Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Centro de Liderança em Políticas de Informação

Dez recomendações para uma Regulamentação Global de IA

Outubro de 2023.

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Dez Recomendações para uma Regulamentação Global de IA

Conteúdos

INTRODUÇÃO

A Inteligência Artificial (IA) [i] está gerando benefícios amplos e crescentes para a sociedade, incluindo a potencialização da pesquisa médica, o enfrentamento da mudança climática, a transformação das indústrias e a modernização de governos. Ao mesmo tempo, a rápida implantação e adoção de novos aplicativos, como chatbots de IA generativa e geradores de imagens, intensificaram preocupações antigas e levantaram novas questões relacionadas a privacidade e proteção de dados, transparência e explicabilidade, direitos humanos, propriedade intelectual, segurança, preconceito, impactos na força de trabalho, geração e disseminação de informação errônea e desinformação e outros efeitos sociais. Em resposta, as organizações estão desenvolvendo controles operacionais e estruturas de governança para garantir desenvolvimento e implantação responsáveis em IA; especialistas do setor estão trabalhando para desenvolver padrões; formuladores de políticas estão redigindo novas leis; e reguladores estão testando os limites das autoridades existentes e propondo novas autoridades[ii]. No entanto, não há consenso entre os países sobre a melhor abordagem para se regulamentar a IA: o foco deve ser regulamentação rígida, modelos de corregulação, certificações e garantias, padrões do setor ou alguma combinação destes fatores[iii]?

O CIPL tem sido um líder em reflexões sobre responsabilidade organizacional e abordagem baseada em risco para políticas e práticas de dados há mais de 20 anos e foi um dos primeiros a contribuir para desafios de escopo e a definir soluções para a governança em IA e práticas neste setor. As principais contribuições do CIPL nesse campo incluem *Inteligência Artificial e Proteção de Dados em Tensão* (outubro de 2018), *Problemas Difíceis e Soluções Práticas* (fevereiro de 2020) e *Inteligência Artificial e Proteção de Dados: Como o RGPD da UE Regulamenta IA* (março de 2020)[iv]. O CIPL também preparou respostas detalhadas para consultas públicas sobre políticas de IA no Brasil, na União Europeia, no Reino Unido e nos Estados Unidos[v].

Com base nessa experiência e em nosso amplo envolvimento com líderes do setor privado que desenvolvem e implantam tecnologias de IA, formuladores de políticas e reguladores, o CIPL oferece neste documento dez recomendações para orientar a formulação de políticas e regulamentação de IA, a fim de permitir uma IA responsável e confiável. Estas dez recomendações resumem a visão do CIPL sobre uma abordagem em camadas ou em três níveis para a regulamentação de IA:

a) regras baseadas em princípios e resultados,
b) responsabilidade organizacional demonstrável e
c) supervisão regulatória robusta e inteligente.

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Essa abordagem fornece regras à prova de futuro baseadas em princípios fundamentais que podem orientar o desenvolvimento ético e a implantação da IA, mesmo com a evolução da tecnologia e dos casos de uso. Descrevemos essa abordagem a seguir.

Recomendações para Regulamentação da IA

O CIPL recomenda uma abordagem baseada em riscos e em níveis para regulamentação da IA que aproveite leis e normas existentes e nas práticas responsáveis das organizações. Essa abordagem deve ser apoiada por supervisão regulatória inovadora e instrumentos de corregulação.

Qualquer abordagem legislativa ou regulatória da IA deve seguir estas recomendações abrangentes:

A. Regras Baseadas em Princípios e Resultados
   1. Criar uma estrutura flexível e adaptável que defina os resultados a serem atingidos em vez de prescrever detalhes sobre como atingi-los
   2. Adotar uma abordagem baseada em riscos que considere riscos e benefícios holisticamente
   3. Basear-se em fundamentos legais vinculativos e não vinculativos
   4. Empoderar indivíduos através de transparência, explicabilidade e mecanismos de compensação

B. Responsabilidade Organizacional Demonstrável
   5. Tornar a responsabilidade organizacional demonstrável um elemento central das regulamentações de IA
   6. Promover a adoção de práticas responsáveis de governança em IA
   7. Repartir a responsabilização cuidadosamente, com foco na parte mais intimamente associada à geração de danos

C. Supervisão Regulamentadora Inteligente
   8. Criar mecanismos para coordenação e cooperação ao longo dos órgãos regulamentadores
   9. Instituir supervisão regulamentadora baseada na cooperação e facilitar inovação regulamentadora contínua
   10. Buscar a interoperabilidade global

I. REGRAS BASEADAS EM PRINCÍPIOS E RESULTADOS

1. Criar uma estrutura flexível e adaptável que defina os resultados a serem atingidos em vez de prescrever detalhes sobre como atingi-los

Para serem eficazes, as regulamentações de IA devem ser capazes de permanecer relevantes enquanto a tecnologia e os casos de uso continuam a avançar. Todas as regras devem ser *tecnologicamente neutras*: uma estrutura excessivamente prescritiva e específica para tecnologias individuais ou modelos e práticas de negócios atuais corre o risco de ficar rapidamente desatualizada e inibir inovações benéficas. De fato, uma abordagem baseada em listas de tecnologias específicas exigirá alterações frequentes para acompanhar as mudanças tecnológicas. Se as regras incluírem listas de tecnologias e aplicativos de alto risco presumido, elas devem permitir que essas presunções sejam refutáveis e que evoluam com o tempo.

As regras devem ainda ser *baseadas em princípios e resultados*. Elas devem permitir que as organizações garantam os resultados necessários (por exemplo, justeza, isenção de preconceitos, transparência, precisão, segurança, supervisão humana) por meio de políticas, procedimentos e

controles internos verificáveis e baseados em riscos que sejam apropriados em seus contextos

específicos, sem prescrever como se atingir esses resultados. Essa abordagem oferece aos desenvolvedores a flexibilidade para inovar, incluindo a capacidade de inovar nos controles reais, nas ferramentas técnicas e nas proteções, mantendo a consistência com os princípios e os resultados principais.[vi]

Da mesma forma, uma estrutura regulatória não deve ser excessivamente prescritiva quanto às metodologias de avaliações de impacto e risco de IA, mas descrever critérios que devem ser considerados na avaliação de riscos e benefícios de um aplicativo de IA e deixar que os órgãos reguladores competentes forneçam orientações mais personalizadas, realistas e práticas em colaboração com aqueles que desenvolvem e implantam tecnologias de IA.

Ao mesmo tempo, as regras devem fornecer o máximo de certeza possível em relação a seu escopo de aplicação. Por exemplo, uma estrutura regulatória para IA deve definir IA de modo que as partes interessadas possam entender claramente quais sistemas são cobertos pelas regras. Na ausência dessa clareza e de um foco baseado em resultados, a ambiguidade regulatória e o excesso de prescrição regulatória arriscam inibir o investimento e a inovação — especialmente para pequenas e médias empresas (PMEs) e start-ups, que são poderosos motores de inovação e investimento em IA.

