

A transparência de sistemas de recomendação de plataformas de Vídeo sob Demanda (VoD): categorias de conteúdos

Transparency of Video on Demand (VoD) platforms recommender systems: content categories

RUIZ, Cinthia; Doutora; PUC-Rio

cinthia.ruiz@gmail.com

QUARESMA, Manuela; Doutora; PUC-Rio

mquaresma@puc-rio.br

Este artigo objetiva investigar a transparência do sistema de recomendação de plataformas de vídeo sob demanda baseado em Inteligência Artificial e Aprendizado de Máquina. Com os algoritmos cada vez mais presentes em nossas vidas e trazendo mudanças de interação, faz-se necessário o olhar sobre a experiência de uso. A falta de transparência dos algoritmos acarreta problemas para o usuário, como modelo mental inadequado, dificuldades para o controle do sistema e falta de confiança nas recomendações. Foi realizado, então, um estudo por meio de um workshop de design para envolver usuários e designers em possíveis soluções com *Explainable AI*. Como resultado, foram apontadas oportunidades para melhorar a explicabilidade das recomendações.

Palavras-chave: Experiência do Usuário; Transparência; Aprendizado de máquina.

This article aims to investigate the transparency of video-on-demand platforms recommender system based on Artificial Intelligence and Machine Learning. With algorithms increasingly present in our lives and bringing changes in interaction, it is necessary to look at the experience of use. The lack of transparency of the algorithms causes problems for the user, such as an inadequate mental model, difficulties in controlling the system and lack of confidence in the recommendations. A study was then carried out through a design workshop to involve users and designers in possible solutions with Explainable AI. As a result, opportunities were identified to improve the explainability of the recommendations.

Keywords: User Experience; Transparency; Machine Learning.

1 Introdução

O avanço tecnológico dos últimos anos permite o armazenamento de grande volume de dados e seu processamento em alta velocidade. Surgem então, diversos produtos baseados em Inteligência artificial. Rouhiainen (2018) diz que a inteligência artificial, ou AI (*Artificial Intelligence*) é a capacidade das máquinas de usar algoritmos para aprender com os dados e tomar decisões, como um ser humano faria. O autor expõe que o aprendizado de máquina, ou ML (*Machine Learning*), é uma das principais abordagens da inteligência artificial, com a qual

máquinas têm a capacidade de aprender com os dados. Os conhecimentos do ML já estão presentes em muitas das interfaces inovadoras com as quais interagimos, como mecanismo de busca, links recomendados em e-commerce, sistemas de streaming de áudio ou vídeo, aplicações médicas etc. Nas, temos algoritmos filtrando o conteúdo e selecionando o mais apropriado a cada perfil, criando uma experiência personalizada para cada usuário.

Segundo Oxborough et al. (2018), a AI está evoluindo em sofisticação, complexidade e autonomia, o que abre oportunidades de transformação para os negócios e a sociedade. Em consequência, surgem questões relacionadas a como aproveitar o seu potencial e quais são os riscos da AI. Assim, com tantas questões envolvendo principalmente a ética e as possibilidades de futuro com o emprego de AI nos sistemas digitais, diversas entidades, como grupos de pesquisadores, órgãos governamentais e instituições privadas, se dedicaram a definir princípios de AI. A intenção dessas listas é fornecer práticas recomendadas para proteger a sociedade e seus indivíduos, garantindo o melhor aproveitamento do potencial da AI. O princípio da transparência aparece em algumas delas, como: *Google AI Principles* (Google PAIR, 2019), *Asilomar AI Principles* (FLI, 2017), *Microsoft AI principles* (Microsoft, 2018), *AI R&D Guidelines for International Discussions* (MIC, 2017), *OECD Principles on AI* (OECD, 2019) e *Recommendation on the ethics of artificial intelligence* (UNESCO, 2021).

Gilpin et al. (2019) afirmam que, conforme as máquinas começam a substituir os humanos em suas decisões, torna-se necessário que esses mecanismos se expliquem. Alertam que, apesar de alcançarem sucesso em diversas áreas, existe uma desconfiança geral sobre seus resultados, problema que poderia ser resolvido com a explicabilidade. Em contraste, Oxborough et al. (2018) defendem que é a criticidade que determina a necessidade da explicabilidade. Os autores apontam que, ao se tratar de sistemas de recomendação, a maioria dos usuários confia nos resultados, sem a necessidade de transparência do algoritmo de ML, já que haveria fácil compreensão e baixo risco. Eles exemplificam com uma plataforma de VoD, na qual uma recomendação errada acarretaria apenas em desperdício de um pouco do tempo do usuário. Parece que os autores desconsideraram os prejuízos a experiência do usuário. Já Bharadhwaj et al. (2018), por defenderem a explicabilidade das recomendações, propuseram um modelo e o testaram com o conjunto de dados da Netflix, demonstrando a eficácia da precisão e explicabilidade das predições. Os autores descrevem uma estrutura para recomendações temporais explicáveis em um modelo de *Deep Learning*, considerando um esquema baseado em vizinhança para gerar explicações.

Esta pesquisa propõe-se a investigar a transparência do sistema de recomendação das plataformas de vídeo sob demanda baseado em AI/ML. Parte-se da hipótese de que trabalhar a transparência dos sistemas de recomendação baseados em ML melhora a experiência do usuário, uma vez que ajuda na formação do correto modelo mental do usuário sobre seu funcionamento, impactando no controle e confiança. Sendo assim, esta pesquisa objetiva trazer recomendações para melhorar a experiência do usuário com sistemas de recomendação baseados em ML das plataformas de vídeo sob demanda, no que tange o princípio da transparência. Para isso, apontam-se possíveis soluções a problemas encontrados na experiência do usuário em relação às predições do ML em plataformas de VoD. Para alcançar seu objetivo, realizou-se um estudo por meio de um workshop de design como método.

2 Inteligência Artificial Explicável (XAI)

A informação disponibilizada pelo ambiente é processada através de uma série de sistemas de processamento humanos, como atenção, percepção e memória de curto prazo. Segundo Wang

(2008), embora não haja um consenso sobre o significado do termo inteligência artificial, há concordância sobre o fato de os seres humanos diferirem de animais e máquinas principalmente em sua habilidade cognitiva, comumente chamada de “inteligência”, e a AI é a tentativa de reproduzir essa habilidade em sistemas computacionais. Ela se espelha no processamento humano de informação, para possibilitar os computadores desempenharem tarefas as quais os humanos já realizam, mas de maneira mais rápida e eficiente.

