на втором — аудиовизуальный.

В другом исследовании¹³ были проанализированы пользовательские запросы к банковскому чат-боту в голосовом (через умную колонку Google Home) и текстовом (через Google Assistant) формате. Пользователи в целом предпочли голосовой интерфейс текстовому, поскольку в случае голосового интерфейса машина гораздо лучше справлялась с распознаванием высказываний. По-видимому, это было связано с тем, что, в отличие от текстового ввода, в случае голосового ввода ошибки пользователя легче сгладить. Например, в случае текстового ввода слово «бланс» сложнее идентифицировать, чем аналогичную ошибку в устной речи, поскольку распознавание устной речи происходит иначе, чем распознавание письменной (система распознает звуки, а не символы).

С другой стороны, в еще одном исследовании¹⁴, где было сопоставлено голосовое и текстовое взаимодействие с цифровым помощником — финансовым менеджером, было выявлено, что пользователи предпочитают текстовый формат. Такой результат мог быть обусловлен как содержанием взаимодействия (в исследовании пользователи должны были искать определенную финансовую информацию), так и типом пользователей (это были профессиональные финансовые менеджеры). Однако данные (пока еще достаточно скудные) о том, как пользователи пользуются реальными банковскими цифровыми помощниками с возможностью голосового взаимодействия, подтверждают, что доверие к голосовым интерфейсам не так уж велико. Например, в случае одного из самых известных и популярных банковских цифровых помощ-

¹¹ **Kim S., Goh J., Jun S.** The Use of Voice Input to Induce Human Communication with Banking Chatbots // HRI'18: Companion of the 2018 ACM/IEEE International Conference on Human-Robot Interaction. Chicago. N.Y.: ACM, 2018. P. 151–152. URL: <https://doi.org/10.1145/3173386.3176970>

¹² **Shrivastava A., Joshi A.** Directedness and Persistence in Audio-Visual Interface for Emergent Users // IndiaHCI'19: Proceedings of the 10th Indian Conference on Human-Computer Interaction. N.Y.: ACM, 2019. Art. 10. URL: <https://doi.org/10.1145/3364183.3364191>

¹³ **Doherty D., Curran K.** Chatbots for Online Banking Services // Web Intelligence. 2019. Vol. 17. № 4. P. 327–342. URL: <https://doi.org/10.3233/WEB-190422>

¹⁴ **Lane D., Renwick R., McAvoy J., O'Reilly Ph.** Voice-Based User Interface vs Text-Based User Interface: Comparing Perceived Usability of Two Natural Language Processor (NLP) Based Question Answering (QA) Systems // Irish Academy of Management 2018 Conference (2018, September 3–5). URL: https://www.researchgate.net/publication/326718916_Voice-based_User_Interface_vs_Text-based_User_Interface_Comparing_perceived_usability_of_two_natural_language_processor_NLP_based_question_answering_QA_systems

ников — Erica от Bank of America, — лишь 13% пользователей используют голосовой интерфейс¹⁵.

Проблемы, с которыми сталкиваются пользователи при использовании цифровых помощников (как с голосовым, так и с текстовым интерфейсом), оказывают существенное влияние на предпочтения в отношении того или иного канала взаимодействия с банком. Как было показано в одном из исследований¹⁶, со временем пользователи склонны отказываться от взаимодействия чат-ботом, поскольку современные чат-боты не могут вести разговор с пользователем достаточно эффективно. Первоначальное позитивное отношение к чат-боту сменяется скептическим отношением или вообще отказом от его использования, потому что первоначально разговорный формат взаимодействия создает ощущение новизны и придает взаимодействию характер игры, однако, когда с течением времени (довольно быстро; в указанном исследовании пользователя интервьюировали через 2 недели после начала использования цифрового банковского помощника) пользователи узнают, как работает система и каковы ее возможности, они теряют к ней интерес.

