Conforme indicado por [Sanneman et al.](https://workofthefuture.mit.edu/wp-content/uploads/2020/11/2020-Research-Brief-Sanneman-Fourie-Shah.pdf) num estudo realizado em grandes fabricantes e PMEs em 2020, a maior produtividade é observada em espaços de trabalho industriais que contam com uma combinação de trabalhadores humanos e robôs: a versatilidade dos humanos, somada à precisão e consistência dos robôs, tende a gerar um rendimento melhor do que mesmo em fábricas automatizadas. Como tal, algum nível de contacto humano com robôs industriais permanecerá uma característica inevitável dos espaços de trabalho industriais. Historicamente, os robôs eram segregados dos trabalhadores humanos durante a operação normal, de modo que a Interação Humano-Robô (IHR) efetivamente não ocorria, mas o surgimento dos cobots (robôs colaborativos) com a Indústria 4.0 levará a uma crescente necessidade de os humanos interagirem com eles, seja ensinando o que fazer, ou até mesmo compartilhando tarefas. Classicamente, a IHR tem sido categorizada, como proposto por [Schmidtler et al. [2015]](https://doi.org/10.3233/OER-150226), com base em quatro pontos de possível sobreposição: espaço de trabalho; tempo de trabalho; objetivo; e contacto. O resultado são três categorias de IHR, como mostrado na Figura 1. Na Coexistência Humano-Robô (HRCx), humanos e robôs compartilham um espaço de trabalho comum, mas não uma tarefa comum. A distinção entre a Cooperação Humano-Robô (HRCp) e a Colaboração Humano-Robô (HRC) é mais fluida e relaciona-se com a natureza das tarefas compartilhadas entre humano e robô, com a compreensão de que a HRCp não requer tanta comunicação entre os dois agentes. Cada categoria de IHR (ou até mesmo nenhuma delas) é mais adequada para tarefas diferentes, pelo que foi sugerido por [Sanneman et al.](https://workofthefuture.mit.edu/wp-content/uploads/2020/11/2020-Research-Brief-Sanneman-Fourie-Shah.pdf) que as fábricas industriais mais produtivas haverão de abranger todas as categorias.

Uma imagem com texto, captura de ecrã, diagrama, Tipo de letra

Descrição gerada automaticamente

Ilustração 4 - A categorização da Interação Humano-Robô (HRI) em HRCx, HRCp e HRC com base nos pontos de possível sobreposição presentes no tipo de interação.

**Fonte:** [Schmidtler et al. [2015]](https://doi.org/10.3233/OER-150226)

Com isso em mente, é interessante explorar a questão do quão conscientes os trabalhadores humanos estarão dos robôs industriais, abrangendo todas as categorias de interação. Nomeadamente, a consciência da presença de robôs industriais no local de trabalho foi determinada como apenas sendo uma subcategoria dentro do conceito amplo de "consciência": a Tabela 1 mostra a categorização da consciência de robôs em humanos que foi usada para orientar o texto abaixo. Foi procurado um entendimento do atual estado da arte e avanços conjeturais para cada subcategoria, com as subcategorias abordadas em ordem crescente do grau de interação que o trabalhador humano terá com o robô para o nível de consciência dado, com foco particular na HRC. A metodologia de pesquisa consistiu em três fases de busca por palavras-chave cada vez mais específicas, com os fundamentos sendo explorados por meio de referências cruzadas; e avanços verificados explorando os artigos que citaram contribuições influentes. Foi dada ênfase especial às limitações das normas de segurança atuais, bem como à motivação para o estudo da Teoria da Mente (ToM) artificial para aplicações industriais.

|  |  |  |
| --- | --- | --- |
| **Subcategorias** | **Descrição da Consciência** | **Componentes** |
| **Opinião** | Perceção concetual de robôs como um coletivo | Confiança, Afinidade, Conformidade |
| **Consciência Situacional** | Localização e movimento de robôs individuais | Atenção, Medidas de segurança passivas |
| **Compreensão de Comportamento** | Planos de robôs individuais | Cinemática, Olhar social |
| **Comunicação Humano-Robô** | Robot “mental state“"Estado mental" do robô | IHR cognitiva, Teoria da Mente (ToM) |

**Tabela 1:** Subcategorias de consciência de robôs em humanos, em ordem crescente de interação.

Para uma parte significativa da IHR, e especialmente para HRC, mais pesquisas foram conduzidas fora do contexto industrial. Por exemplo, existem muitos resultados relevantes obtidos no âmbito da robótica social. Deve-se notar, no entanto, que uma grande parte deles permanece relevante aqui, embora com algumas ressalvas. Notavelmente, os sinais de comunicação entre humanos e robôs industriais não são interpretadas tão fiavelmente quanto entre humanos e robôs sociais (como discutido com mais detalhes abaixo). Dois fatores afetam isto: em primeiro lugar, o ambiente é ruidoso e os rostos das pessoas estão obscurecidos com equipamentos de segurança; em segundo lugar, a escolha dos sensores e sua colocação são limitadas pela necessidade de não prejudicar o robô industrial no desempenho da sua tarefa.

Outro aspeto a considerar é a crescente importância na robótica industrial da garantia que os planos de um robô possam ser inferidos sem comunicar: prever de forma imprecisa o que um robô social fará pode ser irritante; enquanto calcular incorretamente as intenções de um robô industrial pode levar a acidentes fatais. Infelizmente, a compreensão humana da cinemática do robô é uma subcategoria da consciência de robôs em humanos na qual a tradução dos conhecimentos a partir de outras áreas de estudo para a robótica industrial não é tão direta. Robôs industriais tendem a possuir uma gama mais ampla de movimentos, e em particular movimentos que robôs humanoides são incapazes de realizar. Isto é verdade mesmo para robôs industriais não-móveis (Figura 2). Como tal, há pouca "polinização cruzada" que pode ocorrer neste tópico, como será evidente mais adiante, quando se abordar a compreensão humana do comportamento do robô.

Uma imagem com esboço, desenho, farol

Descrição gerada automaticamente

Ilustração 4 - Comparação das possíveis cinemáticas de um robô social com as de um robô industrial. **À Esquerda:** GrowMeUp, um robô social padrão. Sem braços, pode apenas mover o corpo. **À Direita:** Componentes básicos de um sistema padrão de robô industrial fixo. Manipulador com múltiplas articulações para maior liberdade de movimento.

**Fontes:** [Martins et al. [2015]](https://ap.isr.uc.pt/archive/610.pdf)e [OTM Sec. IV Ch. 4, Sec. II, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#basic-components)

A primeira subcategoria de consciência de robôs em humanos definida na Tabela 1 foi a opinião humana sobre robôs. Esta corresponde à consciência de robôs com a menor quantidade possível de interação: é uma impressão interna não-comunicada sobre robôs que nem requer a presença de qualquer robô individual para que exista. Pode corresponder a uma consciência apenas do conceito de robôs, em vez de robôs individuais corporificados. As duas facetas comummente estudadas são a confiança e a afinidade dos humanos em relação aos robôs.

As opiniões das pessoas sobre os robôs geralmente são avaliadas por meio de uma combinação de medidas comportamentais e declarações subjetivas voluntariamente oferecidas. Medidas comportamentais são comummente usadas para inferir a confiança implícita de uma pessoa em relação a um robô. Dependem de experiência para experiência, mas exemplos incluem medidas como tempo de decisão; proxémica (exemplo na Figura 3); propensão para delegar tarefas; e (para robôs humanoides) acompanhamento do olhar.

