

## Multi-stakeholder engagement and corrective measures to contribute to a transparent and inclusive online ecosystem PT (Portuguese)

[00:00:18] Bem-vindo ao módulo cinco. O último deste MOOC sobre Informação e Eleições na Era Digital. Você sabe que este curso é organizado pelo Centro Knight para o Jornalismo nas Américas com o apoio da UNESCO e do PNUD. Meu nome é Albertina Piterbarg, sou especialista eleitoral na UNESCO e instrutora principal de curso. Quero agradecer muito a todos por estarem aqui nas últimas semanas. Estamos muito felizes que você esteja achando este curso interessante e útil e isso é muito gratificante e continuaremos nessa direção. Muito obrigado por seus comentários, por sua participação e novamente por estar presente. OK, vamos começar com o último módulo. Falaremos sobre o engajamento de várias partes interessadas e as medidas corretivas para contribuir com um ecossistema online transparente e inclusivo durante as eleições. Neste módulo final, abordaremos mais práticas que podem ajudar a combater práticas nocivas e prejudiciais durante as eleições. Examinaremos a identificação, medidas de verificação de fatos e monitoramento de conteúdo problemático, monitoramento de mídia relacionado as eleições e verificação de fatos. Medidas corretivas, como sinalização de rotulagem e lista negra. Respostas técnicas e algorítmicas. Medidas regulatórias e não regulatórias. E moderação de conteúdo. Então, vamos começar com medidas de identificação, verificação de fatos e monitoramento. A verificação é uma prática tradicional do jornalismo profissional. Consiste na verificação dupla das fontes de informação na busca de evidências primárias de testemunhas oculares ou na verificação dupla de fatos e números. É uma forma de fazer o controle de qualidade do conteúdo antes que ele seja publicado ou divulgado. Por outro lado, a verificação de fatos ocorre após a publicação ou divulgação do conteúdo, e a verificação de fatos busca responsabilizar figuras públicas, instituições e veículos de notícias pelo que estão divulgando. Os verificadores de fatos buscam fontes confiáveis que possam confirmar ou negar reivindicações ao público. Cada vez mais a verificação de fatos também envolve o desmascaramento proativo. Isso significa publicar debates e as evidências demonstram falsidades, muitas vezes explicando o processo envolvido na mudança para essas conclusões. O monitoramento da mídia relacionada ao eleitorado implica um programa de levantamento de notícias e conteúdo relacionado divulgado nas mídias sociais durante um determinado período de tempo, geralmente anexado ao calendário eleitoral. Então você pode ver neste gráfico que a verificação de fatos, o desmascaramento e a verificação sejam três formas de colaborar para que a qualidade das informações que circulam nas redes sejam muito úteis para os processos eletrônicos. Vamos ver essas três etapas da verificação de fatos. Como dissemos, verificação de fatos, essa análise de conteúdo impulsionada por uma pergunta básica, que é como sabemos se isso é verdadeiro ou falso? E para verificação de fatos, para eficiência da verificação de fatos, precisamos seguir três etapas. Então, uma é determinar qual conteúdos podem, devem ou precisam ser verificados. Isso é muito importante. Encontre os fatos procurando as melhores evidências disponíveis sobre a reivindicação em questão e corrigindo a taxa ao avaliar a reivindicação à luz das novas evidências, geralmente quando é menos estressante. Então é assim que geralmente os profissionais de verificação de fatos fazem isso. Seguindo essas três etapas. A verificação da fonte e das origens do conteúdo visual pode ser muito desafiadora. Ainda mais desafiador do que criar conteúdo. Existem diferentes tipos de conteúdo visual falso ou enganoso em comum. Você tem a hora errada de colocar imagens que são realmente comuns, que são o conteúdo que pode ser desmascarado com facilidade, porque são apenas fotos usadas em outros eventos e é fácil acessar os arquivos. Então você manipulou o conteúdo. Isso significa conteúdo que foi digitalizado e manipulado usando edição de fotos ou edição de vídeo. E o último conteúdo de palco, esse é o conteúdo que foi criado e compartilhado com a intenção de enganar. É apenas um conteúdo de palco,