¹⁵ **Streeter B.** BofA's Surging Digital Assistant Signals a Turning Point in Retail Banking // The Financial Brand. 2019. June 11. URL: <https://thefinancialbrand.com/84350/bank-of-america-chatbot-erica-retail-banking-virtual-assistant/>

¹⁶ **Demirci H.M., Berberoğlu M.** User Experience over Time with Personal Assistants of Mobile Banking Application in Turkey // Intelligent Human Systems Integration 2020: Proceedings of the 3rd International Conference on Intelligent Human Systems Integration (IHSI 2020): Integrating People and Intelligent Systems (February 19–21, 2020, Modena, Italy). Cham: Springer, 2020. P. 646–652. URL: https://doi.org/10.1007/978-3-030-39512-4_100

Способы разрешения проблем при взаимодействии с банковским чат-ботом

В силу новизны голосовых интерфейсов в банковской деятельности, пока практически нет систематических исследований реальных повседневных практик взаимодействия с голосовыми банковскими интерфейсами. Однако некоторое представление о специфике этого взаимодействия можно получить из исследований текстовой коммуникации с банковскими чат-ботами. Результаты этих исследований можно обобщать и распространять на голосовые взаимодействия, поскольку в основе как голосового, так и текстового взаимодействия с банковскими чат-ботами лежат одни и те же технологии искусственного интеллекта, только в случае голосового интерфейса добавляются модули распознавания и синтеза речи. Поэтому значительная часть проблем, с которыми сталкиваются пользователи текстового разговорного интерфейса, встречается и при использовании голосового интерфейса.

Одно из таких исследований¹⁷ было проведено в 2017 году в Национальном университете Цзяотун (Тайвань). Авторы проанализировали взаимодействия 1685 пользователей с банковским чат-ботом в течение 3 месяцев (всего 17 136 взаимодействий; под взаимодействием в исследовании понималась две смежные реплики: ввод со стороны пользователя и отклик со стороны чат-бота; при этом изучались разговоры в целом, а не только две смежные реплики). Чат-бот предоставлял только информацию, не требующую идентификации пользователя: информацию о курсах валют, о кредитных картах, о кредитах на приобретение жилья и об инвестировании. В проанализированных данных было выявлено, что большинство взаимодействий (90,6%) не содержало сбоев, прежде всего потому, что большая их часть заключалась в выборе пользователем вариантов действия, предложенных чат-ботом. В 9,4% случаев, содержащих сбой, подавляющее большинство (97,4%) проблем возникало в ситуациях, когда пользователь набирал текст, а не использовал предлагаемые опции. Сбои также чаще всего (в 88% случаев) возникали, когда пользователи запрашивали информацию, а не предоставляли ее.

В данных были выявлены следующие сбои:

Ошибка распознавания	Неправильное распознавание	Не-распознавание
Ожидаемое содержание	43%	45,2%
Неожиданное содержание		
Дополнительное объяснение	1,6%	2,5%
Повторное начало	0,4%	0,4%
Поддержание темы	0,4%	1,3%
Незаконченное сообщение	0,8%	2,5%
Окончание незаконченного сообщения	0,7%	1,7%

¹⁷ Li C.-H., Yeh S.-F., Chang T.-J., Tsai M.-H., Chen K., Chang Y.-J. A Conversation Analysis of Non-Progress and Coping Strategies with a Banking Task-Oriented Chatbot // CHI'20: Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems. N.Y.: ACM, 2020. P. 1–12. URL: <https://doi.org/10.1145/3313831.3376209>

Для преодоления возникающих сбоев пользователи использовали следующие стратегии:

Переформулирование сообщения

добавление слов	6,68%
убирание слов	4,76%
перефразирование	8,82%
повторение	5,75%
запрос новой темы	5,48%
другое	3,56%

Прекращение коммуникации

временный уход от темы	27,16%
временное превращение разговора	6,74%
смена темы	13,47%
отказ от использования чат-бота	17,58%

Как можно видеть, чаще всего пользователи предпочитают прекращать коммуникацию, сменяя тему. Довольно значительное число пользователей при этом вообще отказывается от дальнейших взаимодействий с чат-ботом. Наибольшее влияние на принятие решения от отказе от использования чат-бота оказывали два типа сбоев: когда пользователи давали дополнительное объяснение, но чат-бот не мог распознать его, и когда пользователи заканчивали незаконченное сообщение, но чат-бот не мог распознать его. Оба эти типа сбоев предполагают предоставление пользователем неожиданного содержания (содержания, которое не требует чат-бот). Также хорошим предсказателем отказа от использования чат-бота является количество следующих друг за другом сбоев, с которыми сталкивается пользователем: 90% тех, кто столкнулся с 3 сбоями подряд, отказываются от дальнейшего взаимодействия с чат-ботом.