2. Adotar uma abordagem baseada em riscos que considere riscos e benefícios holisticamente

Qualquer abordagem regulatória da IA deve buscar proteger os direitos humanos fundamentais e minimizar os riscos para os indivíduos e a sociedade, ao mesmo tempo em que permite o desenvolvimento e o uso da IA para o benefício de ambas essas partes. De fato, os riscos para indivíduos e sociedade também podem se materializar pelo *não* uso de tecnologias eficazes de IA, como aquelas que têm a capacidade de prever e prevenir doenças, ou reduzir danos on-line, ameaças à segurança cibernética e fraudes. Uma abordagem holística baseada em riscos promove esse objetivo, facilitando medidas práticas de proteção que sejam proporcionais aos riscos e benefícios de um determinado sistema de IA. Seu foco está nos possíveis impactos da tecnologia de IA no contexto de casos específicos de uso.

Uma estrutura regulatória baseada em risco para IA forneceria critérios não exaustivos para ajudar as organizações a determinarem a probabilidade e a gravidade de qualquer dano resultante e as medidas necessárias para mitigá-lo. Avaliar e compreender o possível impacto de seus aplicativos de IA permite que as organizações adaptem suas atenuações aos riscos reais e evitem a implementação de medidas desnecessárias. Por exemplo, o algoritmo k-vizinhos mais próximos (k-nearest neighbors, KNN) é um algoritmo de aprendizado de máquina usado em diversas aplicações[vii]: no varejo, para recomendar produtos; na área de saúde, para prever o risco de ataques cardíacos e câncer de próstata; na área financeira, para detectar atividades fraudulentas; na agricultura, para prever o rendimento das colheitas e, no transporte, para prever e otimizar o tráfego. Esses diferentes usos do mesmo algoritmo KNN têm diferentes níveis de risco – a probabilidade e a gravidade do dano na recomendação de músicas ou roupas são diferentes da recomendação em medicina de emergência.

Uma estrutura baseada em riscos também deve avaliar os possíveis benefícios de um sistema de IA para indivíduos, organizações e sociedade. Esses benefícios podem ser comparados com os riscos identificados na implantação (ou não implantação) de IA. Por exemplo, os riscos de veículos autônomos (VAs) dependem dos diferentes ambientes em que são implantados. É possível que haja um risco menor de danos às pessoas na implantação de veículos autônomos em mineração e agricultura, comparativamente com áreas urbanas ou residenciais. Ao mesmo tempo, os veículos autônomos nos primeiros cenários oferecem diferentes benefícios, como redução da oferta de mão de obra, apoio à agricultura sustentável e aumento de produtividade[viii]. O mesmo exemplo é útil para

destacar a importância da avaliação cuidadosa das métricas pelas quais se mede o risco: pode-se

avaliar o risco dos VAs em relação ao status quo (um mundo em que a maioria dos carros é dirigida por pessoas) e/ou um padrão ideal designado para VAs.

Em resumo, os resultados de avaliações holísticas de risco para o uso de uma tecnologia de IA específica podem variar significativamente segundo os casos de uso. Do ponto de vista da formulação de políticas, isso significa que a identificação definitiva com antecedência dos usos de alto ou baixo risco pode ser difícil, pois o contexto da implantação é fundamental. Uma abordagem baseada em risco é preferível a uma abordagem categórica de definição de sistemas de IA que sejam considerados de alto risco automaticamente. Por exemplo, considerar todos os sistemas de IA que estão fazendo inferências com base em dados biométricos como de alto risco poderia abranger usos acessórios, de risco relativamente baixo, como quando a IA é usada para aplicar filtros ou melhorar a qualidade de vídeo em chamadas de vídeo.

Uma abordagem mais adequada seria:

A) descrever fatores, critérios e possíveis danos que as avaliações de risco devem considerar;
B) fornecer, no máximo, uma lista ilustrativa de usos de risco potencialmente mais alto ou mais baixo que possam ser refutados em cada caso específico;
C) fornecer orientação contínua sobre como avaliar riscos e benefícios com base no aprendizado ao longo do tempo.

3. Basear-se em fundamentos legais vinculativos e não vinculativos

Um regime de IA flexível e adaptável deve basear-se em estruturas jurídicas existentes, incluindo regulamentações e legislações ("hard laws") e "soft law" (por exemplo, os Princípios de IA da OCDE). Muitos setores onde a IA encontra aplicação já são altamente regulamentados (por exemplo, saúde, finanças) e as leis e regulamentações existentes já fornecem requisitos, estruturas de conformidade e recursos que se aplicam ao uso de IA. Entretanto, as leis e regulamentações existentes relevantes também podem ter de ser interpretadas de uma nova maneira e adaptadas às realidades de IA. Onde houver lacunas regulatórias relativas aos riscos relacionados a IA, elas devem ser preenchidas com intervenções regulatórias e corregulatórias direcionadas, priorizando os setores em que regulamentações existentes não se aplicam.

Contar com estruturas legalmente vinculativas existentes, na medida do possível, reduz o risco de criação de regras sobrepostas ou conflitantes, que poderiam levar à incerteza jurídica e a proteções inconsistentes. As regras existentes de antidiscriminação, proteção ao consumidor, propriedade intelectual e, principalmente, proteção de dados e privacidade são relevantes para a lida com muitos dos riscos mais importantes associados a IA. Por exemplo, em março de 2020, o CIPL produziu uma análise abrangente descrevendo como o Regulamento Geral sobre a Proteção de Dados da UE já regulamenta a IA em relação ao uso de dados pessoais[ix]. Onde houver lacunas nas estruturas legislativas — como nos Estados Unidos, que atualmente não possuem uma lei federal abrangente sobre privacidade — preenchê-las é uma base importante para uma regulamentação sólida de IA[x]. Onde as estruturas existentes forem relevantes, as agências reguladoras podem promover a conformidade emitindo orientações sobre como essas regras se aplicam a IA. Ao consultar uma série de partes interessadas, os órgãos reguladores podem identificar as circunstâncias em que essa orientação será mais útil.

Além disso, é importante reconhecer que as regras existentes podem exigir alguma adaptação e uma interpretação regulatória evoluída para alinhá-las aos desenvolvimentos da tecnologia de IA. Por exemplo, determinados princípios encontrados em muitas leis de proteção de dados, como a base legal para o processamento, a especificação da finalidade e a limitação do uso, podem estar em tensão com as necessidades dos sistemas de IA e a forma como eles operam.