Para as declarações subjetivas, existe um conjunto de questionários que são tipicamente usados. Para uma experiência sobre a confiança humana em robôs sociais com personalidades diferentes, [Romeo et al.](https://doi.org/10.1109/RO-MAN53752.2022.9900550) propuseram em 2022 um conjunto de quatro questionários para avaliar a confiança humana em robôs: o questionário Disposition to Trust (DtoT) (devido a [Gefen [2000]](https://doi.org/10.1016/S0305-0483(00)00021-9), mede a propensão a confiar); o questionário Adapted Propensity to Trust in Technology (APTT) (devido a [Jessup [2018]](https://corescholar.libraries.wright.edu/etd_all/2209/?utm_source=corescholar.libraries.wright.edu%2Fetd_all%2F2209&utm_medium=PDF&utm_campaign=PDFCoverPages), adapta o questionário Propensity to Trust Technology de Schneider et al. [ver [Jessup](https://corescholar.libraries.wright.edu/etd_all/2209/?utm_source=corescholar.libraries.wright.edu%2Fetd_all%2F2209&utm_medium=PDF&utm_campaign=PDFCoverPages), Apêndice B] de 2017 especificamente para automação); o questionário Negative Attitude towards Robots (NARS) de [Nomura et al.’s](https://doi.org/10.1075/is.7.3.14nom); e o questionário Multi-Dimensional Measure of Trust - Version 2 (MDMTv2) (devido a [Ullman and Malle [2020]](https://research.clps.brown.edu/SocCogSci/Measures/index.html#MDMTlink) uma escala da confiabilidade subjetiva de um robô específico). Para a mesma experiência, Romeo et al. avaliaram a confiabilidade subjetiva de capacidades e a confiabilidade subjetiva moral no robô social de teste em 706 participantes, maioritariamente na faixa etária dos 18 aos 35 anos, usando os questionários acima classificados sobre uma escala de Likert. As médias para a pontuação de confiabilidade de capacidades estavam na faixa dos 5,3-5,6 de uns possíveis 7, enquanto as médias para a confiabilidade moral estavam na faixa dos 5,1-5,4 de uns possíveis 7. Isto verificou-se para todas as três personalidades de robô testadas, sugerindo que a confiança-base das pessoas em robôs é bastante alta.

Uma imagem com diagrama, file, Gráfico, número

Descrição gerada automaticamente

Ilustração 4Tarefa de distanciamento físico realizada em 2011 por Mumm e Mutlu. A proxémica humana foi avaliada no caso de olhares mútuos e desvio de olhares em robôs simpáticos e antipáticos. Este exemplo mostra a aplicabilidade da proxémica para avaliar a afinidade com o robô. **À esquerda:** Diagrama do experimento: o participante tinha a tarefa de se aproximar da parte de trás do robô enquanto ele desviava ou encontrava o olhar da pessoa. **Ao Centro:** Caminhos percorridos por 60 participantes ao longo de 5 tentativas cada. **À direita:** Distância física resultante (medida combinada das distâncias mínimas frontal e traseira entre o robô e o humano) para os quatro casos ("U" para "antipático"; "L" para "simpático"), mostrando um aumento significativo na distância física no caso de olhares mútuos.

**Fonte:** [Mumm and Mutlu [2011]](https://doi.org/10.1145/1957656.1957786)

No entanto, deve-se observar que os robôs testados foram retratados como companheiros, não como colegas de trabalho. Um robô a realizar uma tarefa que poderia ter sido dada a um ser humano pode causar ansiedade e reduzir a aceitação ([Elprama et al. [2017]](https://doi.org/10.1145/3029798.3038309)). Na mesma pesquisa de 2020 realizada por [Sanneman et al.](https://workofthefuture.mit.edu/wp-content/uploads/2020/11/2020-Research-Brief-Sanneman-Fourie-Shah.pdf) em grandes fabricantes e PMEs mencionada no início, foi também relatado que casos de trabalhadores não gostarem, desconfiarem ou mostrarem-se cautelosos em relação à intenção por trás da integração de robôs eram comuns, mas significativamente amenizados quando os próprios trabalhadores propõem o uso e integração de robôs. Estas observações também foram corroboradas em uma síntese bibliográfica sobre o assunto por [Baltrusch et al.](https://doi.org/10.1080/00140139.2021.1984585) em 2022 (Tabela 2), assim como por [Lowe et al.](https://doi.org/10.1002/hfm.20971) em 2023.

Uma imagem com texto, captura de ecrã, menu, documento

Descrição gerada automaticamente

**Tabela 2:** Tabela resumo de análises abordando os efeitos na qualidade do trabalho da Colaboração Humano-Robô (HRC). Referências: [*Wixted and O’Sullivan [2014]*](https://ihfes.org/wp-content/uploads/2020/03/IESConferenceReview2014Volume1.pdf)*;* [*Charalambous et al. [2013]*](https://dspace.lib.cranfield.ac.uk/handle/1826/9480)*;* [*Charles et al. [2015]*](https://dspace.lib.cranfield.ac.uk/handle/1826/9980)*;* [*Groom and Nass [2007]*](https://doi.org/10.1075/is.8.3.10gro)*;* [*Hayes and Scassellati [2013]*](https://doi.org/10.1016/j.cirp.2013.03.048)*;* [*Kadir et al. [2018]*](https://doi.org/10.21278/idc.2018.0319)*;* [*Meissner and Truebswetter [2018]*](https://doi.org/10.3139/104.111971)*;* [*Parasuraman and Riley [1997]*](https://doi.org/10.1518/001872097778543886)*;* [*Wurhofer et al. [2015]*](https://doi.org/10.1007/978-3-319-27048-7_3)

**Fonte:** [Baltrusch et al. [2022]](https://doi.org/10.1080/00140139.2021.1984585)

Alguns trabalhos foram realizados sobre os fatores que afetam a confiabilidade de agentes artificialmente inteligentes. Em 2021, [Rheu et al.](https://doi.org/10.1080/10447318.2020.1807710) fizeram uma revisão de quais fatores eram considerados como afetando a confiabilidade de agentes de conversação. Em resumo, concluíram que não havia uma solução única para aumentar a confiabilidade do agente, uma vez que fatores como a personalidade e o background cultural das pessoas com quem se interagia afetavam os resultados. No entanto, a qualidade de desempenho e nível de conhecimento, antropomorfismo e honestidade sobre erros e imprecisões foram os factores mais consistentes para aumentar a confiabilidade. Surpreendentemente, os sinais de comunicação não verbal não foram vistas como afetando consistentemente a confiabilidade. Mais tarde, em 2023, [Flathmann et al.](http://https/doi.org/10.1080/07370024.2023.2189595) conduziram um estudo sobre o efeito da etiqueta de um companheiro artificialmente inteligente, determinando que a aderência à etiqueta tradicional levava a um aumento na confiança.

Em relação à Colaboração Humano-Robô (HRC), [Goetz et al.](https://doi.org/10.1109/ROMAN.2003.1251796) determinaram em 2003 que a conformidade das pessoas com os pedidos de um robô previamente desconhecido depende da personalidade do robô e da tarefa associada. Nomeadamente, um pedido associado a uma tarefa séria foi mais prontamente atendido quando feito por um robô sério. Também foi determinado por [Bainbridge et al.](https://doi.org/10.1007/s12369-010-0082-7) em 2011 que pessoas tendem a cumprir mais prontamente os pedidos de um robô corporificado imediatamente presente do que de um robô virtual ou remoto.