Com a proliferação de sistemas digitais baseados em AI, surgem diversas questões envolvendo princípios éticos. Além disso, a abordagem centrada no humano leva o foco para a experiência dos usuários com esses sistemas e estuda as suas necessidades. Uma questão em ascensão é a relação da transparência desses sistemas com a confiança dos usuários, já que a maioria dos sistemas são considerados caixas pretas – do inglês *black box*. Nesse cenário, emerge a Inteligência Artificial Explicável (XAI), com a proposta de explicar a lógica da tomada de decisão dos algoritmos.

Lamy et al. (2019) definem XAI como uma área focada em projetar sistemas inteligentes, capazes de explicar seus resultados para um ser humano. Um dos objetivos da XAI é fornecer explicações sobre a lógica das recomendações dos algoritmos de ML, segundo Weitz et al (2020), ajudando assim o usuário a entender e confiar no sistema. Para Gilpin et al. (2019), a XAI permite que o sistema seja transparente, fornecendo explicações sobre suas decisões.

Embora os algoritmos sejam programados pelos humanos, eles encontram relacionamentos entre os dados e criam suas próprias lógicas, que um humano não veria coerência ou sentido. Como resultado, temos previsões baseadas em lógicas que não são comprehensíveis para os humanos. É dizer, nem o desenvolvedor do sistema, nem o usuário final, sabem a lógica seguida pelo ML para chegar até o resultado. O campo da XAI atrai cada vez mais atenção, pois propõe justamente o desenvolvimento de métodos com os quais os resultados sejam transparentes: rastreáveis, explicáveis e comprehensíveis. No entanto, Samek, Wiegand e Müller (2017) ressaltam que os algoritmos de ML com melhor desempenho geralmente são *black box*, que não informam os motivos das previsões.

O ML oferece recomendações para decisões diárias de um público bem diversificado. Ribera e Lapedriza (2019) comentam que com essa mudança no cenário de uso, os usuários de AI vão exigir cada vez mais explicações, em busca de informações fundamentadas e não simplesmente decisões tomadas pelo sistema. Devido à variabilidade do público, os autores defendem explicações diferentes para públicos diferentes, já que os objetivos, expectativas, experiências e necessidades são diferentes. Dessa maneira, o sistema deveria ser capaz de interpretar todas as particularidades de cada usuário, para projetar o nível de detalhe ideal que atendesse as necessidades de cada um.

O estudo de Oxborough et al. (2018) relata que nos últimos 30 anos, a evolução da AI foi interrompida em diversos momentos por limitações tecnológicas. No entanto, o atual avanço tecnológico é propício para o seu desenvolvimento, mas ganhar a confiança da sociedade pode ser uma nova barreira para a evolução da AI. Por essa razão, defende a necessidade da XAI.

É fácil perceber a importância da XAI quando há alto risco envolvido, como no campo da medicina. Um sistema baseado em ML pode auxiliar no diagnóstico de pacientes e recomendar tratamento, mas a visão de um médico especialista é indispensável para a decisão final. A transparência desse sistema é vital para a confiança em suas previsões e essa é fundamental para o engajamento do usuário.

Segundo Ferreira e Monteiro (2020), para o comportamento da XAI de um sistema ser adequado às necessidades e expectativas dos usuários, no momento e formato correto, deve-se considerar quem irá receber a explicação, o porquê de a explicação ser necessária e o contexto no qual a explicação será apresentada. É fundamental aplicar os conhecimentos de UX Design para se projetar a experiência dos usuários com os sistemas com XAI, com foco nas necessidades das pessoas e no entendimento de como a AI pode impactar das decisões humanas.

Samek, Wiegand e Müller (2017) apontam que um aspecto importante da inteligência humana é nossa capacidade de explicar a lógica de nossas decisões para outras pessoas. Segundo os autores, explicar suas decisões é um fator importante nas interações sociais e, muitas vezes, um pré-requisito para estabelecer uma relação de confiança entre as pessoas, como um médico explicar a decisão do tratamento ao seu paciente.

No campo da interação humano-computador, podemos dizer que falta transparência no sistema de ML, quando não fica claro para o usuário quais dados de entrada influenciam nas decisões e consequente saídas da máquina. Ribera et al. (2019) defendem colocar explicações no contexto do usuário para deixá-la comprehensível. A UNESCO (2021) descreve o princípio da transparência e explicabilidade como “essenciais para garantir o respeito, a proteção e a promoção dos direitos humanos, liberdades fundamentais e princípios éticos” (UNESCO, 2021, tradução nossa¹). Segundo ela, a transparência visa fornecer informação adequada aos respetivos destinatários para permitir a sua compreensão e fomentar a confiança. Já a explicabilidade refere-se a tornar comprehensível as entradas e saídas do sistema, assim como todo o seu funcionamento. Dessa maneira, a explicabilidade está diretamente relacionada à transparência, pois os resultados devem ser comprehensíveis e rastreáveis, adequados ao contexto.

Outra questão que favorece a XAI está relacionada ao nível de autonomia. Os sistemas baseados em ML tomam decisões e apresentam opções aos usuários, para ajudar em suas decisões. Le et al. (2019) alertam que há uma preocupação crescente sobre até que ponto a personalização feita pelo algoritmo promove uma exposição seletiva e limita o acesso das pessoas a pontos de vista diferentes, aprisionando-as em um filtro-bolha. Como indicado por Oxborough et al. (2018), a XAI pode aumentar o nível de controle dos usuários, pois eles não vão se limitar a escolher dentre as opções apresentadas. A partir do momento que entendem a lógica das predições, podem questionar as opções da máquina e através de novas entradas, alcançar maior acurácia.

Embora haja progresso nos últimos anos em relação a XAI, Jin et al. (2019) argumentam que as abordagens de XAI dos últimos anos são desenvolvidas para que especialistas em AI visualizem, entendam, depurem ou melhorem os modelos de AI, ignorando os principais usuários, os usuários finais. Witz et al. (2020) também indicam que as abordagens XAI recentes não atendem às necessidades do usuário final.

3 O sistema de recomendação das plataformas de VoD

As plataformas de Vídeo sob Demanda (VoD) oferecem conteúdo audiovisual para serem consumidos via streaming, utilizando a Internet. Diferente da TV aberta, as plataformas de VoD, como a Netflix e a Amazon Prime, dão aos seus usuários opções de escolha do que

¹ Texto original “(...) essential preconditions to ensure the respect, protection and promotion of human rights, fundamental freedoms and ethical principles.”

assistir, quando e onde, pois não precisam se limitar a programação transmitida pelos canais. Segundo Gomez-Uribe e Hunt (2015), os humanos têm dificuldades em escolher quando são apresentadas muitas opções, pois ficam rapidamente sobrecarregados e acabam não escolhendo nada ou fazendo escolhas ruins.