Quanto ao conceito de base legal, pode não haver bases legais suficientes nas leis atuais de

proteção de dados para permitir que os desenvolvedores de IA usem categorias sensíveis de dados

pessoais, como saúde, gênero e etnia, para assegurar que o modelo de IA seja treinado e opere de forma a não resultar em desfechos tendenciosos ou discriminatórios. Além disso, se os dados pessoais forem processados com base em uma base legal específica para fins específicos e usados somente para esses fins ou para fins "compatíveis", geralmente com interpretações restritas sobre o que constitui "compatível", isso pode contradizer a natureza de como os algoritmos de IA operam e aprendem. Dado o potencial da IA para descobrir novos e imprevistos usos de dados, esses princípios podem frustrar desnecessariamente aplicações benéficas de IA, a menos que recebam uma interpretação mais ampla para o contexto de IA. Um exemplo dessa interpretação mais ampla seria aplicar uma definição mais ampla de "compatível", que inclua quaisquer finalidades que não neguem ou entrem em conflito com a finalidade inicial e não aumentem o risco de danos aos indivíduos. Outra solução seria considerar o treinamento algorítmico como uma finalidade em si, separada da finalidade de implantação do algoritmo em um caso de uso específico. Isso permitiria a coleta e o uso mais amplo de dados na fase de treinamento, de modo a garantir adequação ao treinamento e ao funcionamento do algoritmo. Por fim, surgem tensões semelhantes com relação aos princípios de minimização de dados e limitação de retenção, que podem limitar as oportunidades dos algoritmos para "aprender" por meio do reconhecimento de correlações. Em suma, se determinados princípios tradicionais de proteção de dados forem interpretados de forma muito rígida, eles poderão bloquear o desenvolvimento e a implantação de aplicativos de IA benéficos — ou ter consequências não intencionais, como a introdução de preconceitos indesejados, limitando o acesso a diversos dados de treinamento[xi]. Os reguladores precisam ser capazes de fazer evoluir a interpretação dos princípios de proteção de dados existentes por meio de orientações regulatórias desenvolvidas em consulta com desenvolvedores e implantadores de IA.

Por fim, as regras existentes devem ser ampliadas com estruturas de soft law, padrões do setor e ferramentas regulatórias conjuntas desenvolvidas em parceria com as partes interessadas, como códigos de conduta, certificações e modelos de garantia. As normas internacionais podem ajudar a estabelecer requisitos básicos para um desenvolvimento e uma implantação de IA que reflitam entendimentos e valores compartilhados, obtidos por meio de processos de desenvolvimento com várias partes interessadas. Os Ministros Digitais e de Tecnologia do G7 reafirmaram o papel fundamental das normas em sua Cúpula de Hiroshima em abril de 2023[xii] e concordaram, em setembro, em desenvolver um Código de Conduta para organizações que desenvolvem sistemas avançados de IA[xiii], enquanto o Grupo de Trabalho de Certificação multissetorial está liderando um trabalho promissor sobre a Certificação de IA[xiv]. Aproveitar as estruturas de soft law, como os Princípios de IA da OCDE, pode promover o alinhamento internacional das regulamentações de IA: por exemplo, a versão parlamentar da Lei de IA da UE obteve sua definição de "Inteligência Artificial" a partir desses princípios.

4. Empoderar indivíduos através de transparência, explicabilidade e mecanismos de compensação

O CIPL tem defendido o empoderamento individual como um princípio fundamental da regulamentação sólida da privacidade e a mesma situação se aplica a IA. Para que a IA seja confiável e benéfica para todos, as regulamentações, as estruturas de corregulamentação e as práticas do setor devem capacitar os indivíduos por meio de:

- **Transparência.** Os desenvolvedores e implantadores de IA devem oferecer transparência significativa e adequada ao contexto sobre as entradas e operações dos sistemas de IA, preservando a privacidade e a proteção de dados, a segurança, a proteção e os segredos comerciais. Essa transparência contextualizada deve se estender aos usuários comerciais dos sistemas de IA, auditores, reguladores e ao público em geral.
    - Os **sistemas de IA de alto risco** devem documentar como o sistema deve ser usado,

os usos e riscos inadequados conhecidos e as recomendações para os

implantadores sobre como gerenciar esses riscos.

- A **IA Generativa** exige medidas para garantir que os usuários entendam as práticas e limitações dos dados dos modelos. O ideal é que desenvolvedores e implantadores ofereçam transparência por meio de vários mecanismos, incluindo políticas, termos de serviço, notificações de produtos e plataformas de recursos centralizados.

- **Explicabilidade.** A explicabilidade é um aspecto da transparência e um meio de aumentar a responsabilidade e a confiança. Ela exige que desenvolvedores e implantadores expliquem de forma significativa como os sistemas de IA afetam as decisões e os resultados que impactam indivíduos, tendo em mente compensações, como entre explicabilidade e segurança e explicabilidade e precisão. Quanto mais complexo e preciso for o algoritmo, mais difícil poderá ser explicar como ele realmente funciona. Também pode haver restrições técnicas à explicabilidade em algumas circunstâncias. Por exemplo, nem sempre é possível explicar como os grandes modelos de linguagem (LLMs) geram resultados específicos com base em entradas individuais ou parâmetros do modelo. As organizações terão que documentar as contrapartidas relevantes para demonstrar como e por que priorizaram a precisão em detrimento da explicabilidade. Um caso em questão pode ser os algoritmos de IA usados na área de saúde e medicina, em que a IA pode possibilitar determinados benefícios à saúde que não podem ser obtidos por ferramentas que não sejam de IA, mas que podem não ser explicáveis. Nessas circunstâncias, a precisão pode ter precedência sobre a explicabilidade. Em suma, dependendo do contexto, dos riscos e dos possíveis benefícios de um caso de uso específico, exigir total explicabilidade como condição de uso pode não ser apropriado em todos os casos.

- **Feedback e compensação do usuário.** Quando os indivíduos não entenderem uma decisão tomada pela IA ou acreditarem que foram prejudicados pela IA, deve haver opções claras para feedback do usuário, consultas, reclamações, mais transparência, direito de contestar a decisão, exigência de revisão humana e, em última análise, reparação, bem como ação das autoridades executivas, quando apropriado e necessário. Os desenvolvedores e usuários corporativos devem considerar como permitir mais transparência, revisão humana no caso de uso contestado de IA, bem como oportunidades de coleta de reclamações e reparação como parte do projeto de soluções de ponta a ponta que aproveitam a IA.