Em resumo, foram apresentados métodos para medir a opinião das pessoas sobre robôs, e foi observado que a afinidade e a confiança dos trabalhadores humanos em relação aos robôs dependem de se a implantação e integração foram realizadas considerando os interesses dos trabalhadores. Também foi observado que a qualidade de desempenho de um robô e a honestidade sobre erros melhoram sua confiabilidade; e que seriedade apropriada aumenta a conformidade com seus pedidos.

Uma imagem com texto, menu, captura de ecrã, documento

Descrição gerada automaticamente

**Tabela 3:** Exemplos de robôs colaborativos industriais implantados.

**Fonte:** [Hentout et al. [2019]](https://doi.org/10.1080/01691864.2019.1636714)

A segunda subcategoria de consciência de robôs em humanos definida na Tabela 1 foi a consciência situacional humana entre robôs. Corresponde à perceção de robôs sem interação. É relevante para todas as três categorias de Interação Humano-Robô (IHR) e está relacionada à consciência da presença de robôs.

Atualmente, o nível de consciência dos trabalhadores em relação à presença e ao funcionamento de robôs industriais é geralmente mais que suficiente. Durante a operação normal, na grande maioria dos casos, os robôs industriais atualmente implantados permanecem imóveis sempre que os trabalhadores precisam de estar próximos deles ([OTM Sec. IV Ch. 4, Sec. VI, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#safety)). Os cobots móveis serviriam como contra-exemplo, especialmente para casos de IHR de nível HRC, caso a sua implantação em espaços de trabalho industriais não fosse tão limitada até ao presente (ver Tabela 3, assim como [Murashov et al. [2016]](https://doi.org/10.1080/15459624.2015.1116700) e [Sanneman et al. [2020]](https://workofthefuture.mit.edu/wp-content/uploads/2020/11/2020-Research-Brief-Sanneman-Fourie-Shah.pdf)). Fora da operação normal, durante a integração ou manutenção de um sistema robótico, um trabalhador geralmente estará a prestar atenção explicitamente ao robô em questão, portanto, mais uma vez, o nível de consciência dos trabalhadores é mais que suficiente.

Para os casos de fabricação, integração, manutenção e operação, é implementada uma coleção de avaliações de risco necessárias e métodos de redução de risco ([ISO 12100 [2010]](https://www.iso.org/standard/51528.html)), juntamente com normas de segurança específicas para robôs a serem seguidas ([ISO 10218-1:2011 [2016]](https://www.iso.org/standard/51330.html) e [ISO 10218-2:2011 [2016]](https://www.iso.org/standard/41571.html)), complementadas com normas aplicadas especialmente em cobots ([ISO/TS 15066:2016 [2016]](https://www.iso.org/standard/62996.html)). Todos estes elementos juntos definem uma série de dispositivos de proteção necessários; procedimentos; e manutenção, todos de acordo com o nível de risco, garantindo que os trabalhadores não se depararão com um robô num estado perigoso. Isto inclui a separação de robôs dos trabalhadores, mas também a limitação das capacidades físicas do mesmo quando um trabalhador está próximo.

Uma imagem com file, Gráfico, diagrama, captura de ecrã

Descrição gerada automaticamente

**Figura 4:** Estimativas para estoques de robôs nos Estados Unidos, Espanha e Portugal de 1993 a 2019, de acordo com diferentes métodos. Os números estão expressos em milhares. As estimativas da IFR são obtidas com dados da Federação Internacional de Robótica. As estimativas da PIM são obtidas com o Método de Inventário Permanente, sendo que as percentagens representam diferentes taxas de depreciação.

**Fonte:** [Jurkat et al. [2021]](https://doi.org/10.13140/RG.2.2.25801.72803)

Na realidade, atualmente é difícil um trabalhador não estar ciente da presença de robôs na sua vizinhança, e ainda mais difícil para essa falta de consciência representar um perigo. Como exemplo, nos Estados Unidos, o número de fatalidades de trabalhadores envolvendo robôs de 1993 a 2017 manteve consistentemente uma média de 8,2 por cada intervalo de 5 anos ([Layne [2023]](https://doi.org/10.1002/ajim.23470)), ao mesmo tempo que o stock total de robôs nos Estados Unidos aumentou exponencialmente de 37 mil para 293 mil (Figura 4). Aliás, [Lowe et al](https://doi.org/10.1002/hfm.20971) determinaram em 2023 que a substituição de equipamentos tripulados por robôs industriais resultou numa redução de fatores de risco e lesões.

É importante observar, no entanto, que os dados e normas atualmente disponíveis estão principalmente focados em robôs parados, não colaborativos e pouco autónomos. À medida que os robôs colaborativos se tornam mais comuns nos espaços industriais, esses dados e normas podem tornar-se desatualizados. É crucial que as regulamentações e padrões acompanhem o avanço da tecnologia e crescente presença de robôs colaborativos, para garantir a continuada segurança e eficácia nos ambientes de trabalho industriais.

Uma imagem com texto, captura de ecrã, número

Descrição gerada automaticamente

**Tabela 4:** Lesões associadas às fatalidades de trabalhadores relacionadas a robôs nos Estados Unidos.

**Fonte:** [Layne [2023]](https://doi.org/10.1002/ajim.23470)

O Manual Técnico da OSHA (Administração de Segurança e Saúde Ocupacional) compila uma lista de categorias de perigos associados aos sistemas de robôs industriais ([OTM Sec. IV Ch. 4, Sec V, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#hazards)). Esses perigos podem ser agrupados em perigos devido ao comportamento do robô (impacto, colisão, esmagamento, aprisionamento, etc.); falhas de funcionamento ou fadiga (perigos de projéteis e perigos elétricos, hidráulicos ou pneumáticos); e perigos no ambiente de trabalho (escorregões, tropeços, quedas e perigos ambientais). Destes, os perigos devido ao comportamento do robô são de maior interesse para este trabalho, porque são os que podem ser facilmente alterados por meio de software, e também porque as colisões com braços e pernas humanos são o tipo mais comum de acidente envolvendo sistemas robóticos ([Meziane et al. [2014]](https://doi.org/10.1109/ROSE.2014.6952980)).

Os dados sobre fatalidades de trabalhadores relacionadas a robôs nos Estados Unidos de 1992 a 2017 compilados por [Layne](https://doi.org/10.1002/ajim.23470) em 2023, juntamente com estudos conduzidos na Suécia e no Japão para a Associação de Indústrias de Robótica ([OTM Sec. IV Ch. 4, Sec. V.D, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#accident)), concordam que os acidentes envolvendo robôs mais comummente ocorreram durante a fabricação e integração das aplicações robóticas: ou seja, fora das condições normais de operação, enquanto os trabalhadores ainda não estão familiarizados com o comportamento do robô (ver Tabela 4). No entanto, a operação e manutenção de robôs também são citadas como uma fontes usuais de perigo, sempre que os trabalhadores precisem de estar próximos de um robô ligado ([OTM Sec. IV Ch. 4, Sec. V.B, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#haz-process)). Naturalmente, isto é especialmente relevante para a introdução de cobots no ambiente de trabalho, já que o seu desempenho de tarefas complexas e maior autonomia proporcionam uma fonte persistente de situações perigosas durante a operação e manutenção. As fontes mais comuns de acidentes relacionados com o comportamento do robô durante a operação e manutenção incluem erro humano, negligência ou complacência humana, e má visibilidade ou salvaguarda inadequadas ([OTM Sec. IV Ch. 4, Sec. V.C, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#sources)).