A grande inovação trazida por essas plataformas é o sistema de recomendação, que tem o objetivo de selecionar e classificar vídeos, para oferecer uma lista de recomendações de maneira personalizada a cada usuário. Sua proposta é agilizar, facilitar e orientar a escolha de cada usuário, de acordo com seu perfil, em meio ao grande volume de conteúdo disponível pelas plataformas. O sistema de recomendação das plataformas é composto por vários algoritmos diferentes, que atendem a diferentes casos de uso e se completam para criar a experiência única. São os algoritmos de ML que possibilitam a personalização, com base na análise de dados, seguindo abordagens supervisionadas e não supervisionadas. Os dados analisados podem ser fornecidos diretamente pelos usuários, como *feedback* de classificação, ou indiretamente, como que conteúdo foi assistido e quando, que recomendações não foram acessadas etc. As plataformas de VoD combinam personalização – que podem ser desde os conteúdos apresentados, até títulos das categorias, ordenação e imagens – com popularidade. Além disso, as categorias combinam relevância e diversidade, podendo obedecer a regras diferentes, já que algumas são intituladas por gêneros e outras consideram outros fatores para apresentar o grupo de conteúdos recomendados.

O algoritmo de ML se propõe a conhecer o usuário, aprender com ele e fazer previsões cada vez mais certeiras do tipo de conteúdo que ele escolheria, baseado em análise de dados de comportamento e comparações. Em teoria, essa operação parece ideal para um design centrado no ser humano, pois economiza esforço cognitivo do usuário e tempo na tomada de decisões, apresentando uma navegação personalizada para cada indivíduo. Seria como uma tentativa de resolver o problema da sobrecarga de informação e do paradoxo da escolha. Konstan e Riedl (2012) indicam que um ponto forte dos sistemas de recomendação é o fato de reduzirem a carga de trabalho dos usuários frente a tantas opções disponíveis. No entanto, afirmam que os usuários geralmente ficam mais satisfeitos quando estão no controle do sistema, mesmo que implique mais esforço. Alguns problemas já são estudados por pesquisadores, como: a formação de bolhas, perda de controle, modelo mental fraco, falta de transparência, entre outros.

Os sistemas de recomendação das plataformas de VoD utilizam alguns tipos de filtros como o baseado em conteúdo e o colaborativo. O filtro baseado em conteúdo compara um novo conteúdo com as preferências do usuário, para recomendá-lo. Já o colaborativo, identifica afinidade entre seus usuários, avaliando os registros de seus interesses e comparando as similaridades. Existem ainda as recomendações híbridas, que combinam duas ou mais técnicas para superar os limites individuais combinando as estratégias, como relatam Samih, Adadi e Berrada (2019).

Para manter o engajamento dos usuários, o sistema de recomendação precisa ser certo. Para cumprir esse objetivo, ele deve selecionar conteúdo relevante para o usuário, além de apresentar novidades, ou seja, conteúdo ainda não assistido pelo usuário, mas aderente ao seu interesse. Outro fator importante é o ineditismo, com conteúdos que o surpreendam positivamente. Ocorre quando são conteúdos não óbvios de acordo com o perfil dele, mas que capturem seu interesse. Por fim, a diversidade nos tipos de conteúdos apresentados aumenta as chances de algum ser escolhido e evita o tédio de uma seleção com itens muito semelhantes uns dos outros.

Tintarev e Masthoff (2015) apontam alguns fatores que aumentam a confiança dos usuários em um sistema de recomendação: precisão do algoritmo, transparência e a possibilidade de interação com sistemas de recomendação. Os autores comentam que a transparência não necessariamente precisa ser uma justificativa da recomendação. Às vezes, uma explicação mais detalhada do item recomendado, já seja suficiente para o usuário entender sua lógica. Segundo eles, essas explicações melhoram a confiança e a lealdade do usuário, consequentemente, aumentando a satisfação, já que agilizam e facilitam a escolha.

Embora os sistemas de ML pareçam autônomos ao tomar decisões, eles são feitos para humanos e por humanos. São decisões humanas que direcionam o processamento e mineração de dados, a seleção de metas de otimização e o diálogo projetado com os usuários finais com seus mecanismos de *feedback* implícitos e explícitos. Dessa maneira, são influenciados por humanos, desde as decisões de projeto até a interação. Segundo Cramer e Thorn (2017), as decisões humanas afetam os resultados dos sistemas de ML em praticamente todas as etapas do processo. Eles apontam ainda que o ecossistema que envolve qualquer sistema de aprendizado de máquina é um assunto completamente humano, seja conscientemente projetado ou não. Por isso, já há muita pesquisa recente investigando os problemas que ocorrem na interação do usuário com sistemas baseados em AI.

Além disso, Budiu (2018) alerta que a falta de transparência dos algoritmos atrapalha o entendimento dos usuários sobre o funcionamento do sistema, pois eles não sabem o quanto as recomendações são aleatórias ou influenciadas pela sua interação. Esse fato é agravado porque as ações do usuário, como assistir ou curtir um conteúdo, não impactam imediatamente no resultado das recomendações, já que o algoritmo depende de tempo e volume para aprender. Os sistemas de recomendação das plataformas de VoD baseados em ML são considerados caixas pretas, o que prejudica então a formação do modelo mental dos usuários sobre o funcionamento.

A AI não é uma novidade, mas encontrou um cenário muito propício para sua retomada e evolução, com o avanço tecnológico dos últimos anos e as oportunidades do mercado. Ao mesmo tempo que ela traz muito benefícios, as soluções de interação dos usuários com os sistemas baseados em AI ainda estão em processo de amadurecimento. Já são encontrados muitos estudos no sentido de projetar uma boa experiência de uso com os sistemas de AI, demonstrando que há preocupação do mercado e da academia em focar nas necessidades dos usuários.

4 Método

Para investigarmos as possibilidades de trabalhar a transparência do sistema de recomendação das plataformas de Vod, refletida nos títulos das categorias de conteúdos oferecidas aos usuários, foi planejado um workshop de design. O workshop de design é uma forma de design participativo que consolida formas de cocriação (MARTIN E HANINGTON, 2012). Através de uma coleta de dados qualitativa, busca-se envolver os usuários para aproveitar o conhecimento e experiência deles.