Безусловно, голосовое взаимодействие будет иметь свою специфику. Пользователи будут использовать дополнительные стратегии преодоления затруднений, возможные благодаря речи, например менять интонацию, громкость и скорость произнесения высказываний. Тем не менее базовые трудности, с которыми они будут сталкиваться, и способы их преодоления будут, скорее всего, идентичны тем, которым наблюдаются в случае текстового взаимодействия с банком.

Персонализация голосовых банковских интерфейсов

Один из способов повышения доверия и удобства использования голосовых интерфейсов в банке — использование аватаров, с которыми пользователи могут вести разговор. Виртуальные банковские помощники начали появляться совсем недавно, однако уже есть несколько систем, близких к полноценному релизу. Например, в 2018 году новозеландский банк ANZ вместе с фирмой Soul Machines разработал цифрового помощника Джейми — консультанта, которая может предоставлять пользователям различную банковскую информацию через веб-сайт и мобильное приложение.



Джейми отвечает на общие вопросы, не требующие предоставления персональной информации: как перевести деньги, открыть счет, заменить карту и т.д. За первые 100 дней работы Джейми обслужила 12 000 запросов от пользователей¹⁸. Однако успешно выполнить смогла только 60% запросов.

Такая персонализация тесно связана с персонализацией: пользователи, общающиеся с виртуальной сотрудницей банка, должны, по идее, чувствовать более персональное отношение банка к ним. Банк при этом выступает в виде приветливого аватара, который никогда не злится и готов оказать помощь в любое время. Компания-разработчик Джейми Soul Machines называет своих аватаров «цифровыми людьми».

Персонализация достигается тремя способами:

1. С помощью *аватара*, имеющего четкие индивидуальные характеристики (в случае Джейми аватар гиперреалистичен).
2. С помощью *голосового* интерфейса и *разговорного* искусственного интеллекта, которые создают у пользователя впечатления «нормальной» коммуникации.

¹⁸ **Anthony J.** ANZ's Artificial Intelligence Entity Jamie Took 12,000 Customer Inquiries in its First 100 Days // Stuff. 2018. November 18. URL: <https://www.stuff.co.nz/business/business-top-stories/108681909/anzs-artificial-intelligence-entity-jamie-took-12000-customer-inquiries-in-its-first-100-days>

3. С помощью *нарративизации*: цифровой помощник говорит о себе «я» и демонстрируют определенные предпочтения, например говорит какой цвет ей (чаще всего это женщина, как в случае Джейми) нравится.

Однако насколько эффективна персонификация? В проведенном в 2016 году исследовании 851 продуктовых отзывов на Amazon Alexa¹⁹ было выявлено, что персонификация неразрывно связана с двумя факторами: с количеством взаимодействий с цифровым помощником и с количеством людей, проживающих в домохозяйстве. Персонифицируют цифрового помощника в первую очередь, те, кто активно с ней общается, и те, кто не живет один. Персонификация также приводит к большей удовлетворенности цифровым помощником.

Проблема с персонификацией заключается в том, что она создает завышенные ожидания относительно способностей цифрового помощника. Эти ожидания довольно быстро не оправдываются, в том числе — по чисто коммерческим причинам (например, разработчики делают так, что их цифровые ассистенты не имели своего «мнения» по чувствительным темам вроде феминизма²⁰, чтобы не вызывать негативные реакции у пользователей). Обнаружение того, что у цифрового помощника нет «личности» в общепринятом значении слова, может приводить к тому, что пользователи перестают доверять технологии в целом, даже если она успешно решает узкие задачи²¹.