Essas fontes de acidentes mais comuns explicam por que o design de robôs intrinsecamente seguros concentra-se sobretudo na implementação de métodos para que os robôs assegurem ativamente a segurança humana, seja limitando o peso dos manipuladores e operadores terminais (end-effectors); controlando velocidades e forças de impacto; ou ensinando-lhes comportamentos para contornar ou minimizar o risco de ações humanas perigosas detetadas ([Hentout et al. [2019]](https://doi.org/10.1080/01691864.2019.1636714)). Tirando o caso da proteção durante a manutenção ([OTM Sec. IV Ch. 4, Sec. VI.D, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#safety-planning)), muito menos atenção é dedicada na literatura ao desenvolvimento de métodos passivos, como alarmes e sinalização de presença eficazes. Considerando que historicamente os acidentes têm geralmente ocorrido devido ao comportamento inesperado de robôs aos quais humanos estavam ativamente a prestar atenção, e não devido a falta de consciência da sua presença, poucos potenciais acidentes poderiam ser prevenidos por métodos passivos.

Uma imagem com interior, máquina, brinquedo

Descrição gerada automaticamente

**Figura 5:** Rob@Work 3, exemplo de um cobot móvel.

**Fonte:** [Fraunhofer IPA](https://www.ipa.fraunhofer.de/content/dam/ipa/en/documents/Expertises/Roboter--und-Assistenzsysteme/Product_sheet_rob@work_3.pdf)

No entanto, como referido por [Layne](https://doi.org/10.1002/ajim.23470), a introdução de cobots móveis no ambiente de trabalho (exemplo na Figura 5) traz mais oportunidades para acidentes causados pela falta de consciência dos trabalhadores em relação à presença deles, sobretudo devido à distração humana ou à baixa visibilidade. A recomendação da OSHA de instalar dispositivos de consciencialização - barreiras de corda; luzes intermitentes e placas, apitos e buzinas - de acordo com o risco avaliado, é uma maneira de lidar com esse problema emergente. Independentemente disso, fica claro que os métodos para alertar adequadamente os trabalhadores sobre a presença de um robô são muito menos padronizados do que aqueles que simplesmente separam ou limitam o robô.

Uma imagem com rede entrelaçada, aço, Abrigo para animais, Acessórios para animal de estimação

Descrição gerada automaticamente

**Figura 6:** Robô em gaiola. **Fonte:** [Sanneman et al. [2020]](https://workofthefuture.mit.edu/wp-content/uploads/2020/11/2020-Research-Brief-Sanneman-Fourie-Shah.pdf)

O método clássico de garantir a segurança dos trabalhadores próximos de robôs é colocá-los em jaulas (Figura 6). É um método verificado como sendo eficaz durante a operação normal, mas obviamente impraticável para a Colaboração Humano-Robô (HRC). Com isso em mente, a ISO 10218-1/2:2011 define padrões para quatro modos de operação colaborativa, com robôs fora de jaulas físicas, mas ainda tendo o seu comportamento limitado por regras enquanto na proximidade de pessoas. Os modos são Parada Monitorizada com Classificação de Segurança (SMS); Controlo Guiado Manualmente (HGC); Monitorização de Velocidade e Separação (SSM); e Limitação de Potência e Força (PFL). O procedimento deles é representado na Figura 7. A ISO/TS 15066 também fornece fórmulas específicas para os limiares de distância apropriados no SSM e limiares de dor para o PFL. Posteriormente, o NIST propôs alguns ajustes nos critérios de ativação do SSM e do PFL, conforme referido por [Marvel](https://doi.org/10.1109/ICSENS.2017.8234264) em 2017.

Em trabalhos relacionados, [Bdiwi et al.](https://doi.org/10.1016/j.cirp.2017.04.009) sugeriram em 2017 os modos de operação colaborativa de que um cobot estacionário deve ser capaz, dependendo de para qual dos quatro níveis propostos de interação humana o cobot é usado (Tabela 5). Propuseram dividir a Colaboração Humano-Robô (HRCp) em dois níveis distintos: em tarefas compartilhadas sem interação física; e tarefas compartilhadas com entrega de objetos. Informações relacionadas também estão disponíveis em [OTM Sec. IV Ch. 4, Sec. VIII.B, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#risk-collab). Além disso, em 2017, [Marvel](https://doi.org/10.1109/ICSENS.2017.8234264) mencionou alguns projetos de engenharia em progresso no NIST relacionados à conformidade com a norma ISO/TS 15066. Incluiu o desenvolvimento de sistemas de sensores RGB-profundidade-térmicos calibrados para a localização de humanos em ambientes de chão de fábrica, de modo a auxiliar na conformidade com o SSM; e referiu-se ao projeto do Assessment Framework for Robotic Systems, desde então já concluído, cujo objetivo era pesquisar métodos para alinhar elementos de deteção às necessidades para a realização de PFL.

Uma imagem com esboço, desenho, design

Descrição gerada automaticamente

**Figura 7:** Os quatro modos de operação colaborativa definidos pela ISO 10218-1/2:2011. a) é o SMS, b) é o HGC, c) é o SSM, d) é o PFL.

**Fonte:** [Hägele et al. [2016]](https://doi.org/10.1007/978-3-319-32552-1_54)

É importante salientar que a ISO/TS 15066 não altera o framework estabelecido pela ISO 10218-1/2:2011. Como tal, o padrão atualmente em uso limita a colaboração Humano-Robô (HRC), porque todos os modos definidos de operação colaborativa exigem que o robô reduza significativamente o seu desempenho sempre que um humano se aproxime. No entanto, a ISO 10218-1/2:2011 está para ser substituída pela norma [ISO/FDIS 10218-1](https://www.iso.org/standard/73933.html), que provavelmente abrangerá mais casos que permitirão ao robô maior liberdade de movimento.

Finalmente, além das normas de segurança atualmente em vigor, [Hentout et al.](https://doi.org/10.1080/01691864.2019.1636714) sintetizaram em 2019 a bibliografia produzida acerca de táticas de segurança para cobots industriais durante a década de 2008 a 2017, propondo três categorias principais de estudo: táticas de segurança intrínseca; táticas de pré-colisão; e táticas de pós-colisão. As táticas de segurança intrínseca mencionadas em sua maioria giravam em torno da redução do peso das partes móveis ou da adição de amortecedores, tornando mais fácil alcançar o PFL. O foco atual nas táticas de pré-colisão é nas estratégias baseadas em sensores proprioceptivos e exteroceptivos (mais informações em [Indri et al. [2015]](https://doi.org/10.1109/ETFA.2015.7301539) e [Avanzini et al. [2014]](https://doi.org/10.1109/TCST.2014.2300696)), embora estratégias reativas também tenham sido propostas e validadas com sucesso ([Haddadin et al. [2008]](https://doi.org/10.1109/IROS.2008.4650764) e [Flacco et al. [2012]](https://doi.org/10.1109/ICRA.2012.6225245)). Além do que foi estipulado pelas normas do PFL, foram propostas estratégias de pós-colisão, um método inovador para reduzir o risco de lesões por meio de um modelo de resposta de deformação elástica não linear bio-inspirado, proposto por [Courrèges et al.](https://doi.org/10.1109/INDIN.2016.7819164) em 2016.