Com o objetivo de entender as expectativas dos usuários e aproveitar os insumos de suas experiências com as tarefas relacionadas ao problema de falta de transparência dos sistemas de recomendação, foram conduzidas quatro sessões de workshop de design com grupos de quatro a seis colaboradores cada, com o requisito de serem usuários de pelo menos uma plataforma de VoD baseada em ML, independente da experiência (tempo e frequência de uso). Para evitar o efeito do *design fixation* estudado por Jansson e Smith (1991), não foram

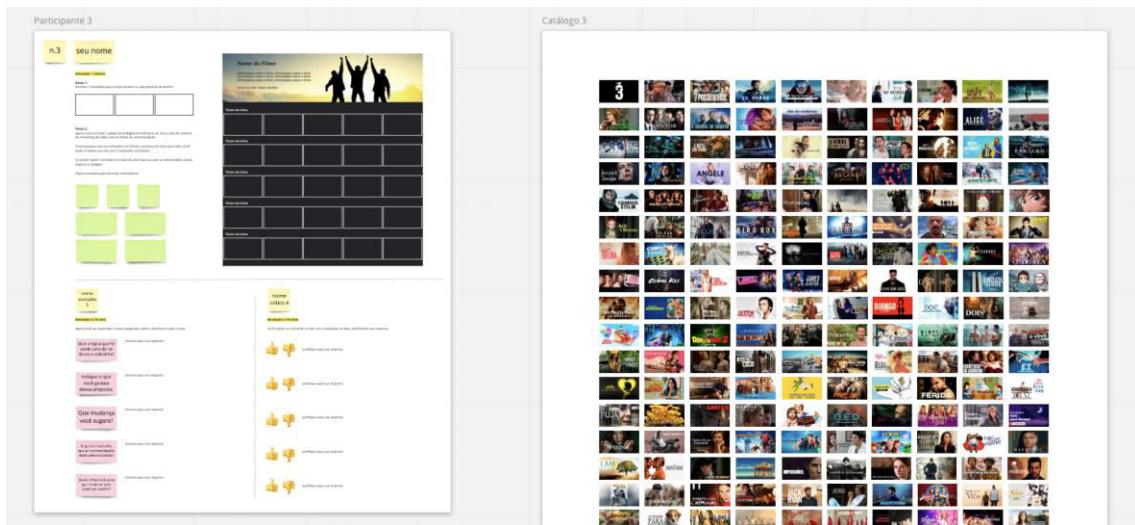
recrutadas pessoas que já trabalhassem com sistema de recomendação baseado em AI/ML. Os autores descrevem o *design fixation* como uma barreira no processo de design conceitual, a qual limitaria os designers a um conjunto de ideias ou conceitos já conhecidos, sendo contraproducente para a geração de novas ideias e resolução de problemas.

Os participantes foram recrutados por um formulário divulgado a voluntários de estudos anteriores e em plataformas de mídias sociais. Foi utilizado o método bola de neve para ampliar a amostragem. Os grupos foram formados de acordo com a disponibilidade dos participantes, buscando envolver usuário de plataformas de VoD e designers.

Cada sessão foi planejada para ser realizada em um único encontro virtual, com duração de aproximadamente duas horas. Para gravação, foi utilizada a funcionalidade já existente no software de reunião remota Zoom (www.zoom.us). Também, foi criado um *board* no Miro (www.miro.com), preparado com as imagens e espaços necessários para a execução da dinâmica da sessão de cocriação (Figura 1), conduzida de acordo com um roteiro:

1. Recepção – 10 minutos: Os participantes acessavam o link do *board* no Miro e eram recepcionados em um quadro com uma mensagem de boas-vindas, composta por agradecimento, orientações sobre o funcionamento da dinâmica e conceitos sobre interação e transparência em sistema de recomendação com AI/ML.
2. Introdução – 5 minutos: Era feita a leitura e explicação do conteúdo do quadro.
3. Criação individual – 20 minutos: O participante assumia seu primeiro papel como criador. Cada um deveria escolher três imagens de conteúdo disponíveis no catálogo do *board*. Era necessário que tais imagens representassem seu gosto individual. Em seguida, deveria fazer a função da AI e popular os espaços da simulação de tela entregue a ele, nomeando as linhas de recomendação. O resultado seria uma tela como gostariam que a AI montasse para eles.

Figura 1 – Captura de tela exibindo o *board* no Miro utilizado pelos participantes.



Fonte: As autoras.

4. Avaliação – 10 minutos: O participante assumia seu segundo papel, o de avaliador. Cada participante deveria direcionar-se para a tela criada por outro participante, indicado pela moderadora. Deveria então, avaliar a proposta de outro participante, guiado por perguntas escritas no *board*.

5. Crítica a avaliação – 10 minutos: O participante assumia seu terceiro e último papel, o de crítico da avaliação. Cada participante deveria, mais uma vez, movimentar-se para outro quadro, indicado pela moderadora. Deveria opinar sobre a avaliação feita por outro participante, referente a proposta de tela.
6. Exposição – 20 a 30 minutos: Cada participante explicava oralmente sua crítica à avaliação da tela proposta. Os pontos eram debatidos por todos os participantes.
7. Discussão – 15 a 20 minutos: Algumas questões relacionadas à transparência eram trazidas para que os participantes discutissem sobre elas.
8. Finalização – 5 minutos: Agradecimento.

Após o encerramento da dinâmica, todo o material produzido foi coletado e analisado. As observações referentes às propostas de telas, avaliações e comentários dos participantes foram organizados em um *board* do Miro e então realizada uma análise *bottom-up*, através de um diagrama de afinidades (Martin e Hanington, 2012).

5 Resultados

Foram mapeadas descobertas referentes ao modelo dos participantes, manifestadas em suas propostas, críticas e comentários. Notam-se características comuns e divergentes entre alguns fatores considerados para a formação dos títulos das categorias de conteúdos, como: a influência do histórico ou fatores externos, gêneros, tipos, personalização e transparência. Além das categorias, os participantes abordaram também a organização do conteúdo por filme ou série e suas repetições em diferentes categorias. Foram discutidas ainda ideias sobre funcionalidades e necessidades, como a integração com redes sociais e configurações. Também surgiram comentários sobre como fazer escolhas, formação de bolha e confiança nas recomendações. As descobertas são a seguir apresentadas:

As pessoas entendem e esperam que as recomendações sejam feitas de acordo com o seu perfil, influenciadas pelo seu comportamento de uso. Quase todos os participantes desenvolvem suas propostas deixando evidente a influência do histórico na maioria das categorias criadas. Isso demonstra que já aprenderam que o sistema de recomendação é personalizado, diretamente influenciado pelo comportamento de uso, além disso, gostam desse funcionamento. Entretanto, alguns participantes têm dificuldades em perceber a personalização quando o título da categoria não é explícito, como “porque você gostou de...”. Conclui-se que por mais que o algoritmo se baseie em histórico, o título da categoria pode não ser transparente o suficiente para o usuário entender a sua lógica.