II. RESPONSABILIDADE ORGANIZACIONAL DEMONSTRÁVEL

5. Tornar a responsabilidade organizacional demonstrável um elemento central das regulamentações de IA

Para garantir a responsabilidade dentro do ecossistema mais amplo, as regulamentações devem facilitar o uso demonstrável pelas organizações de estruturas de responsabilidade e programas de governança que forneçam ferramentas e processos para que as organizações implementem todos os requisitos legais relevantes e outros padrões. Assim como em outras áreas de conformidade corporativa tradicional e ética nos negócios — e, mais recentemente, nas esferas de dados, segurança e digital —, a responsabilidade deve ser incorporada e implementada em todos os estágios do ciclo de vida da IA e da "pilha de tecnologia" da IA, incluindo infraestrutura, modelos e aplicativos do centro de dados da IA[xv].

Há uma variedade de estruturas de responsabilidade que fornecem modelos úteis para a elaboração de programas de responsabilidade organizacional e de governança de IA, incluindo a NIST AI Risk Management Framework (Estrutura de Gestão de Risco de IA do NIST) dos EUA, a

Model AI Governance Framework (Estrutura de Governança de IA Modelo) de Cingapura e a

Estrutura de Responsabilidade do próprio CIPL, descrita nos Anexos 1-3 deste relatório[xvi].

As organizações também precisam ser capazes de *demonstrar* responsabilidade internamente, para seus diretores executivos e conselhos corporativos, bem como externamente, para acionistas, investidores, reguladores e o público em geral. Certificações, auditorias, códigos de conduta e avaliações são ferramentas úteis para demonstrar responsabilidade. De fato, esses mecanismos de responsabilidade são essenciais na política e regulamentação digital, inclusive para desenvolvedores e implantadores de inteligência artificial, pelos seguintes motivos:

- Eles demonstram a todos os atores de toda a organização o compromisso e a capacidade de garantir que os produtos e serviços atendam a critérios específicos.
- Eles permitem que as organizações traduzam requisitos legais baseados em princípios e resultados em controles demonstráveis e baseados em riscos, garantindo uma regulamentação mais eficaz e uma melhor conformidade na prática.
- Elas desempenham um papel importante no fornecimento de segurança jurídica e no fortalecimento da confiança, inclusive em contextos business-to-business.

Qualquer regulamentação de IA deve incluir explicitamente a responsabilidade demonstrável como um elemento central, além de permitir o desenvolvimento e o uso de estruturas de corregulamentação, como esquemas de certificação e códigos de conduta, que facilitem e demonstrem essa responsabilidade.

6. Promover a adoção de práticas responsáveis de governança em IA

Embora um conjunto básico de práticas de responsabilidade deva ser exigido para as organizações desenvolverem e implantarem IA, os formuladores de políticas e os reguladores também devem encorajar e incentivar proativamente a adoção de práticas, estruturas, ferramentas e tecnologias de responsabilidade mais amplas. Eles devem trabalhar com as partes interessadas para codesenvolver ferramentas e estruturas para criação e demonstração de responsabilidade em IA. O objetivo deve ser criar um ambiente em que as organizações vejam a adoção de estruturas de responsabilidade bem desenvolvidas como diferenciais para criar valor e aprofundar a confiança em suas práticas de dados, além de cumprir obrigações legais e regulamentares básicas.

Os formuladores de políticas e os órgãos reguladores também devem entender impulsionadores e desafios relacionados às práticas tecnológicas responsáveis e às soluções tecnológicas, como as Privacy Enhancing Technologies (tecnologias de proteção da privacidade, PETs), e tomar medidas para incentivar seu desenvolvimento e sua adoção mais ampla[xvii].

Deve-se considerar uma ampla gama de possíveis incentivos à responsabilidade[xviii], incluindo:

- Reconhecer formalmente a responsabilidade demonstrada ou certificada como um fator atenuante em ações de fiscalização e na avaliação de sanções e/ou níveis de multas;
- Usar a responsabilidade demonstrada como uma forma de "licença para operar", dando às organizações responsáveis maior liberdade para desenvolver e implantar modelos de IA de forma responsável;
- Permitir o uso mais amplo de dados em projetos de IA para pesquisas socialmente benéficas que tenham sido validadas por avaliações de risco, mitigações, supervisão e controles relevantes em programas de responsabilidade;
- Permitir que as partes que adquirem sistemas de IA cumpram os requisitos de devida diligência, adquirindo sistemas que tenham sido certificados de acordo com padrões reconhecidos para IA responsável.
- Usar a responsabilidade demonstrada pela IA como um critério de elegibilidade para projetos de compras públicas de modo a incentivar os contratados a obterem a certificação

de IA responsável.

7. Repartir a responsabilização cuidadosamente, com foco na parte mais intimamente associada à geração de danos

A adoção de mecanismos de responsabilidade organizacional por todos os atores do ecossistema de IA levará a uma melhor conformidade e a melhores resultados na prática e, provavelmente, resultará em menos necessidade de se recorrer a questões relacionadas à responsabilização. No entanto, quando a responsabilização suscita preocupações, há debates ativos em andamento sobre a distribuição adequada entre as partes no ecossistema de IA.

Em princípio, a responsabilização deve ser atribuída principalmente à parte mais intimamente associada à geração do dano em questão, mas a atribuição de responsabilização pode ser complexa na prática. A análise será moldada pelos padrões e precedentes legais existentes, bem como pela extensão em que as partes divulgam informações relevantes por meio de requisitos de transparência e relatórios.

Dependendo das circunstâncias, a responsabilização pode ser atribuída ao desenvolvedor, ao implantador, aos usuários finais ou a alguma combinação destes. Os desenvolvedores podem ser o foco apropriado da responsabilização por sistemas que foram insuficientemente testados quanto a possíveis danos ou fornecidos aos usuários com indicações enganosas sobre aptidões. Por outro lado, os usuários compartilham a responsabilidade pela forma como usam os sistemas de IA, pois determinam se usarão um sistema para um uso de maior risco ou de formas que sejam expressamente contraindicadas pela orientação fornecida pelos desenvolvedores.

Como em outras áreas do comércio, os contratos — incluindo práticas específicas e emergentes de contratação de IA — desempenharão um papel importante na distribuição de responsabilidades e obrigações das partes no ciclo de vida de desenvolvimento e implantação de IA. Por exemplo, se um desenvolvedor proibir contratualmente um caso de uso de alto risco de seu produto, o risco de uso indevido deve ser transferido para o usuário que violou intencionalmente os termos desse contrato. Em cenários em que terceiros fornecem modelos de IA ou soluções habilitadas para IA, a responsabilidade entre desenvolvedores de modelos e implantadores deve ser especificada em contratos.