Resumidamente, foi observado que as medidas de segurança industrial atualmente implementadas garantem que os trabalhadores geralmente tenham consciência suficiente dos robôs em seu redor, e que não estejam em perigo mesmo que não tenham. No entanto, notou-se que isso é apenas verdadeiro para os robôs industriais atualmente implantados: segregados e imóveis. Os métodos de segurança atualmente definidos para cobots móveis em mariotariamente tornam-nos incapazes de realizar as suas tarefas. Foi também mostrado que a principal fonte de acidentes relacionados a robôs é o erro humano, e algumas táticas de segurança conjeturais foram discutidas.

Uma imagem com texto, captura de ecrã, número, design

Descrição gerada automaticamente

**Tabela 5: À Esquerda:** Quatro níveis de interação propostos por Bdiwi et al. O Nível 1 corresponde a HRCx, o Nível 4 corresponde a HRC, e os Níveis 2 e 3 se enquadram em HRCp. **À Direita:** Funções de segurança propostas para os quatro níveis de interação.

**Fonte:** [Bdiwi et al. [2017]](https://doi.org/10.1016/j.cirp.2017.04.009)

A terceira subcategoria de consciência de robôs em humanos definida na Tabela 1 foi a compreensão humana do comportamento do robô. Consiste em percecionar e compreender o comportamento do robô sem necessariamente comunicar com ele. É de interesse mesmo para a integração de cobots móveis, por outro lado capazes de comunicar a sua intenção, pois a comunicação é um processo demorado que requer a atenção de ambos os agentes: se houver muitas pequenas interações entre humanos e robôs, pode ser preferível os trabalhadores entenderem intuitivamente a extensão do movimento de um robô ou os seus planos sem interagir com eles, seja para preservar mais facilmente a segurança dos trabalhadores, ou para delegar tarefas de forma mais eficaz. Esta área de pesquisa na robótica industrial é comparativamente recente, tendo se tornado relevante apenas com o advento da Indústria 4.0, pelo que conhecimento concreto dentro dela é comparativamente escasso.

O aspeto dos robôs afeta a perceção dos trabalhadores acerca das suas capacidades, em particular o nível de antropomorfismo que apresentam. Em 2004, um estudo de [Hinds et al.](https://doi.org/10.1207/s15327051hci1901&2_7) concluiu que trabalhadores se sentem mais confiantes a delegar tarefas a robôs com design humanóides. Portanto, um design humanóide num robô industrial é capaz de transmitir uma perceção exagerada das capacidades do robô e induzir situações perigosas, em particular durante a integração do robô. Por outro lado, é importante observar que a habituação a novas tecnologias, como visto por [Nass and Moon](https://doi.org/10.1111/0022-4537.00153), tende a reduzir a tendência à antropomorfização. Portanto, à medida que os trabalhadores se familiarizam mais com a colaboração humano-robô (HRC), espera-se que a confiança atribuída aos robôs esteja em linha com as suas habilidades reais.

Uma imagem com diagrama, file, Gráfico, captura de ecrã

Descrição gerada automaticamente

**Figura 8:**  Abordagem e dados de enfrentamento do robô para humano, mostrando a contradição entre os resultados de dois experimentos. **À Esquerda:** Dados de abordagem de [Walters et al. [2006]](https://doi.org/10.1080/09540090600879513). **Ao Centro:** Dados de enfrentamento de [Walters et al. [2006]](https://doi.org/10.1080/09540090600879513), mostrando aversão ao enfrentamento frontal. **À Direita:** Dados de enfrentamento de [Koay et al. [2007]](https://aaai.org/papers/0004-exploratory-study-of-a-robot-approaching-a-person-in-the-context-of-handing-over-an-object/), mostrando preferência pelo enfrentamento frontal.

**Fonte:** [Walters et al. [2006]](https://doi.org/10.1080/09540090600879513) and [Koay et al. [2007]](https://aaai.org/papers/0004-exploratory-study-of-a-robot-approaching-a-person-in-the-context-of-handing-over-an-object/)

Começa a surgir uma noção da importância da compreensão humana da cinemática dos robôs. [Sisbot and Alami](https://doi.org/10.1109/TRO.2012.2196303) chamaram a atenção em 2012 para a necessidade de robõs apresentarem movimentos seguros e de intenção evidente, e em 2016 [Murashov et al.](https://doi.org/10.1080/15459624.2015.1116700) destacaram a importância do movimento físico previsível em cobots para a saúde mental humana: movimentos inesperados e surpreendentes de máquinas capazes de causar lesões graves não apenas produziriam um ambiente de trabalho inseguro, mas também fariam com que fosse inconsistente e stressante. Abaixo estão alguns resultados detalhando o que constituiria um movimento de robô seguro, confortável e compreensível.

Em 2006, [Walters et al.](https://doi.org/10.1080/09540090600879513) realizaram experiências acerca da distância social preferida e direção de aproximação na interação humano-robô com um robô de aparência mecânica, concluindo que, para a maioria das pessoas, a distância social preferida é comparável à da interação humano-humano. Também foi proposto por [Walters et al. [2006]](https://doi.org/10.1080/09540090600879513) e [Koay et al. [2007]](https://aaai.org/papers/0004-exploratory-study-of-a-robot-approaching-a-person-in-the-context-of-handing-over-an-object/) que, dependendo do nível de experiência com robôs das pessoas, é preferida uma aproximação lateral ou uma aproximação frontal: pessoas com menos experiência em interação humano-robô foram vistas como percepcionando uma aproximação frontal como perigosa e agressiva; e essa perceção diminuiu com o aumento da habituação, levando a uma preferência por aproximações frontais após algum tempo (Figura 8).

Vale a pena reiterar aqui que o momento mais propenso a acidentes na interação humano-robô industrial é quando o trabalhador humano ainda não está familiarizado com o comportamento do robô específico ([OTM Sec. IV Ch. 4, Sec V.D, 2023](https://www.osha.gov/otm/section-4-safety-hazards/chapter-4#accident)). Uma técnica evidente para a redução de acidentes nessas circunstâncias pode ser o reconhecimento audiovisual individual dos trabalhadores para prever a sua familiaridade com o sistema robótico. Caso não sejam muito familiarizados, o movimento do robô pode ser significativamente abrandado, ou idealmente, perguntas podem ser feitas ao trabalhador para garantir que quaisquer comandos proferidos estejam a ser entendidos corretamente (por exemplo, "Para o meu lado esquerdo ou para o seu?").

Também foi demonstrado convincentemente que pessoas têm preferência por poder ver claramente o operador terminal de um robô na sua proximidade. Na experiência de [Koay et al. [2007]](https://aaai.org/papers/0004-exploratory-study-of-a-robot-approaching-a-person-in-the-context-of-handing-over-an-object/), um robô tinha a tarefa de entregar um objeto a uma pessoa, e 75% dos participantes mostraram preferência por receber o objeto com o robô diretamente à frente deles. Isto foi corroborado numa experiência de [Dufour et al. [2020]](https://doi.org/10.1016/j.robot.2020.103626), onde dos 15 participantes incumbidos de realizar uma tarefa simples na proximidade de um robô, todos relataram sentir-se mais seguros e confortáveis quando o operador terminal estava sempre claramente à vista (Figura 9).