Quando as categorias são subjetivas e/ou abrangentes, as pessoas têm dificuldades em reconhecer a personalização. Muitos participantes criaram categorias subjetivas e/ou abrangentes, como “você também vai gostar”, “Diretores aclamados”, “Tempo e Espaço”, “Filmes Reflexivos: discussões profundas e temas sociais” etc. A maioria fez a proposta combinando e variando tipos de categoria, provavelmente influenciados como a Netflix é atualmente. Comparando-se todas as propostas e suas avaliações, nota-se que os participantes, no papel de avaliadores e críticos, não reconhecem a personalização quando o título da categoria não se refere diretamente ao motivo da recomendação, ou seja, não reflete o princípio da transparência. Além disso, os avaliadores divergem se categorias abrangentes ajudam na escolha, por atribuir autonomia ao usuário, ou se atrapalham, por deixá-lo perdido. Um participante relata: “Os diferentes tipos de conteúdo que podem ser encontrados em uma categoria só. Por exemplo: Filmes premiados – você encontra drama, ação, animação em um lugar só. Acho que ajuda o usuário que não sabe exatamente o que assistir e pode avaliar o seu

humor enquanto vai passando pelos filmes.” Aqui percebe-se uma contradição, pois a proposta do sistema de recomendação é exatamente filtrar o conteúdo para facilitar a escolha do usuário. Outros participantes defendem que se perdem em categorias abrangentes.

A categorização baseada em gênero cinematográfico ainda é muito forte no modelo mental das pessoas para a escolha de conteúdo. A maioria dos participantes fez sua proposta com pelo menos alguma categoria influenciada pelos gêneros tradicionais, de produções cinematográficas. Para exemplificar, alguns propuseram categorias com os gêneros, como “Documentário”, “Policial” e “Comédia Romântica”, outros combinaram com outros fatores, como “Dramas Femininos”, “Comédias *Feel Good*” e “Séries intrigantes com pitadas de comédia”. Dessa maneira, argumentam que gênero facilita a escolha pois já estamos acostumados. O problema desse tipo de categoria é que os avaliadores não reconheceram a personalização, mesmo quando havia conteúdo do mesmo gênero na escolha do participante para representar o seu gosto.

As categorias podem evidenciar a personalização, mas variar o nível de transparência. São propostas pelos participantes categorias relacionadas ao histórico de maneira mais ampla, como “você também vai gostar”, ou bem focadas em um conteúdo assistido, como “porque você assistiu a...”. Apenas um participante fez sua proposta com todas as categorias bem específicas em um atributo, por exemplo: Porque você assistiu/curtiu “nome de filme ou série que assistiu recentemente”. Um participante comenta que mais do que “porque você assistiu”, prefere “porque você curtiu”. Dessa maneira, aumenta a chance de o sistema recomendar conteúdo com afinidade com o usuário, mas é dependente de seu *feedback* ativo.

As categorias claramente influenciadas pelos outros usuários não atendem ao princípio de transparência. Os participantes não entendem as regras de categorias como “Em Alta” e “Top10”. Alguns comentam que gostariam que essas categorias fossem personalizadas, inferindo que não são. Além disso, surgem comentários sobre o período não ser identificado e sobre o que seria estar em alta, “para quem?”. Adicionalmente, um participante comenta que a categoria “Em Alta” causa FOMO – do inglês, *Fear of Missing Out*.

A tela ideal dosa os tipos de categorias que a compõe. A grande maioria dos participantes gosta de compor a tela com variedade no nível de personalização, além de combinar com outros fatores. Alguns trazem as categorias que não são recomendação, como “Continuar Assistindo” e “Minha Lista” (seleção do usuário). Aparecem também as propostas com a categoria de “Lançamentos” e categorias claramente influenciadas por outros usuários, como “Em Alta” e “Top10”. Alguns participantes sugerem ainda que informações da ficha dos filmes aparecem no título de categorias, como ator, diretor, ano de produção, origem e indicação ao Oscar, rotulando a lógica do agrupamento.

A tela ideal apresenta as categorias e o conteúdo de acordo com o perfil do usuário. Alguns participantes comentam que as categorias deveriam aparecer de acordo com o gosto do usuário. Um comenta: “o algoritmo deveria considerar as diferenças de como as pessoas querem suas telas”. Isso quer dizer que, além das linhas de recomendação baseadas no perfil, as categorias de “Lançamentos”, “Em Alta” e “Top10”, por exemplo, só deveriam aparecer se o usuário faz suas escolhas com elas. Os agrupamentos de conteúdo e títulos também atenderiam ao gosto do usuário. Alguns comentam que não basta o título da categoria acertar o gosto do usuário, os conteúdos oferecidos nela também precisam ter afinidade com o perfil. Por fim, dois participantes comentam que a ordem em que as categorias são dispostas é muito importante.

Além do perfil do usuário, o sistema de recomendação poderia considerar o momento da escolha. Alguns participantes propõem categorias que referenciam ânimo do usuário “Para passar o tempo”, pelo que ele quer sentir “Emocionante”, o tipo de companhia “Para assistir com amigos” ou até o tempo que tem disponível “Para maratonar”. Vale ressaltar que quando não está evidente no título, as pessoas podem não perceber a personalização. É o que acontece com os avaliadores em grande parte dessas categorias. Às vezes, até consideram que nenhum tipo de personalização seria necessário, pois abrangeriam momentos em que qualquer usuário poderia se encaixar.

A transparência só é alcançada quando a categoria é focada em um conteúdo do histórico. Poucos participantes conseguiram inovar criando um título de categoria transparente. Muitos usaram o padrão conhecido de “Porque você assistiu a determinado conteúdo”, mas surgiram categorias combinando outros fatores como “Novidades similares a nome de filme ou série que assisti recentemente”, quando o participante queria bastante personalização. Não identificamos nenhuma proposta que considerasse o histórico do usuário de forma geral e fizesse combinações mais amplas para recomendar uma categoria com título transparente. Percebe-se pelas propostas e discussões que a maneira que os participantes criaram as telas personalizadas para elas, não apresentavam transparência suficiente para os outros perceberem. Mesmo que cada pessoa tenha criado a seu gosto, os avaliadores não entendem como personalizado sempre que a categoria é abrangente, pelo fato de supostamente servir para um número grande de pessoas e a falta de transparência nos títulos.

A falta de transparência das categorias impacta na confiança que os usuários têm nas recomendações. Muitos participantes comentam que desconfiam da recomendação quando o conteúdo recomendado não é do gosto deles, até mesmo na categoria de “Top10”. Como as pessoas não entendem o motivo de um conteúdo fora do perfil delas ser recomendado, intuem que haja um interesse promocional da plataforma de VoD, até mesmo como manipulação do usuário. Alguns participantes encaram apenas como erro do sistema quando recebe uma recomendação fora do seu perfil. Já outros questionam a veracidade do “Top10”, a lógica do “Em Alta” e do índice de relevância da Netflix. Por fim, um participante comenta ainda que a AI ajuda a escolha, mas tira a autonomia do usuário.