como atuar, preparar ou fazer um cenário em que as informações gravadas ou fotografadas são totalmente falsas. Existem muitas ferramentas diferentes disponíveis para desmascarar a desinformação e a misinformation usuais. Existem várias pesquisas de imagens que permitem ver se essa imagem está em alguns bancos de dados ou onde estiveram arquivadas em outras versões. E isso é muito fácil de fazer, mesmo com o Google. E então você tem que olhar a localização, o que permite determinar onde o vídeo ou a imagem foi capturada. Muitas vezes, a borda de um local requer a referência cruzada de características visuais e pontos de referência do conteúdo com imagens de satélite com vista da rua e conteúdo visual que esteja disponível. Então, é como cruzar diferentes tipos de imagens e pode dar uma olhada na geolocalização e, em seguida, você tem a análise forense de imagens. Algumas ferramentas permitem detectar inconsistências nos metadados da imagem sugerindo manipulação. Como um exemplo de nossa resposta global, podemos mencionar a Rede Internacional de Verificação de Fatos na Poynter, [pointer.org](http://pointer.org) é uma das redes globais de verificação de fatos mais importantes e muito comprometidos em promover a excelência na verificação de fatos globalmente. Eles criaram e desenvolveram um código de princípios que você pode ver na tela. Existem cinco princípios importantes para uma verificação profissional e confiável de cheques. Um deles é estar comprometido com o apartidarismo, justiça, a transparência dos serviços, a transparência do financiamento e da organização, a transparência de metodologia, e políticas de correções abertas e honestas. Mesmo que se enganem, precisam contar ao público para não perder credibilidade. Portanto, esses são os cinco princípios mais importantes para a verificação de fatos. Vamos ver verificação de fatos e monitoramento de mídia eletrônica. Vamos dar uma olhada mais próxima para essa questão. O monitoramento da mídia eleitoral geralmente ocorre por um curto período de tempo. Por exemplo, durante a campanha eleitoral ou outra fase do ciclo eleitoral, como o pós-eleitoral ou durante as eleições, ou outros eventos mas geralmente, por um curto período de tempo. Isso tem o potencial de ser implementado em outros casos, é claro. Por exemplo, digamos que durante o processo de registro. A metodologia de monitoramento tradicional de mídia seguida por um conjunto predeterminado de mídia pode ser combinado com ferramentas automatizadas. Mas isso também é bom para realizar exercícios de monitoramento envolvendo o componente humano e isso é uma combinação de metodologias. O monitoramento de mídia relacionado à eleições geralmente funciona com base em uma amostra que define quais mídias serão monitoradas, quando, como e por quê, e geralmente envolve o uso de indicadores qualitativos e quantitativos. O qualitativo pode ser refletido em palavras-chave e indicadores quantitativos podem ser gráficos e dados e a serem processados para chegar a conclusões. O monitoramento de mídia requer financiamento da infraestrutura para a infraestrutura. É muito caro ativar o hardware e o software. Obviamente, existem outras maneiras de monitorar a mídia de forma muito mais fácil e viável, mas, em geral é um processo caro e, às vezes, é muito provável que você tenha um terceiro que possa estar envolvido no monitoramento automático de recursos humanos e em tendências. O monitoramento de mídia é muito útil e muito eficaz quando combinado com mecanismos e plataformas de alerta precoce e vimos isso em muitos países. Há amostras e indicadores, idealmente uma série de palavras-chave que identificam e funcionam como indicadores do nível de hostilidade ou discurso de ódio presente nos períodos eleitorais. Os relatórios de monitoramento de mídia são extremamente úteis, especialmente quando combinados com mecanismos e plataformas de alerta precoce. Como eu disse antes, a verificação de fatos pode fazer parte dos exercícios de monitoramento da mídia, complementando o conjunto de indicadores. Portanto, existem novas tecnologias que podem ser combinadas para obter melhores resultados, identificar resultados e ser capaz de resolver os problemas relacionados às novas tecnologias na era digital. Ok tudo bem, agora vamos fazer uma pausa de 5 minutos e depois disso continuaremos com as medidas corretivas.