Por fim, [Papenmeier et al.](https://doi.org/10.1007/s12369-018-0493-4) determinaram em 2018 que um robô com um perfil de velocidade decrescente é percepcionado como tendo movimento mais previsível, e que é percecionado como sendo menos autónomo quando a sua direção de frente não coincide com sua direção de movimento.

Uma imagem com pessoa, vestuário, homem, interior

Descrição gerada automaticamente

**Figura 9:** Duas configurações de movimento do manipulador do robô próximo ao humano testadas em [Dufour et al. [2020]](https://doi.org/10.1016/j.robot.2020.103626). **À Esquerda:** Movimento do manipulador do robô com o efetuador final e a ferramenta fora da visão clara do humano. **À Direira:** Movimento do manipulador do robô com o efetuador final e a ferramenta em vista clara do humano.

**Fonte:** [Dufour et al. [2020]](https://doi.org/10.1016/j.robot.2020.103626)

O olhar importa na comunicação entre humanos, e tem-se visto importar também na interação humano-robô (HRI). Em 2017, [Admoni and Scassellati](https://doi.org/10.5898/JHRI.6.1.Admoni) sintetizaram o estado da arte na bibliografia acerca de olhar no âmbito de HRI, examinando também a resposta humana ao olhar do robô. Foi determinado que pessoas apercebem-se consistentemente quando o olhar de um robô está direcionado para elas ([Imai et al. [2002]](https://doi.org/10.1109/ROMAN.2002.1045657)), e que também são sensíveis ao olhar do robô quando este está direcionado para objetos ([Mutlu et al. [2009b](https://doi.org/10.1145/1514095.1514110)] e [Movellan and Watson [2002]](https://doi.org/10.1109/DEVLRN.2002.1011728)). Foi ainda observado que, ao olhar para um objeto durante uma interação com um robô, uma pessoa espera que o robô olhe para o mesmo objeto em cerca de 2,5 segundos ([Pfeiffer-Leßmann et al. [2012]](https://www.semanticscholar.org/paper/An-Operational-Model-of-Joint-Attention-Timing-of-a-Pfeiffer-Le%C3%9Fmann-Pfeiffer/725415987af4d4d944085573c989cad85435bc1f)).

Além disso, o olhar é comprovadamente importante para a atribuição de tarefas em HRC. Um robô que apresenta contacto visual mútuo foi visto como sendo mais propenso a receber assistência durante uma tarefa de ensino ([Ito et al. [2004]](https://doi.org/10.1109/ROMAN.2004.1374806)). O olhar de um robô foi também usado por [Mutlu et al.](https://doi.org/10.1145/1514095.1514109) em 2009 para atribuir papéis específicos durante a interação com parceiros humanos numa conversa, bem como por [Staudte and Crocker](https://doi.org/10.1145/1514095.1514111) em 2009 para melhorar a compreensão da fala do robô pelo humano.

Em resumo, foi visto que, pelo menos por enquanto, um robô industrial deve preferencialmente ter uma aparência de máquina; que as pessoas tendem a preferir a mesma distância social que teriam com um humano, assim como uma aproximação frontal com direção facial correspondente durante o movimento e um perfil de velocidade decrescente; e que ter operadores terminais visíveis é particularmente importante para o conforto humano na proximidade de um robô. Alguns resultados sobre o olhar foram abordados, e uma medida de segurança foi proposta, usando a perceção do robô sobre a familiaridade do humano como base.

Uma imagem com texto, captura de ecrã, Tipo de letra, file

Descrição gerada automaticamente

**Tabela 6:** Matriz de confusão binária para os resultados de eventos de comunicação entre humanos e robôs.

**Fonte:** [Mohan et al. [2010]](https://doi.org/10.1163/016918610X527185)

A quarta subcategoria de consciência de robôs em humanos definida na Tabela 1 foi a comunicação humano-robô. Corresponde à consciência de robôs com a maior quantidade possível de interação: é a tentativa de alcançar um entendimento mútuo das intenções com um robô individual por meio de feedback verbal e não-verbal.

Em 2010, [Mohan et al.](https://doi.org/10.1163/016918610X527185) apresentaram um modelo para classificar os resultados de eventos de comunicação humano-robô com base numa matriz de confusão binária. A premissa era que, num evento de comunicação humano-robô, o humano pode dar uma instrução correta (aquela cuja semântica corresponde ao significado pretendido), ou uma instrução errada (aquela cuja semântica não corresponde ao significado pretendido), e o robô pode obedecer ou recusar / fazer algo diferente. O resultado é a matriz de confusão apresentada na Tabela 6.

Uma imagem com texto, captura de ecrã, Tipo de letra, design

Descrição gerada automaticamente

**Figura 10:** Modelo da característica operacional do recetor do risco do comportamento do robô.

**Fonte:** [Mohan et al. [2010]](https://doi.org/10.1163/016918610X527185)

Com esse modelo, Mohan et al. propõem uma maneira de avaliar o risco relativo de um robô com base nas taxas com que obedece às duas classes de instruções, conforme indicado na Figura 10. Destaca como, mesmo que um robô industrial obedeça perfeitamente a todas as instruções fornecidas, não será necessariamente seguro para HRC: como mencionado acima, a fonte mais comum de acidentes relacionados ao comportamento do robô é o erro humano, formalizado neste modelo como um evento de Interferência Falsa Positiva (FPI, sigla em inglês).

Um evento de Interferência Verdadeira Negativa (TNI, sigla em inglês) pode variar desde pedir esclarecimentos sobre uma instrução ambígua, conflitante ou perigosa, até a ter uma compreensão mútua ou familiaridade suficiente com um trabalhador para conseguir reconhecer contextualmente o que provavelmente teria pretendido dizer com uma instrução claramente incorreta. Esses são resultados que se conseguem por meio de uma comunicação eficaz entre humanos e robôs, o que motiva o seu estudo dentro do contexto da cobótica industrial e da consciência de robôs em humanos.

Em 2008, [Green et al.](https://doi.org/10.5772/5664) propuseram um modelo abrangente da colaboração humano-humano, no âmbito da investigação de colaboração homem-máquina para a NASA. O modelo categorizou todos os sinais de comunicação possíveis em três canais - áudio, visual e ambiental - cada um deles sujeito às características dos meios de comunicação, estas afetando a resultante forma de colaboração. Naturalmente, uma variedade de sinais de comunicação sem restrições em todos os três canais é o ideal para conseguir uma compreensão mútua eficaz, mas o ambiente de alto ruído e a distância entre estações de trabalho, típicos de espaços industriais ([Farrell [2021]](https://www.plantengineering.com/articles/effective-communication-in-a-high-noise-environment/)), tornam essa situação ideal impraticável.

Em vez disso, os meios de comunicação comparativamente desimpedidos têm que compensar pela falta de fidelidade (sobretudo) nos sinais comunicativos de áudio. Para o canal visual, isto consiste em compreender e emular sinais não-verbais legíveis por humanos ([Martin [1997]](https://doi.org/10.1109/MSPEC.1997.576013)): gestos e linguagem corporal; bem como sinais comunicativos indiretos, como a direção do olhar; expressões faciais; e ritmos de interação. Para o canal ambiental, consiste em o robô ser capaz de entender relações espaciais; entender e emular manipulações de objetos; e (especialmente para robôs industriais de autonomia variável) interpretar corretamente interações com sistemas de feedback de força.