Другая проблема с персонификацией заключается в том, что пользователи, реагируя на демонстрацию «личностных» особенностей аватара, начинают относиться к нему или ей как в обладателю этих качеств и демонстрировать в отношении их определенные формы поведения (часто — стереотипные). Несмотря на все попытки разработчиков сделать аватары максимально «нейтральными», сама задача персонификации заставляет придавать цифровому помощнику характеристики, провоцирующие соответствующее поведение. Это часть того, что получило название «*нечестной антропоморфизации*»²², которая чревата чрезмерно негативной или позитивной реакцией в отношении аватара в зависимости от его качеств. Нечестная антропоморфизация — это демонстрация час-ботом (или любой другой системой) человеческих характеристик, которыми система на самом деле не обладает и которое создают у пользователя ложное впечатление о том, с кем они взаимодействуют. Это может происходить как в силу того, что разработчики скрывают реальные возможности системы, так и в силу того, что они стремятся сделать систему более «привычной» пользователю. Чаще всего такая нечестная антропоморфизация решает маркетинговые задачи, но ее последствия могут не только прямо противоречить поставленным целям (пользователи будут отказываться от использования таких систем), но и сказываться на характере взаимодействия (например, пользователи могут взаимодействовать с системой, разговаривающей женским голосом, исходя из своих стереотипных представлений об обслуживающей роли женщин).

¹⁹ **Purington A., Taft J.G., Shruti Sannon, Bazarova N.N., Hardman Taylor S.** “Alexa is my new BFF”: Social Roles, User Satisfaction, and Personification of the Amazon Echo // CHI EA’17: Proceedings of the 2017 CHI Conference Extended Abstracts on Human Factors in Computing Systems. N.Y.: ACM, 2017. P. 2853–2859. URL: <https://doi.org/10.1145/3027063.3053246>

²⁰ **Hern A.** Apple Made Siri Deflect Questions on Feminism, Leaked Papers Reveal // The Guardian. 2019. September 6. URL: <https://www.theguardian.com/technology/2019/sep/06/apple-rewrote-siri-to-deflect-questions-about-feminism>

²¹ **Toader D.-C., Boca G., Toader R., Măcelaru M., Toader C., Ighian D., Rădulescu A.T.** The Effect of Social Presence and Chatbot Errors on Trust // Sustainability. 2019. Vol. 12. № 1. Art. 256. URL: <https://www.mdpi.com/2071-1050/12/1/256>

²² **Leong B., Selinger E.** Robot Eyes Wide Shut: Understanding Dishonest Anthropomorphism // FAT*’19: Proceedings of the Conference on Fairness, Accountability, and Transparency. N.Y.: ACM, 2019. P. 299–308. URL: <https://doi.org/10.1145/3287560.3287591>

Резюме

- Взаимодействие с голосовыми интерфейсами характеризуется *дискретностью*: пользователи уделяют максимальное внимание интерфейсу, когда слушают, и минимальное — когда говорят. Предоставление пользователю слишком большого объема информации на пиках внимания и высокая частота этих пиков значительно увеличивают когнитивную нагрузку на пользователя и приводит к сбоям во взаимодействии.
- Взаимодействие с голосовыми банковскими интерфейсами имеет два базовых *ограничения*: 1) пользователям легче воспринимать цифры в виде текста, чем на слух, и 2) пользователи беспокоятся о приватности передаваемой финансовой информации и поэтому избегают взаимодействия с голосовыми банковскими интерфейсами в публичных местах, предпочитая быть получателем, а не отправителем банковской информации.
- Голосовые интерфейсы первоначально оцениваются пользователями более позитивно, чем графические и текстовые, но с течением времени пользователи *отказываются* от них, предпочитая общаться с банком посредством текста. Это означает, что, хотя голосовые интерфейсы привлекают своей новизной и необычностью (по сравнению с уже привычными формами взаимодействия), они рассматриваются, скорее, как *игрушка*, чем как серьезный механизм коммуникации.
- Еще одной причиной отказа от использования голосовых интерфейсов являются *слабые коммуникативные способности* современных систем разговорного искусственного интеллекта. Неспособность распознать, что говорит пользователь, приводит к отказу от использования.
- Попытки разработчиков сделать голосовой банковский интерфейс более *персонализированным* сталкиваются с проблемой «нечестной антропоморфизации», когда голосовой помощник создает завышенные ожидания относительно своих коммуникативных способностей, формируя у пользователя впечатление полноценной личности, которое современные алгоритмы, однако, не способны поддерживать.