As plataformas de VoD não explicam para o usuário como o sistema de avaliação impacta no resultado das recomendações. Os participantes divergem sobre qual seria o melhor modelo de avaliação, se um binário como o “*like*” ou um que permita a gradação como “5 estrelas”. Os participantes intuem que de alguma forma sua avaliação refletirá no funcionamento do algoritmo, mas não sabem o quanto. Um comenta que não sabe como usar a avaliação e questiona sobre o que é considerado na avaliação para suas recomendações, se o voto de todos os usuários ou se apenas o dele. Já outro afirma que gosta de avaliar para que o conteúdo seja mais recomendado para outros usuários, inferindo a influência entre usuários.

Dar mais controle ao usuário pode melhorar a experiência com o sistema de recomendação. Muitos participantes comentam que gostariam de interferir na recomendação, seja através de filtros, sistema de “tagueamento” ou “ranquear” suas preferências. Os participantes sugerem funcionalidade para usuário indicar se não concorda que o conteúdo faça parte de uma categoria, excluir conteúdo e categoria recomendados dos quais não gostam. Um participante exalta que a personalização precisa ser controlada pelo usuário: “Quero algo muito a ver com o que eu amo, ser customizável”. Outro participante comenta sobre a relação da AI com o usuário: “A AI deixa o usuário de lado. Grande erro achar que sabe o que eu gosto, sem me perguntar.” Outro ainda fala sobre os projetistas: “Quem projeta, coloca a sua visão de

mundo. Será que estão sendo diversos?”. Por fim, um participante faz um comentário mais extremo “Eu nunca gosto do que me recomenda. Prefiro categorias para não me influenciar. Prefiro me sentir no comando.” Todos esses comentários compõem a oportunidade dos sistemas de VoD permitirem mais interação do usuário com o sistema de recomendação, para facilitar o aprendizado do algoritmo e consequentemente, melhorar a experiência de uso. A possibilidade de o usuário controlar o sistema, pode ser uma boa maneira de facilitar seu entendimento.

A funcionalidade de configuração do sistema pelo usuário é uma oportunidade não explorada pelas plataformas de VoD. Muitos participantes manifestam o desejo de configurar o sistema de alguma forma, para tornar as recomendações dos algoritmos mais certeiras com seu perfil e agilizar o aprendizado do ML, evitando ainda o *Cold Start*. Um participante comenta “Parei de assistir porque as recomendações pioraram. Eu gostaria de configurar”. Alguns sugerem indicar o que querem e o que não querem por um formulário com múltiplas escolhas. Um participante gostaria de indicar alguns títulos de conteúdo que gosta, não categorias; outro sugere um filtro de ano de produção; enquanto outro gostaria de marcar o que já assistiu fora da plataforma, para ensinar seu gosto e não receber a recomendação de conteúdo já assistido. Alguns participantes discutem que nem todos os usuários querem perder esse tempo inicial com configuração, pois estariam ansiosos para explorar a plataforma de VoD. Para solucionar esses casos, surgem soluções como a configuração inicial ser opcional, ficar disponível para o usuário acessar a qualquer momento ou ser ao longo da experiência, como a possibilidade de o usuário editar diretamente a home. Referente a isso, um participante comenta “Por que não podemos configurar? Como eu quero que a AI funcione para mim?”. Apesar de muitos participantes concordarem com esse posicionamento, não é uma unanimidade. Alguns dizem que ainda preferem aguardar que o aprendizado normal do algoritmo, ou até mesmo que não gostam de configurar, mas o aprendizado poderia ser mais rápido, como ocorre com os streamings de música. O problema é que o aprendizado exige tempo e volume. Naturalmente os streamings de música têm a possibilidade de coletar mais dados do comportamento do usuário ao longo dos dias que um streaming de vídeo, devido ao tempo de seus conteúdos.

Embora o foco do workshop fosse transparência, outras questões relevantes e referentes à experiência com as plataformas de VoD foram trazidas pelos participantes:

A mistura de filmes e séries, muitas vezes, é um inconveniente. Com frequência, os participantes evidenciam nas categorias a separação do conteúdo de filme e série, ou comentam o incômodo em encontrar ambos os tipos em uma mesma categoria. Como solução, propõem separar as categorias de filme e séries ou aplicar alguma identificação na imagem de capa.

O compartilhamento de recomendação de conteúdo entre usuários é uma oportunidade ainda não explorada pelas plataformas de VoD. Surge o comentário de que boa parte dos usuários recorrem a meios externos para decidir o que assistir, exemplo: blogs e páginas de críticos. Alguns participantes discutem que gostariam de compartilhar comentários e avaliações de conteúdo com outras pessoas. “Amigos sabem do que eu gosto, e se eu confio no gosto do amigo, eu quero ver o que ele gostou”. Propõem funcionalidade para avaliar a recomendação dos amigos e bloqueá-la, ou até mesmo ter conexão com as redes sociais, como faz por exemplo, o streaming de áudio Deezer. Além disso, um participante sugere uma categoria com histórico dos assistidos, para servir de consulta e dicas para sua rede.

É positiva a dose de novidade recomendada ao usuário para evitar a formação de bolha. Alguns participantes citam o problema da formação de bolha, caso a recomendação se prenda exclusivamente ao histórico do usuário. Eles comentam que gostam de receber recomendação de conteúdo diferente do habitual e desconhecido, para ampliar as suas escolhas.

As plataformas de VoD atualmente não praticam a AI explicável. Poucas categorias apresentam transparência suficiente ao ponto de os usuários entenderem a lógica da recomendação apresentada a eles. Este workshop de design se propôs a explorar a transparência das recomendações feitas ao usuário. Como achados, obteve diversos pontos de oportunidade para melhorar a experiência de uso com as plataformas de VoD. Apesar de levantar boas ideias para melhorar a transparência do sistema e soluções para alguns problemas discutidos, nenhuma proposta dos participantes trouxe uma solução inovadora para o problema da falta de transparência. Mesmo não envolvendo profissionais atuantes no desenvolvimento de sistema de recomendação baseado em AI/ML para evitar o *design fixation*, os participantes já estão acostumados como as plataformas de VoD se apresentam hoje. É possível que esse fator tenha influenciado as soluções propostas por eles. Apesar disso, foi efetivo para evidenciar os problemas causados na experiência de uso ocasionados pela falta de transparência, além de indicar fatores a serem considerados pelos algoritmos de recomendação e funcionalidades desejadas para melhorar a experiência do usuário.