Mas esses sinais examinados em isolamento não proporcionam a melhor compreensão mútua possível. Dois contra-exemplos:

* U*m robô social que recebe a instrução "Vá para a esquerda" pode prever a qual "esquerda" o humano se refere com base no olhar do humano, mas apenas se o robô perceber que a instrução e o olhar estão relacionados.*
* *Num local de trabalho no exterior, um trabalhador humano pode franzir a testa ao se ferir, mas também pode franzir ao levar com uma forte rajada de vento. Um caso pode requerer assistência médica, enquanto que o outro é só desconfortável.*

Cada contra-exemplo destaca um fator importante para uma compreensão mútua eficaz na comunicação, para além do acesso a uma variedade de sinais de comunicação. O primeiro aponta para a necessidade de uma análise monolítica dos sinais; o segundo aponta para a necessidade de entender fatores ambientais além da própria comunicação.  
  
A avaliação dos sinais deve, portanto, ser sensível ao contexto. A avaliação sensível-a-contexto dos sinais de comunicação disponíveis, com o objetivo de construir um modelo de IA das vontades e conhecimentos de um humano, aproxima-se conceptualmente da Teoria da Mente (ToM): a habilidade geral de inferir estados mentais de outras pessoas (conforme definido por [Quesque and Rossetti](https://doi.org/10.1177/1745691619896607) em 2020). Isto motiva a investigação de ToM para a colaboração humano-robô industrial. É essencial para cobots móveis eficazes e seguros, e mesmo que os robôs continuem a ser usados exclusivamente para tarefas repetitivas de baixa responsabilidade, é a maneira mais eficaz de prevenir acidentes causados por erros humanos, como mencionado anteriormente.  
   
  
Mais tangivelmente, em 2019, [Bianco and Ognibene](https://doi.org/10.1109/CEEC47804.2019.8974334) identificaram quatro potenciais vantagens funcionais da implementação da Teoria da Mente (ToM) em robôs: gestão de crenças falsas; proatividade e antecipação de comportamento; percepção ativa; e construção de conceitos. Relevantemente à cobótica, ToM permitiria que cobots compreendessem como outros agentes representam o mundo (gestão de crenças falsas); agissem antecipadamente ao comportamento humano (proatividade); permanecessem atenciosos a sinais de comunicação (perceção ativa); e distinguissem comportamentos humanos planeados de reações (construção de conceitos, também veja [Lake et al. [2016]](https://doi.org/10.1017/S0140525X16001837)).

Uma imagem com vestuário

Descrição gerada automaticamente  
  
  
Figura 11: Conjunto padronizado de gestos para HRC proposto por [Barattini et al. [2012]](https://doi.org/10.1109/ROMAN.2012.6343743).  
  
Fonte:[Barattini et al. [2012]](https://doi.org/10.1109/ROMAN.2012.6343743)

A pesquisa sobre a Teoria da Mente Artificial para HRC industrial ainda está nos seus estágios iniciais. Antes que a Teoria da Mente possa ser alcançada, métodos devem ser desenvolvidos para capturar de forma confiável a variedade de sinais de comunicação de alta fidelidade disponíveis no ambiente industrial: isso é pesquisado no campo da Interação Humano-Robô Cognitiva. Em 2019, [Hentout et al.](https://doi.org/10.1080/01691864.2019.1636714) sintetizaram o estado-da-arte da Interação Humano-Robô Cognitiva industrial nessa altura, listando as suas subcategorias como sendo o reconhecimento de ações humanas; o reconhecimento de gestos; o reconhecimento facial; comandos de voz; e olhar e aceitação.

Para o reconhecimento de ações humanas, citam arquiteturas desenvolvidas com base em florestas aleatórias para obter múltiplos reconhecimentos em tempo real ([Akkaladevi and Heindl [2015]](https://doi.org/10.1109/CGVIS.2015.7449900)); e em arquiteturas baseadas em ontologias (termo que se refere a descrições formais do que os objetos são e como se relacionam com outros, ver [*Umbrico et al. [2020]*](https://doi.org/10.1016/j.procir.2020.04.045)) para inferir regras semânticas de comportamento ([Ramirez-Amaro et al. [2015]](https://doi.org/10.1080/01691864.2014.1003096)). Maioria dos trabalhos citados de reconhecimento de gestos usavam interfaces de Realidade Aumentada (AR), e apenas para o reconhecimento de movimentos das mãos. No entanto, dois trabalhos são excecionais nesse aspeto: um deles foi a construção de um framework para controlar robôs por meio do reconhecimento de gestos do corpo detetados por Kinect ([Tsarouchi et al. [2017]](https://doi.org/10.1080/0951192X.2016.1187297)); e outro usou um conjunto padronizado previamente proposto de gestos HRC humanos-para-robô e robô-para-humanos (Figura 11 e [Barattini et al. [2012]](https://doi.org/10.1109/ROMAN.2012.6343743)) para o reconhecimento de gestos baseado em Kinect por meio de dynamic time warping ([Wu et al. [2013]](https://doi.org/10.1109/ICASSP.2013.6638079)).

O reconhecimento de trabalhadores (reconhecimento facial aplicado à HRC industrial) ainda apresenta desafios por enfrentar. Em primeiro lugar, tende-se a usar só para a autenticação de utilizadores: os seus usos para garantir a segurança de trabalhadores não familiares não foram considerados até 2019. Em segundo lugar, a identificação facial não foi citada como tendo sido implementada com sucesso para pessoas com equipamentos de cabeça ou óculos, ou mesmo apenas se olhassem para o lado ([Bdiwi [2014]](https://doi.org/10.1016/j.procir.2014.10.099)). O comando de voz foi considerado utilizável num ambiente de fábrica de teste, desde que o ser humano e o robô estivessem a não mais de 3 metros de distância um do outro ([Barattini et al. [2012]](https://doi.org/10.1109/ROMAN.2012.6343743)). Embora isso reduza significativamente a viabilidade de sinais de comunicação de áudio para cobots móveis, é razoável para tarefas de ensino num robô parado.

Por fim, a pesquisa sobre olhar e aceitação citada por Hentout estava principalmente relacionada à perceção humana do olhar do robô e desde então foi superada por trabalhos mencionados acima. Em particular, grande parte dela está compilada na síntese bibliográfica de [Admoni and Scassellati’s](https://doi.org/10.5898/JHRI.6.1.Admoni) Como exemplo de um trabalho mais recente de reconhecimento de olhar em robôs, em 2022, [Lombardi et al.](https://doi.org/10.3389/frobt.2022.770165) usaram features devolvidas pelo OpenPose numa máquina de vetores de suporte com kernel RBF para alcançar uma precisão de 97% no reconhecimento mútuo de olhar humano-robô. No entanto, esse trabalho foi realizado no âmbito da robótica social, pelo que o uso de capacetes de segurança não foi considerado.