III. SUPERVISÃO REGULAMENTADORA INTELIGENTE

8. Criar mecanismos para coordenação e cooperação ao longo dos órgãos regulamentadores

A IA é usada em setores regidos por diferentes regulamentações e reguladores. Por exemplo, as autoridades de proteção de dados (DPAs) terão competência geral sobre o processamento de dados pessoais usando IA. Outros órgãos reguladores têm competências mais específicas para cada setor em relação aos aplicativos de IA, como saúde, habitação, serviços financeiros, telecomunicações, produtos farmacêuticos e disciplinas transversais, como propriedade intelectual. Na maioria das circunstâncias, não deve haver necessidade de um novo regulador de IA abrangente, pois isso provavelmente resultaria em excesso de regulamentação, sobreposição, inconsistência e falta de segurança jurídica. Em vez disso, é mais apropriado

a) aprimorar as competências e capacidades dos reguladores existentes para que estejam prontos para a supervisão e fiscalização de IA; e
b) permitir a coordenação e a colaboração de alto nível da política de IA entre as autoridades existentes.

Embora cada órgão regulador deva manter a competência sobre sua própria alçada (por exemplo, para fins de segurança jurídica, as DPAs devem manter a competência geral sobre aplicativos de IA

que envolvam o processamento de dados pessoais e/ou impactem a privacidade dos indivíduos), um

órgão central de coordenação governamental permanente deve ser criado para definir políticas e metas de IA de alto nível aplicáveis a todos os setores e indústrias e facilitar o alinhamento, a coordenação regulatória e a ação conjunta entre diferentes órgãos reguladores, quando necessário e apropriado. O órgão de coordenação pode oferecer aos reguladores um espaço para discutir as compensações entre diferentes objetivos políticos, incluindo eficiência, produtividade, justiça, privacidade, segurança e resiliência. Ele também pode esclarecer a quem as partes devem recorrer para obter orientação em circunstâncias específicas de desenvolvimento e implantação de IA.

Essa abordagem seria benéfica tanto para organizações quanto para órgãos reguladores, promovendo a consistência nas abordagens regulatórias, bem como políticas e orientações holísticas e interdisciplinares que são mais fáceis de implementar e monitorar por órgãos reguladores especializados e pelo setor ao longo do tempo. Essa abordagem também pode ser útil para harmonizar novas leis e regulamentações com as já existentes.

Um exemplo de coordenação entre regulamentações é o Digital Regulation Cooperation Forum (Fórum de Cooperação em Regulamentação Digital, DRCF) do Reino Unido. Ele inclui um CEO e uma equipe permanentes, atividades conjuntas, orientação conjunta e outras ações regulatórias, bem como projetos de colaboração formal e destacamentos de pessoal. A IA tem sido um foco do trabalho do DRCF, conforme evidenciado por seu fluxo de trabalho plurianual sobre transparência algorítmica[xix]. Outros países, como Austrália, França, Irlanda e Holanda, também estabeleceram mecanismos de cooperação para reguladores[xx] .

9. Instituir supervisão regulamentadora baseada na cooperação e facilitar inovação regulamentadora contínua

À medida que a tecnologia continua a evoluir, os reguladores, as técnicas regulatórias e as ferramentas também precisam evoluir.

a) Os reguladores precisam melhorar suas habilidades, capacidades e a forma como operam em um mundo onde há interesses conflitantes e múltiplos em jogo. A tarefa das autoridades de proteção de dados, por exemplo, não deve se limitar à proteção dos direitos fundamentais dos indivíduos, mas também deve permitir a utilização responsável dos dados e o desenvolvimento da tecnologia de IA em benefício da sociedade e da economia, de tal forma que proteja os direitos fundamentais. Isto requer uma mudança na mentalidade regulatória, nas prioridades regulatórias e na ação regulatória. Essa mudança é essencial para que os reguladores atuais se mantenham relevantes e eficazes em um novo mundo digital.

b) Os reguladores devem adotar uma abordagem baseada no risco para serem estratégicos e eficazes. Isto exige a compreensão dos riscos e benefícios dos sistemas de IA e a concentração nas áreas que apresentam os maiores riscos para os indivíduos e a sociedade, preservando simultaneamente os benefícios da tecnologia de IA e o seu avanço. Também é necessário que os reguladores priorizem todo o seu trabalho – estratégia regulatória, orientação, supervisão e aplicação e que se concentrem nas áreas que criam os maiores riscos para os indivíduos e a sociedade.

c) Os mecanismos tradicionais de supervisão baseados exclusivamente ou principalmente na aplicação *ex post* podem já não ser suficientes em uma sociedade digital e habilitada para a IA. A confiança na resolução das falhas do mercado apenas através da aplicação da lei não levará aos resultados desejados. Dado o ritmo de avanço da tecnologia de IA e a necessidade de compreender os seus riscos e benefícios, existe uma necessidade premente de uma abordagem mais cooperativa baseada no envolvimento construtivo contínuo entre reguladores e entidades reguladas, compartilhamento de experiências e informações sobre

desenvolvimentos tecnológicos, além do trabalho em conjunto para desenvolver metas de conformidade realistas e interpretações das regras aplicáveis. Isto exige que tanto reguladores como entidades reguladas sejam transparentes e estejam prontos para se envolverem no compartilhamento construtivo de informações em tempo real à medida que as tecnologias e as práticas empresariais mudam. Investir em medidas *ex ante*, como o incentivo à responsabilização proativa e demonstrável, provavelmente alcançará melhores resultados do que a dispendiosa aplicação *ex post*. É claro que a aplicação da legislação deve continuar a ser uma opção regulatória e um expediente importante em relação a violações repetidas, graves e negligentes, que causam danos reais aos indivíduos e à sociedade.

d) Ferramentas regulatórias inovadoras, como sandboxes e prototipagem de políticas, podem ser eficazes para a supervisão regulatória de novas tecnologias, como a IA. Elas fornecem aos reguladores uma compreensão mais profunda e experiência em primeira mão de aplicações e desenvolvimentos de IA, voltados para o mercado geral. Elas também proporcionam um porto seguro para a indústria testar riscos e benefícios da inovação responsável com uma ligação direta ao regulador competente. Os governos devem fornecer financiamento e recursos adequados para que os reguladores desenvolvam e se envolvam em sandboxes regulatórias e sejam capazes de dimensionar estas atividades para grupos maiores de participantes, inclusive em uma base setorial.