6 Discussão

Nossa pesquisa concluiu que os usuários entendem e até já esperam que as recomendações de conteúdo sejam feitas de maneira personalizada, influenciadas pelo seu comportamento de uso. Apesar disso, não sabem o quanto a plataforma é personalizada, no sentido de que não identificam os fatores que impactam em cada categoria recomendada. Além do mais, como apontado por Gilpin et al. (2019) e identificado pela nossa pesquisa, a falta de transparência implica na redução de confiança dos usuários, em relação ao que lhe é recomendado pelo sistema. Defendemos que a falta de transparência das plataformas de VoD é um dos principais problemas a ser solucionado para melhorar a experiência de seus usuários.

Mesmo sem entender o funcionamento do sistema, muitos usuários gostam de receber recomendações. O problema acontece quando os conteúdos recomendados não são do interesse deles. Isso pode ocorrer no início, quando o algoritmo ainda não teve tempo e volume para aprender o suficiente sobre o perfil do usuário, o chamado *cold start*. Também pode acontecer a qualquer momento durante a utilização do sistema, prejudicando a experiência do usuário com a plataforma de VoD, e consequentemente, o engajamento. Uma solução discutida no workshop seria a possibilidade de o usuário poder configurar o sistema, desde seu primeiro contato, sem obrigatoriedade, ficando disponível a qualquer momento. Dessa maneira, atenderia às pessoas que querem investir mais tempo na configuração, para terem uma recomendação certeira ou as que preferem aguardar e toleram o erro do sistema durante o aprendizado. Nesse sentido, Tory (2017) comenta sobre os diferentes de perfis das pessoas, umas que preferem configurar frente às que toleram o erro.

Além disso, surgiram boas propostas em nosso workshop, como a possibilidade de avaliar conteúdos não presentes na plataforma, para ensinar o gosto do usuário. Atualmente, por não entenderem as regras do sistema, os usuários não sabem como ajudar o sistema de recomendação. Mesmo a funcionalidade de *feedback* ativo, o “*like*”, não é explicada aos usuários das plataformas de VoD. Muitos intuem que indicando seu gosto, influenciará de

alguma maneira as recomendações, mas não sabem o quanto, em quais categorias, ou se impactará também na recomendação para outros usuários. Defendemos que dar mais controle aos usuários poderia melhorar a experiência de uso das plataformas. Entretanto, o controle não pode ser obrigatório, mas sim opcional, para atender às necessidades dos usuários que preferem investir seu tempo, para acelerar o processo de aprendizagem dos algoritmos.

Quando os usuários fazem uso das recomendações, há diversas informações sobre os conteúdos fornecidas pelas plataformas, que o ajudam na escolha, como sinopse, elenco, capa etc. Ferreira et al. (2020), indicam que informações relacionadas à estrutura narrativa, como número de episódios/temporadas, se são independentes ou sequenciais etc., ajudam os usuários a decidir se estão dispostos a assistir determinado conteúdo naquele momento. Um inconveniente muito comentado pelos participantes da nossa pesquisa é a falta de separação ou indicação se o conteúdo seria filme ou série. Poderia ser resolvido com uma *tag* na capa, que já evitaria o erro do usuário e auxiliaria na escolha quando o fator tempo é importante para a decisão. As categorias nas quais os conteúdos são distribuídos também direcionam os usuários para facilitar sua escolha. O gênero cinematográfico ainda é muito forte no modelo mental das pessoas, que esperam estar refletidos nos títulos das categorias. Como conclusão de nosso workshop, a tela ideal é composta por diferentes tipos de categorias, variando os fatores de influência para a formação dessas categorias. Um relato bastante comum dos participantes envolvidos nesta pesquisa, foi a influência do ânimo no momento da escolha, fator desconsiderado pelas plataformas. Seria mais uma oportunidade para as plataformas explorarem.

Percebemos em nossa pesquisa, que os usuários apresentam comportamentos diferentes para escolher conteúdos. Se os usuários têm expectativas e necessidades diferentes entre si, a tela ideal deveria apresentar as categorias e o conteúdo de acordo com o perfil do usuário. Isso significa ampliar a personalização que é oferecida hoje. Por exemplo, se hoje a Netflix apresenta algumas categorias que são comuns a todos, como “top10”, ela deveria apresentá-la apenas aos que gostam dela. Poderia ser possível a inclusão e exclusão de categorias pelo controle do usuário. Alguns demonstram preferências em relação à especificidade dos títulos das categorias, ordenação das categorias e combinação de fatores, por exemplo, que deveriam ter suas particularidades respeitadas para melhorar a experiência. Há um grande campo a ser explorado para um trabalho de *UX Writing* direcionado aos títulos das categorias.

7 Conclusão

Embora o princípio da transparência seja recomendado em diversas listas de princípios de AI, somando-se a todas as publicações que defendem o uso de XAI com foco no usuário final, as plataformas VoD ainda não empregam bem a explicabilidade de suas recomendações. Constatamos que a maioria dos participantes envolvidos está ciente de que existe um mecanismo automático, mas não está claro como funciona e o nível de personalização para aproveitar ao máximo sua capacidade. Trabalhar a transparência da AI com a explicabilidade do sistema é essencial para ganhar confiança e melhorar a experiência dos usuários com o sistema de recomendação. Isso porque, atualmente, os usuários não estão no controle, já que se não entendem o funcionamento do sistema, não forma um modelo mental adequado e não sabem como influenciá-lo. Em nosso workshop de design, levantamos boas ideias que melhoraram a transparência, mas não alcançamos uma solução final inovadora.

Além da transparência, uma opção de configuração poderia ser implementada para aumentar o controle do usuário. Em nosso workshop de design, surgem algumas alternativas de configuração, mas para todas é importante respeitar a vontade do usuário. Ele deve escolher quando quer usar e o quanto de seu tempo quer investir para fornecer dados que auxiliem os algoritmos. Além do mais, essa seria uma maneira de também melhorar o entendimento que o usuário tem sobre o funcionamento do sistema.

Outras oportunidades ainda não exploradas pelas plataformas foram identificadas para melhorias a experiência de uso: a possibilidade de troca entre usuários através da plataforma e a consideração do fator ânimo pelo mecanismo de recomendação. Foi muito comentado pelos participantes de nossa pesquisa a valorização de dicas de conteúdo para assistir, com outros usuários de suas redes de relacionamento, além de pesquisas na internet. Os usuários poderiam manter esse costume sem ter que sair da plataforma, compartilhando recomendações com outros usuários. O fato de o ânimo influenciar na escolha do conteúdo a assistir no momento também apareceu bastante. Para isso, seria necessário apresentar opções considerando diferentes ânimos.