Aliás, a pesquisa sobre ToM está atualmente significativamente mais madura dentro da robótica social, devido às melhores condições para os sinais de comunicação, comparado com o ambiente industrial. As arquiteturas contemporâneas de robótica social geralmente usam várias modalidades para conseguir aprendizagem e execução de planos *online*, para além do mero reconhecimento correto e reprodução de sinais de comunicação em vários canais. Em 2019, [Bianco and Ognibene](https://doi.org/10.1109/CEEC47804.2019.8974334) avaliaram as sete arquiteturas de robótica social então mais recentes relativo à sua implementação das quatro vantagens funcionais de ToM mencionadas acima (Tabela 7). Seis das sete geriam crenças falsas; quatro das sete apresentavam construção de conceitos; duas das sete eram capazes de agir de forma pro-ativa; e finalmente nenhuma delas realizava perceção ativa.

Uma imagem com texto, Tipo de letra, número, file

Descrição gerada automaticamente  
Tabela 7: Vantagens funcionais do ToM implementadas em sete arquiteturas recentes de robótica social. Referências: [Petit et al. [2013]](https://doi.org/10.1109/TAMD.2012.2209880); [Görür et al. [2017]](https://researchspace.csir.co.za/dspace/bitstream/handle/10204/9653/G%C3%B6r%C3%BCr_19463_2017.pdf); [Devin and Alami [2016]](https://doi.org/10.1109/HRI.2016.7451768); [Lemaignan et al. [2017]](https://doi.org/10.1016/j.artint.2016.07.002); [Rabinowitz et al. [2018]](http://proceedings.mlr.press/v80/rabinowitz18a/rabinowitz18a.pdf); [Vanderlest and Winfield [2018]](https://doi.org/10.1016/j.cogsys.2017.04.002); [Demiris and Khadhouri [2006]](https://doi.org/10.1016/j.robot.2006.02.003)Fonte: [Bianco and Ognibene [2019]](https://doi.org/10.1109/CEEC47804.2019.8974334)

Relativo a avanços conjeturais no campo do ToM industrial, aqui estão três trabalhos recentes.

Em 2018, [Winfield](https://doi.org/10.3389/frobt.2018.00075) e a sua equipa apresentaram os resultados de uma série de experiências com robôs equipados com um framework *digital twin* de simulação de eventos como substituto a ToM: os robôs foram capazes de evitar colisões de forma muito eficaz dentro de multidões, assim como em cenários de encontros individuais estilo-passeio; foram capazes de tomar escolhas razoáveis quando apresentados com dilemas éticos; e demonstraram ser capazes de inferir os objetivos de outro robô.

Em 2021, [Chen et al.](https://doi.org/10.1038/s41598-020-77918-x) aplicaram uma arquitetura de rede neuronal profunda para interpretar o *input* visual de um agente IA a observar um robô em movimento. Ao contrário das experiências de Winfield, onde os sujeitos foram dotados de conhecimento prévio sobre as multidões nas experiências de prevenção de colisões, neste caso o agente IA não tinha conhecimento prévio acerca do robô, e mesmo assim foi capaz de inferir de forma consistente os seus planos e antecipar as suas ações.

Por fim, em 2020, [Umbrico et al.](https://doi.org/10.1016/j.procir.2020.04.045) desenvolveram a ontologia HRC sensível ao contexto SOHO, posteriormente estendida em 2023 com padrões de design de ontologia que caracterizam melhor as restrições comportamentais e a dinâmica da colaboração, bem como um procedimento geral de extração de conhecimento para inferência de planos ([Umbrico et al. [2023]](https://doi.org/10.3233/SW-233394)). A ontologia SOHO foi experimentalmente avaliada em diferentes cenários industriais, como em ambientes de trabalho da indústria metalúrgica e automobilística, e mostrou-se capaz de inferir com sucesso planos dentro desses ambientes.

Resumidamente, foi observado que é importante para os robôs avaliar a validade das instruções recebidas para evitar erros humanos, e que a interpretação dos sinais de comunicação deve ser sensível ao contexto, motivando o estudo da Teoria da Mente artificial. O estado da arte da sua componente fundamental, a interação humano-robô cognitiva, foi apresentado, e atenção especial foi dada ao fato de que o reconhecimento de trabalhadores é um campo pouco explorado. O estado da arte em Teoria da Mente para a robótica social foi apresentado como substituto para futuros trabalhos em Teoria da Mente robótica industrial, mas também foram apresentados alguns avanços recentes em direção à Teoria da Mente robótica industrial. Notavelmente, a simulação de eventos de gémeos digitais em vez da Teoria da Mente mostrou resultados impressionantes, e o desenvolvimento de uma ontologia baseada em contexto para interação humano-robô colaborativa foi destacado.

|  |  |  |  |
| --- | --- | --- | --- |
| **Subcategoria** | **Componentes** | **Observações Principais** |  |
| **Opinião** | ***Confiança / Afinidade*** | 1. Depende do método de implantação e integração 2. Melhora com honestidade sobre erros |  |
| ***Conformidade*** | 1. O robô deve ser apropriadamente sério para a tarefa que está delegando |  |
|  |
| **Consciência**  **da situação** | ***Atenção*** | 1. Suficiente por enquanto, mas os padrões são inadequados para a introdução de cobots em espaços industriais 2. O erro humano é a principal causa de acidentes, mas raramente (por enquanto) devido à falta de atenção |  |
| ***Medidas de segurança*** | 1. Efetivas para a segurança dos trabalhadores, mas reduzem a produtividade dos cobots |  |
| **Compreensão do comportamento** | ***Cinemática*** | 1. Robôs com aparência de máquina são atualmente preferíveis 2. Abordagem de frente com direção de frente é preferível 3. Perfil de velocidade decrescente é preferível 4. Distância social semelhante à interação humano-humano é preferível 5. O efetuador final do robô próximo deve estar preferencialmente sempre visível |  |
| ***Olhar social*** | 1. As pessoas prestam atenção ao olhar dos robôs e esperam que os robôs sigam o olhar dos humanos 2. O olhar é importante no ensino e na atribuição de papéis |  |
| **Comunicação humano-robô** | ***Interação humano-robô cognitiva*** | 1. A inferência de regras semânticas de comportamento foi realizada 2. O reconhecimento de gestos foi alcançado com gestos padronizados de HRC 3. O reconhecimento de trabalhadores é pouco explorado |  |
| ***Teoria da Mente (ToM)*** | 1. A simulação de eventos gémeos digitais resultou em evitar colisões, resolver dilemas éticos e inferir objetivos 2. Redes neurais profundas (DNNs) em dados visuais foram usadas para inferir com sucesso planos de agentes previamente desconhecidos 3. SOHO é uma ontologia de HRC baseada em contexto atualmente em desenvolvimento |  |

**Table 8:** Resumo das principais observações por subcategoria

A Tabela 8 resume as principais observações por subcategoria abordadas neste texto. Não estão incluídas nesta tabela descrições das medidas de segurança atualmente implementadas para robôs industriais estacionários e cobots; a sugestão de um método de previsão de acidentes com base no reconhecimento do trabalhador para HRC; e uma explicação da motivação para a aplicação de ToM a robôs industriais. Optou-se por não explorar alguns aspetos relacionados aos humanos na interação humano-robô e colaboração humano-robô, nomeadamente abordagens baseadas em realidade virtual; métodos baseados em realidade aumentada; e abordagens ciberfísicas. Embora certamente abram caminho para uma comunicação humano-robô muito mais eficaz, considerou-se preferível explorar primeiro abordagens mais "económicas".