Sandboxes regulatórias: sandboxes regulatórias são mecanismos importantes para exploração e experimentação regulatória, pois fornecem uma plataforma de testes para a aplicação de leis a produtos e serviços inovadores em ambientes da vida real, sob a supervisão de um regulador[xxii]. Elas podem ser usadas para ajudar a abordar e resolver alguns dos aspectos mais desafiadores da implantação de IA no contexto dos requisitos legais em vigor, especialmente aqueles que parecem inconsistentes ou em conflito com relação às novas tecnologias.
Exemplos incluem:

- O Gabinete do Comissário de Informação do Reino Unido administra um programa sandbox estabelecido desde 2020, com foco particular em tecnologias emergentes e biometria[xxiii];
- O Data Regulatory Sandbox, da Singapore Infocomm Media Development Authority (IMDA), permite que as empresas obtenham orientação regulatória relativa a tecnologias inovadoras e com uso intensivo de dados. A IMDA também opera uma sandbox específica para promover o desenvolvimento e a adoção de tecnologias que melhoram a privacidade (PETs)[xxiv].
- A Autoridade Norueguesa de Proteção de Dados (Datatilsynet) lançou uma área restrita regulatória especial para aplicativos de IA[xxv];
- O governo colombiano desenvolveu uma área restrita regulatória para promover a privacidade desde a concepção e o padrão em projetos de IA[xxvi];
- A CNIL francesa opera uma área restrita que concluiu projetos em saúde digital e tecnologia educacional. Em 2023, anunciou uma nova iniciativa focada em IA[xxvii];
- O projeto de lei da UE sobre IA permitiria, e poderá, em última análise, exigir que os estados membros estabeleçam sandboxes regulatórias para IA. A Espanha foi o primeiro estado membro a testar uma sandbox de IA[xviii].

As sandboxes regulatórias devem ser concebidas de forma a incentivar a inovação, o compartilhamento de informações e outros modos de cooperação. Qualquer quadro regulatório de IA deve fornecer uma base legal expressa para que os reguladores estabeleçam sandboxes, incluindo sandboxes de regulamentação cruzada com reguladores apropriados e relevantes, incluindo proteção de dados, concorrência, meios de comunicação, consumidores, saúde/indústria farmacêutica, telecomunicações e reguladores financeiros. As regulamentações devem ter em conta a forma como as autoridades legais ou as prioridades de aplicação podem afetar a

participação das empresas. Ao mesmo tempo, para garantir a confiança do público nos resultados,

as sandboxes devem incluir garantias de que os indivíduos continuarão a ser protegidos contra danos, mesmo durante a experimentação de políticas.

Prototipagem de políticas: Trata-se de projetos-piloto que mobilizam agentes públicos e privados para explorarem, avaliarem e desenvolverem conjuntamente diferentes modelos legislativos de governança antes da sua efetiva promulgação. O processo normalmente envolve a seleção de um grupo de participantes, tais como empresas de tecnologia em estágio inicial, para desenvolver e aplicar protótipos de políticas em parceria com o governo, a indústria e especialistas acadêmicos. O programa OpenLoop da Meta tem sido um dos principais praticantes da prototipagem de políticas, inclusive para a proposta de Lei sobre IA da UE[xxix]. A IMDA, de Singapura, tem um programa de prototipagem de políticas em sua Sandbox Regulatória de Dados que se concentra em notificação, consentimento e divulgação; transparência e explicabilidade de IA; além de transparência e consentimento no metaverso, incluindo contextos para aplicação de interesse legítimo como base para processamento[xxx].

10. Buscar a interoperabilidade global

Dada a natureza global da tecnologia de IA – desde os dados que utiliza para formação, até a pesquisa e o desenvolvimento, infraestruturas informáticas e aplicativos que atravessam fronteiras – é claro que nenhum governo pode abordar satisfatoriamente a política e a regulamentação de IA de forma isolada. A cooperação em nível internacional é essencial para garantir que os indivíduos e as sociedades em todo o mundo possam contar com os benefícios de uma IA confiável e responsável e que novos riscos sejam avaliados e mitigados de forma contínua. Este trabalho se beneficiaria de um fórum internacional dedicado que permitisse às partes interessadas governamentais e outras cooperar na política de IA.

Além disso, a cooperação internacional deve promover a interoperabilidade das políticas e regulamentações sobre IA. Tal como o CIPL observou no contexto da proteção de dados, a interoperabilidade global permite a prestação responsável de serviços transfronteiriços, amplia o acesso, reduz os custos de conformidade, aumenta a segurança jurídica e assegura uma proteção consistente dos direitos e interesses dos indivíduos[xxxi]. Diferentes jurisdições terão suas próprias prioridades, tradições jurídicas e conjunto de regulamentações existentes, mas podem ser capazes de se unir em torno de princípios e abordagens fundamentais ao considerar a política e regulamentação da IA – semelhantes aos que o CIPL promoveu neste documento. Podem também tomar medidas para codificar a interoperabilidade através de mecanismos de reconhecimento e certificação, nomeadamente através da participação no sistema Global de Regras de Privacidade Transfronteiriças (CBPR) no contexto da proteção de dados e dos fluxos de dados transfronteiriços confiáveis[xxxii]. Tem havido esforços encorajadores em prol da interoperabilidade da IA através da iniciativa do G7 acima mencionada, dos Princípios de IA da OCDE, de acordos comerciais e econômicos, como o Acordo de Parceria Econômica Digital (DEPA)[xxxiv], além da Parceria Global em IA[xxxv].

ANEXO I – ESTRUTURA DE RESPONSABILIDADE DO CIPL

ANEXO II – MAPEAMENTO DE MELHORES PRÁTICAS DE GOVERNANÇA DE IA PARA A ESTRUTURA DE RESPONSABILIDADE DO CIPL

A tabela a seguir descreve exemplos de atividades responsáveis de IA realizadas por organizações selecionadas de diferentes setores, geografias e tamanhos, com base na Estrutura de Responsabilidade do CIPL e mapeadas em relação a cada elemento de responsabilidade. As práticas não pretendem ser padrões obrigatórios do setor, mas servem como exemplos específicos que são calibrados com base nos riscos, contexto do setor, modelo de negócios, tamanho e nível de maturidade das organizações.

ELEMENTO DE RESPONSABILIDADE	**PRÁTICAS RELACIONADAS**
Liderança e Supervisão	• Compromisso público e manifestação da liderança a fim de respeitar a ética, os valores e os princípios específicos no desenvolvimento, implantação e uso de IA • Processos de IA institucionalizados e tomada de decisão com critérios de dimensionamento • Conselhos e Comitês de IA/Ética/Supervisão (internos ou externos) - para revisar casos de uso arriscados de IA e melhorar continuamente as práticas de IA • Nomeação de um membro do conselho para supervisão de IA • Nomeação de um líder responsável de IA, agente de IA ou defensor de IA • Criação de um conselho interdisciplinar interno de IA ou comitê de IA • Determinação da inclusão e da diversidade no desenvolvimento de modelos de IA e nas equipes de produtos de IA
Avaliação de Risco	• Avaliação de impacto algorítmico ou ferramentas de avaliação de equidade para monitorar e testar continuamente algoritmos para evitar preconceitos humanos, discriminação injusta e desvio de conceito ao longo de todos os ciclos de vida de IA • Avaliação de impacto ético/avaliação de impacto nos direitos humanos/avaliação de impacto na proteção de dados • Desenvolvimento de metodologias padronizadas de avaliação de risco, que levem em consideração os benefícios e a probabilidade e gravidade dos fatores de risco para os indivíduos e/ou sociedade, o nível de supervisão humana envolvida em decisões individualmente automatizadas com efeitos legais, bem como a sua explicabilidade de acordo com o contexto e auditabilidade • Documentação de compensações (por exemplo, precisão – minimização de dados, segurança – transparência, impacto em poucos – benefício para a sociedade) para processamento de alto risco como parte da avaliação de risco • Avaliação da qualidade dos dados por meio de KPIs • Avaliação de dados em relação ao propósito - qualidade, procedência, pessoal ou não, fontes sintéticas, internas ou externas • Estrutura para preparação de dados e avaliação de modelo – incluindo engenharia de recursos, validação cruzada, backtesting, KPIs validados por empresas • Estreita colaboração entre especialistas de negócios e de dados (analistas de dados, engenheiros de dados, engenheiros de TI e de software) para avaliar regularmente as necessidades e os resultados de precisão para garantir que o modelo possa ser usado corretamente