Nos vemos em 5 minutos. Bem vindos(as) de volta. Vamos continuar agora dando uma olhada nas medidas corretivas. Como parte de seus esforços para aumentar a transparência e aumentar a conscientização sobre as crescentes ameaças à liberdade de expressão online. Grandes empresas de Internet começaram a implementar uma série de medidas corretivas. Isso pode ser o resultado da iniciativa unilateral dessas empresas, ou também pode ser o resultado do marco legal em vigor em um determinado país ou região. Mas essas medidas corretivas são muito relevantes para os processos eleitorais. As plataformas sociais, incluindo as mais conhecidas, como Facebook, Twitter, Instagram, Snapchat, etc., implementaram mecanismos de reclamações e apelações que funcionam com base no contrato assinado pelo indivíduo quando usa, quando se inscreve, quando decide usar esses serviços. Esses mecanismos, teoricamente, acionam medidas corretivas internas do lado dessas empresas. Por exemplo, as plataformas online geralmente permitem que os usuários sinalizem conteúdos falsos, violentos, ofensivos, etc. A rotulagem permite que os usuários identifiquem anúncios obscuros, particularmente importante quando relacionados à propaganda política. E a lista negativa (negativelist) pode implicar até mesmo a remoção de conteúdo ou de um usuário específico de uma plataforma de mídia social. No entanto, todas essas práticas não devem afetar o direito à liberdade de expressão e o direito ao acesso à informação, o que é um equilíbrio muito delicado e muito difícil de medir e manter. Esta é a questão-chave de todas essas medidas. Você pode ver na tela um exemplo de sinalização nas mídias sociais, no Twitter, que é muito popular para debates políticos durante as eleições. Você pode ver na tela como o problema é relatado. Respostas por meio de algoritmos técnicos para detecção semi-automática de desinformação online e sua utilidade prática em todo o ciclo de vida das campanhas de desinformação, incluindo respostas de conteúdo, análise de credibilidade, disseminação da rede, medição do impacto nas crenças e ações dos cidadãos e métodos de desmascaramento. Essas respostas técnicas podem ser implementadas pelas próprias plataformas sociais e pelos próprios mecanismos de busca, mas também podem ser ferramentas de terceiros, por exemplo, plug-ins de navegador ou métodos experimentais de pesquisas acadêmicas. Mas, em qualquer caso, a regulamentação contratual interna de empresas privadas, que geralmente é automatizada com revisores humanos limitados, nem sempre é muito transparente ou fácil de entender por todos. A maioria das ferramentas automáticas para detecção de desinformação tem como objetivo fornecer informações para a tomada de decisões humanas. No nível do conteúdo, eles fornecem informações para permitir a análise humana e a manipulação. A suposição por trás das abordagens técnicas e algorítmicas é que elas podem conter e reduzir o compartilhamento de desinformação, misinformation e conteúdo que seja prejudicial. As empresas de Internet implantam uma variedade de modelos de detecção automatizada para tipos e serviços de conteúdo. Isso inclui ferramentas para identificar a disseminação artificial de informações, bem como para identificar conteúdo que atenda aos critérios de classificação, com rotulagem ou remoção. Mas isso é muito desafiador, como mencionei antes, pois o conteúdo circula em milhares de idiomas e o contexto nem sempre é fácil de acessar automaticamente. Vamos ver agora as medidas regulatórias e não-regulatórias, porque a combinação de conteúdo extenso durante o período eleitoral também é uma questão muito complexa, pois pode ameaçar os direitos à liberdade de expressão e o acesso à informação. O processo de desenvolvimento de leis e regulamentos deve seguir uma abordagem aberta, transparente e participativa. Por esse motivo, cada vez mais os países estão tentando fazer cumprir certas regulamentações durante os processos eleitorais porque a situação é muito complicada para garantir que eleições livres e justas não sejam prejudicadas. Mas, ao mesmo tempo, essas medidas precisam ser compatíveis com o artigo 19 da Liberdade de Expressão e do Pacto Internacional sobre Direitos e Direitos Civis. No entanto, a maioria das legislações e regras que regem as eleições e as ferramentas de autorregulação de mídia relacionadas