Por fim, todas essas questões deveriam respeitar as particularidades de cada usuário, através de uma personalização além da praticada atualmente. As plataformas precisam considerar além do que exibir, também como exibir de acordo com o perfil. Apresentar apenas categorias que tenham afinidade com o usuário, títulos e combinações seguindo o gosto do usuário; ordenação e qualquer característica da tela, que atendam as necessidades específicas individuais.

Como nossa pesquisa envolveu participantes brasileiros, seria interessante um estudo transcultural para conferir se descobertas semelhantes ocorreriam entre usuários de outros países ou se haveria diferenças culturais. A amostra representa uma limitação, portanto, não podemos generalizar. Ela retrata também um período específico, então, diversas evoluções podem ocorrer nas plataformas aqui citadas pelos participantes.

Esta pesquisa apontou diversas recomendações referente aos problemas encontrados, mas levantou ainda muitos questionamentos que merecem ser explorados. Em primeiro lugar, gostaríamos de explorar mais as soluções de transparência para os sistemas de recomendação das plataformas de VoD, para identificar se tornaria evidente o nível de personalização e possibilidades de controle. Da mesma forma, demonstrar as diferenças de comportamento e possibilidade de variações que atendessem às diferentes necessidades.

Agradecimentos

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

8 Referências

BUDIU, R. Can Users Control and Understand a UI Driven by Machine Learning? 2018. Disponível em: <<https://www.nngroup.com/articles/machine-learning-ux/>>. Acesso em: 22 nov. 2020.

CRAMER, H.; THORN, J. Not-So-Autonomous, Very Human Decisions in Machine Learning: Questions when Designing for ML. **AAAI Spring Symposium – Technical Report**, 2017. V. SS-17-01-, p. 412–414.

FERREIRA, F. et al. **Investigating Multimodal Features for Video Recommendations at Globoplay**. Fourteenth ACM Conference on Recommender Systems. **Anais...** New York, NY, USA: ACM, 22 set. 2020.

FERREIRA, J. J.; MONTEIRO, M. S. What Are People Doing About XAI User Experience? A Survey on AI Explainability Research and Practice. In: MARCUS, A.; ROSENZWEIG, E. (Org.). **9th International Conference on Design, User Experience, and Usability (DUXU 2020)**. Cham: Springer, 2020, p. 56–73.

FLI FUTURE OF LIFE INSTITUTE. Asilomar AI Principles. 2017. Disponível em: <<https://futureoflife.org/ai-principles/>>. Acesso em: 7 maio 2020.

GILPIN, L. H. et al. Explaining explanations: An overview of interpretability of machine learning. **Proceedings – 2018 IEEE 5th International Conference on Data Science and Advanced Analytics, DSAA 2018**, 2019. P. 80–89.

GOMEZ-URIIBE, C. A.; HUNT, N. The Netflix Recommender System: Algorithms, Business Value, and Innovation. **ACM Transactions on Management Information Systems**, v. 6, n. 4, 2015.

GOOGLE PAIR. People + AI Guidebook. 2019. Disponível em: <pair.withgoogle.com/guidebook>. Acesso em: 7 maio 2020.

JANSSON, D. G.; SMITH, S. M. Design fixation. **Design Studies**, v. 12, n. 1, p. 3–11, jan. 1991.

JIN, W. et al. Bridging ai developers and end users: An end-user-centred explainable ai taxonomy and visual vocabularies. **Proceedings of the IEEE Visualization, Vancouver, BC, Canada**, 2019. p. 20-25.

KONSTAN, J. A.; RIEDL, J. Recommender Systems: From Algorithms to User Experience. **User Modeling and User-Adapted Interaction**, 2012. V. 22, n. 1–2, p. 101–123.

LAMY, J. B. et al. Explainable artificial intelligence for breast cancer: A visual case-based reasoning approach. **Artificial Intelligence in Medicine**, 2019. V. 94, n. August 2018, p. 42–53.

LE, H. et al. **Measuring political personalization of Google news search**. The Web Conference 2019 - Proceedings of the World Wide Web Conference, WWW 2019. **Anais...** Association for Computing Machinery, Inc, 13 maio 2019.

MARTIN, B.; HANINGTON, B. **Universal Methods of Design: 100 Ways to Research Complex Problems, Develop Innovative Ideas, and Design Effective Solutions**. Beverly: Rockport Publishers, 2012.

MICROSOFT. Microsoft AI Principles. 2018. Disponível em: <<https://www.microsoft.com/en-us/ai/our-approach-to-ai>>. Acesso em: 7 maio 2020.

MINISTRY OF INTERNAL AFFAIRS AND COMMUNICATIONS (MIC), The G. Of J. AI R&D Principles. 2017.

OECD COUNCIL RECOMMENDATION ON ARTIFICIAL INTELLIGENCE. OECD Principles on AI. 2019. Disponível em: <<https://www.oecd.org/going-digital/ai/principles/>>. Acesso em: 7 maio 2020.

- OXBOROUGH, C. et al. **Explainable AI: Driving business value through greater understanding.** [S.I.]: [s.n.], 2018.
- RIBERA, M.; LAPEDRIZA, A. Can we do better explanations? A proposal of User-Centered Explainable AI. New York: ACM, 2019. p. 7.
- ROUHAINEN, L. **Artificial Intelligence: 101 Things You Must Know Today About Our Future.** [S.I.]: [s.n.], 2018.
- SAMEK, W.; WIEGAND, T.; MÜLLER, K.-R. Explainable Artificial Intelligence: Understanding, Visualizing and Interpreting Deep Learning Models. 2017.
- SAMIH, A.; ADADI, A.; BERRADA, M. Towards a knowledge based Explainable Recommender Systems. **ACM International Conference Proceeding Series**, 2019.
- TINTAREV, N.; MASTHOFF, J. Explaining Recommendations: Design and Evaluation. **Recommender Systems Handbook**. Boston, MA: Springer US, 2015, p. 353–382.
- WANG, Pei. What Do You Mean by “AI”? **Frontiers in Artificial Intelligence and Applications** 171(1):362-373, 2008. In: Conference: Proceedings of the 2008 conference on Artificial General Intelligence 2008: Proceedings of the First AGI Conference, 2008
- WEITZ, K. et al. “Let me explain!”: exploring the potential of virtual agents in explainable AI interaction design. **Journal on Multimodal User Interfaces**, 2020.
- XU, F. et al. Explainable AI: A Brief Survey on History, Research Areas, Approaches and Challenges. In: TANG, J. et al. (Org.). **Natural Language Processing and Chinese Computing: 8th CCF International Conference, NLPCC 2019 Dunhuang, China, October 9–14, 2019 Proceedings, Part II**. Dunhuang: Springer, 2019, p. 563–574.