Políticas e Procedimentos	• Adoção de políticas e procedimentos específicos de IA sobre como projetar, usar ou vender IA • Políticas sobre a aplicação de privacidade e segurança desde a concepção no ciclo de vida da IA • Regra que define o nível de verificação de entrada e saída de dados • Teste piloto de modelos de IA antes do lançamento • Uso de dados protegidos (por exemplo, dados criptografados, pseudonimizados, tokenizados ou sintéticos) em alguns modelos • Uso de conjuntos de dados menores, mas de alta qualidade • Uso de modelos federados de aprendizagem de IA, considerando a compensação entre segurança de dados e responsabilidades do usuário • Considerações especiais para organizações que criam e vendem modelos, software e aplicativos de IA • Listas de verificação ou ferramentas de devida diligência/autoavaliação para parceiros de negócios que usam IA

	• Definição de etapas de dimensionamento em relação a relatórios, governança e análise de risco

	• Fase de ideação entre todas as partes interessadas (cientistas de dados, negócios, usuário final, funções de controle) em que são discutidos necessidades, resultados, regras de validação, manutenção, necessidade de explicabilidade, orçamento
Transparência	• Diferentes necessidades de transparência para indivíduos, reguladores, parceiros de negócios e internamente nas diferentes fases do ciclo de vida de IA com base no contexto • Divulgações adequadas comunicadas de maneira simples e fácil de se entender • Consideração de que a IA deve ser inclusiva e acessível para pessoas com necessidades especiais/deficiências • Estabelecimento de uma trilha de transparência para explicabilidade de decisões e amplo funcionamento do algoritmo para tornar o sistema de IA auditável • Explicação sobre o que é uma decisão de IA/ML, se houver possibilidade de confusão (teste de Turing) • Fornecimento de informações contrafactuais • Compreensão das expectativas dos clientes e implementação com base na sua prontidão para adotar IA • Implementação de transparência em níveis • Da caixa preta à caixa de vidro – observando os dados, bem como o algoritmo/modelo • A aspiração de explicabilidade ajuda a compreender a caixa preta e cria confiança • Definição de critérios de implantação de tecnologias de IA dentro da organização com base em cenários de uso e comunicação ao usuário • Produção de cartões modelo (documentos curtos que acompanham os modelos de IA para descrever o contexto em que o modelo deve ser usado, qual é o procedimento de avaliação) • Centro de dados para transparência sobre governança de dados, acessibilidade de dados, linhagem de dados, modificação de dados, qualidade de dados, definição etc. • Adaptação da transparência ao risco identificado: por exemplo, marca d'água para produção generativa de IA
Treinamento e Conscientização	• Treinamento de cientistas de dados, incluindo como limitar e lidar com preconceitos • Treinamento multifuncional – profissionais e engenheiros de privacidade • Treinamento de ética e equidade para equipes de tecnologia • Uso de casos em que a implantação problemática de IA foi interrompida • Papel dos “tradutores” nas organizações, explicando o impacto e o funcionamento da IA
Monitoramento e Verificação	• Capacidade para humanos envolvidos no projeto, na supervisão e na reparação • Capacidade de compreensão humana dos negócios e processos usando IA • Capacidade de auditoria humana de entrada e saída • Capacidade de revisão humana de decisões individuais com efeitos legais • Monitoramento do ecossistema desde o fluxo de dados, processo de dados e fluxo de dados • Confiança em diferentes técnicas de auditoria • Confiança em técnicas de testes contrafactuais • Pré-definição de controles de auditoria de IA • Equipe de auditoria interna especializada em IA e outras tecnologias emergentes • Os processos devem permitir o controle ou a intervenção humana no sistema de IA sempre que for tecnicamente possível e razoavelmente necessário • Monitoramento do modelo (backtesting e ciclo de feedback) e processo de manutenção
Resposta e Aplicação	• Processos e procedimentos para receber e tratar comentários e reclamações • Mecanismos de reparação para remediar uma decisão de IA • Reparação por um humano, não um bot • Canal de Feedback

[i] Para este relatório, o CIPL utiliza o termo "inteligência artificial" de uma forma consistente com a definição de "sistemas de IA" desenvolvida pelo Instituto Nacional de Padrões e Tecnologia dos EUA (NIST) no seu Quadro de Gestão de Risco 1.0, adaptada de uma definição comparável desenvolvida pela Organização para a Cooperação e Desenvolvimento Econômico (OCDE): "Um sistema de IA [é referido] como um sistema projetado ou baseado em máquina que pode, para um determinado conjunto de objetivos, gerar resultados como previsões, recomendações ou decisões que influenciam ambientes reais ou virtuais. Os sistemas de IA são projetados para operar com vários níveis de autonomia". Ver https://nvlpubs.nist.gov/nistpubs/ai/NIST.AI.100-1.pdf.

[ii] Por exemplo: A "Garante", Autoridade Italiana de Proteção de Dados, baniu o ChatGPT em 30 de março de 2023, disponível em https://www.gpdp.it/web/guest/home/docweb/-/docweb-display/docweb/9870832 ; o Comitê Europeu para a Proteção de Dados (EDPB) cria força-tarefa sobre o Chat GPT, disponível em https://edpb.europa.eu/news/news/2023/edpb-resolves-dispute-transfers-meta-and-creates-task-force-chat- gpt_en; A Autoridade de concorrência e mercado do Reino Unido lança revisão inicial de modelos de inteligência artificial, disponível em https://www.gov.uk/cma-cases/ai-foundation-models-initial-review ; A Agência Canadense de Privacidade de Dados (OPC) lançou uma investigação sobre o ChatGPT, disponível em https://www.priv.gc.ca/en/opc-news/news-and-announcements/2023/an_230404/; Declaração conjunta sobre os esforços de aplicação contra discriminação e preconceito em sistemas automatizados da Agência de Proteção aos Consumidores de Produtos Financeiros dos EUA (CFPB), Divisão de Direitos Civis do Departamento de Justiça, Comissão de Oportunidades Iguais de Emprego (EEOC) e Comissão Federal de Comércio (FTC), disponível em https://www.ftc.gov/system/files/ftc_gov/pdf/EEOC-CRT-FTC-CFPB-AI-Joint- Statement%28final%29.pdf.