nem sempre se aplicam a plataformas digitais, mídias sociais e mídias de mensagens. Esse é outro desafio. Por exemplo, o período silencioso que muitos países impõem imediatamente antes do dia da eleição, ou mesmo depois, nem sempre são respeitados pelas redes sociais. Mesmo que um governo proíba a publicação dos resultados das pesquisas antes de uma eleição ou durante o dia da votação. Usuários de mídias sociais podem tornar esses resultados públicos mais cedo, sem respeitar as regras eleitorais. Da mesma forma, a publicidade política geralmente se espalha online e por meio de mensagens sociais e até mesmo durante os períodos de silêncio. E o que é problemático aqui é que esse conteúdo, como vimos nos módulos anteriores, pode ser falsos, com dados e com pesquisas e resultados inventados, o que pode afetar e desacreditar negativamente todo um processo eleitoral. Além disso, como vimos em modelos anteriores, essas ações também podem fazer parte de uma campanha mais ampla de desinformação orquestrada para desestabilizar e prejudicar a credibilidade das instituições democráticas e de um governo específico em um determinado país. A legislação contra a desinformação vêm se desenvolvendo nos últimos anos, e muitos países adotaram legislação contra a desinformação, como regulamentação de conteúdo por meio de bloqueio ou remoção, criminalização com base em difamação ou discurso de ódio, mandatos regulatórios ou legais para monitorar mídias sociais e leis de privacidade de dados e também regulamentos financeiros políticos relacionados à regulamentação de mídias sociais e de empresas de tecnologia ou plataformas de mídia social específicas durante períodos específicos de tempo. No entanto, as iniciativas regulatórias têm sido muito criticadas em alguns casos porque às vezes tendem a dar um poder desproporcional de censura ao estado ou a um governo. E mesmo isso significa, às vezes, dar esse poder a um determinado partido político em detrimento dos outros. Portanto, qualquer tentativa de regular o conteúdo online deve equilibrar os direitos de liberdade de expressão e acesso à informação com a proteção de outros direitos civis e políticos, como participação política, privacidade e liberdade de discriminação. As restrições à liberdade de expressão online devem atender às três partes descritas no Artigo 19 do Pacto Internacional sobre Direitos Civis e Políticos, que consiste: leis e regulamentos devem ser prescritos por lei; toda atenção a um objetivo legítimo pode precisar ser necessária e proporcional; e a autorregulação é um mecanismo totalmente diferente e é um setor de conformidade voluntário, em nível industrial, onde a legislação não necessariamente desempenha um papel na aplicação dos padrões. Como vimos, existem empresas que implementam medidas sozinhas. Então, em outras palavras, essas são iniciativas privadas de conteúdo, moderação. Mas o que exatamente é a moderação de conteúdo. A moderação de conteúdo é o processo de monitoramento se o conteúdo enviado a um site está em conformidade com suas regras e diretrizes e é adequado para aparecer no site. Se for adequado de acordo com a estrutura legal em vigor na área do país ou da região. Isso envolve a definição de regras e diretrizes para todo o conteúdo que for exibido no site obedeça à filtragem de qualquer coisa que seja considerada prejudicial, sensível ou imprópria. Essencialmente, a moderação de conteúdo ajuda a garantir que o conteúdo publicado no site não seja ilegal, impróprio, não assedie ou prejudique outras pessoas. É comumente usado em aplicativos que dependem muito do conteúdo gerado por usuários, como fóruns, plataformas de mídia social, sites de namoro e mercado online. Principalmente as redes de mídia social e outros intermediários da Internet têm se concentrado no desenvolvimento de regras de autorregulação, geralmente na forma de um código de conduta ou dos termos de serviço sobre questões como a remoção de conteúdo e as práticas de processamento. Um exemplo muito interessante de um desses mecanismos é um grande conselho de supervisão que é um órgão independente de moderação de conteúdo. Em 2020, a META criou um conselho de supervisão para o Facebook e o Instagram composto por personalidades altamente renomadas de direitos humanos e liberdade de expressão. A ideia principal era ter um

órgão independente para lidar com as decisões de moderação mais complexas. Inicialmente, no âmbito restrito de qual conteúdo é retirado pela empresa. O conselho aceita casos de usuários que esgotaram o processo de apelação no Facebook ou Instagram. Eles examinam a questão de forma independente para eles, fazem entrevistas gerenciais e tomam decisões, por exemplo, sobre o conteúdo. E eles também fornecem recomendações políticas que cobram a META de criar melhorares regras e agir de forma equitativa, transparente e tratar todos os usuários corretamente. Então, esse é um exemplo interessante. E para finalizar nosso MOOC depois dessas cinco semanas, falaremos um pouco sobre a abordagem regulatória híbrida para processos eleitorais. Essa é uma abordagem muito específica que não pode ser aplicada em todos os lugares, mas está tendo muitos bons resultados e foi muito analisada. E esse é um cenário que o marco legal existente não impõe requisitos restritivos adicionais às plataformas de mídia social, mas afirma os resultados esperados. Por exemplo, um código de prática é uma prática para empresas de mídia social em consulta com essas empresas, bem como com outros partidos políticos e políticos. Uma necessidade mais ampla pode ser desenvolvida, revisada e tornada pública e de forma colaborativa, criando um ecossistema mais forte, onde cada ator está fazendo parte das decisões das medidas tomadas contra conteúdos nocivos. Portanto, todas as partes interessadas eleitorais, temos um papel a desempenhar nisso. Temos um papel a desempenhar nessa abordagem, particularmente no monitoramento eletrônico por essas autoridades de mídia, organizações da sociedade civil, governo e muitas vezes também partidos políticos. Isso implica também que todos esses compromissos são, obviamente, voluntários e dependem do quadro jurídico existente. Não podemos generalizar com eles porque, em muitos contextos, essa abordagem não seria possível. Um exemplo de iniciativa regulatória é a adoção de códigos de conduta e compromissos acordados, por exemplo, por instituições estatais, partidos políticos e grandes empresas de mídia social, anunciantes, veículos de comunicação e outros atores. Esses códigos de conduta geralmente visam garantir a transparência da propaganda política e fortalecer as ações para fechar contas falsas ou impedir atividades nocivas. Rotulando mensagens disseminadas pelos bots e colaborando com verificadores de fatos na academia para conduzir campanhas de desinformação e aumentar a visibilidade e o alcance do conteúdo verificado. Vamos fazer uma breve recapitulação de nossas últimas informações no curso sobre Informação e Eleições na Era Digital. Durante esse módulo, examinamos o engajamento de várias partes interessadas e as medidas corretivas para contribuir para um ecossistema transparente e inclusivo. Muitas das diferentes medidas para ser capaz hoje em dia de combater a desinformação, a msinformation e outras práticas nocivas durante as eleições são missões de verificação, monitoramento e identificação de fatos, medidas corretivas e medidas regulatórias e não-regulatórias, incluindo a abordagem co-regulatória híbrida. Muito obrigada por estar aqui. Estou muito orgulhosa de poder ser sua treinadora principal durante essas cinco semanas. Novamente, agradeço todos os seus comentários, todas as suas sugestões, suas contribuições. Tudo tem sido muito positivo e digno. E muito obrigada. Espero que o curso tenha sido bom e útil para você. Nos vemos em breve!