[iii] De acordo com a OCDE, mais de 800 iniciativas e estratégias sobre políticas referentes à IA foram elaboradas em 69 países, territórios e na União Europeia https://oecd.ai/en/dashboards/overview.

[iv] CIPL, "First Report: Artificial Intelligence and Data Protection in Tension" (Primeiro relatório: inteligência artificial e proteção de dados em conflito), outubro de 2018, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_first_ai_report_-_ai_and_data_protection_in_tension 2_.pdf; "Segundo Relatório: Hard Issues and Practical Solutions (Questões difíceis e soluções práticas)," fevereiro 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_report_-_artificial_intelligence_and_data_protection_-_hard_issues_and_practical_solutions 27_february_2020_.pdf; "Artificial Intelligence and Data Protection: How the GDPR Regulates AI" (Inteligência artificial e proteção de dados: como a GDPR regula a IA), março de 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note-how_gdpr_regulates_ai 12_march_2020_.pdf.

[v] CIPL, "Response to NTIA Request for Comment on AI Accountability Policy" (Resposta à solicitação da NTIA de comentário sobre a política de responsabilidade em IA), junho de 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_ntia_ai_accountability_policy_june2023.pdf; "CIPL's Top Ten Recommendations for Regulating AI in Brazil" (As principais dez recomendações do CIPL para regulamentar a IA no Brasil), outubro de 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/[en]_cipls_top_ten_recommendations for_regulating_ai_in_brazil 4_october_2022_.pdf; "Response to UK DCMS Proposed Approach to Regulating AI" (Resposta à abordagem proposta pela DCMS do Reino Unido sobre a regulamentação da IA), setembro de 2022, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_uk_dcms_proposed_approach_to_regulating_ai_23_09_22.pdf; "CIPL Response to the EU Commission's Consultation on the Draft AI Act" (Resposta do CIPL à consulta da comissão da UE em relação ao projeto de lei sobre IA), julho de 2021, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_consultation_on_the_draft_ai_act 29_july_2021_.pdf.

[vi] Não existem definições universalmente aceitas para desenvolvedores, implantadores e usuários de IA; na verdade, uma taxonomia de 2023 desenvolvida em conjunto pela UE e pelos EUA descreveu as definições de implantação, desenvolvedor e usuário como "pendentes" (ver "EU-U.S. Terminology and Taxonomy for Artificial Intelligence: First Edition" (Terminologia e Taxonomia UE-EUA para Inteligência Artificial: primeira edição), https://www.nist.gov/system/files/documents/noindex/2023/05/31/WG1%20AI%20Taxonomy%20and%20Terminology%20Subgroup%20List%20of%20Terms.pdf). Neste artigo, usamos o termo "desenvolvedores" para nos referirmos às partes que projetam e constroem sistemas de IA, "implantadores" como partes que disponibilizam tais sistemas para uso e "usuários" como os usuários finais que operam esses sistemas em uma base contínua. Uma única entidade poderia desempenhar cada uma destas funções em pontos diferentes ou simultaneamente.

[vii] *K-Nearest Neighbors Algorithm* (K-vizinhos mais próximos-KNN), IBM, disponível em https://www.ibm.com/uk- en/topics/knn#:~:text=Next%20steps-,K%2DNearest%20Neighbors%20Algorithm,of%20an%20individual%20data%20point.

[viii] *3 ways autonomous farming is driving a new era of agriculture* (3 maneiras pelas quais a agricultura autônoma está impulsionando uma nova era da agricultura), Fórum Econômico Mundial, 2022, disponível em: https://www.weforum.org/agenda/2022/01/autonomous-farming-tractors-agriculture/

[ix] CIPL, "Artificial Intelligence and Data Protection: How the GDPR Regulates AI" (Inteligência artificial e proteção de dados: como a GDPR regula a IA), março de 2020, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_how_gdpr_regulates_ai 12_march_2020_.pdf.

[x] Para saber mais sobre a intersecção entre IA e a regulamentação sobre proteção de dados, ver CIPL AI First Report - *Artificial Intelligence and Data Protection in Tension* (Primeiro relatório do CIPL sobre IA - Inteligência artificial e proteção de dados em conflito); CIPL AI Second Report - *Hard Issues and Practical Solutions (Segundo relatório do CIPL sobre IA - questões difíceis e soluções práticas)*; e CIPL/Hunton Andrews Kurth White Paper - *How the GDPR Regulates AI* (Relatório branco CIPL/Hunton Andrews Kurth - como *a GDPR regulamenta a IA*) – todos disponíveis aqui https://www.informationpolicycentre.com/ai-project.html.

[xi] Para saber mais sobre este tópico, ver CIPL, "First Report: Artificial Intelligence and Data Protection in Tension" (Primeiro relatório: inteligência artificial e proteção de dados em conflito), outubro de 2018, cipl_first_ai_report_-_ai_and_data_protection_in_tension 2_.pdf (informationpolicycentre.com), e CIPL, "Second Report: Hard Issues and Practical Solutions" (Segundo relatório: questões difíceis e soluções práticas)," fevereiro de 2020, AI Project - Centre for Information Policy Leadership (informationpolicycentre.com).

[xii] Ministerial Declaration The G7 Digital and Tech Ministers' Meeting 30 April 2023 (g7digital-tech-2023.go.jp)

[xiii] G7 Hiroshima AI Process: G7 Digital & Tech Ministers Statement" (Processo de IA do G7 em Hiroshima: declaração dos ministros digitais e de tecnologia do G7), setembro de 2023, acessado em 3e39b82d- 464d-403a-b6cb-dc0e1bdec642-230906_Ministerial-clean-Draft-Hiroshima-Ministers-Statement68.pdf (politico.eu).

[xiv] Grupo de Trabalho de Certificação, "Unlocking the Power of AI – Steps for Effective Certification to Help Drive Innovation and Trust" (Desbloqueando o poder da IA – etapas de uma certificação eficaz para ajudar a impulsionar a inovação e a confiança), junho de 2023, https://www.responsible.ai/post/white-paper-draft-from-the-certification-working-group.

[xv] Para uma discussão adicional sobre a governança entre estas camadas do Stack Tecnológico da IA, ver Microsoft, "Governing AI: A Blueprint for the Future" (Governando a IA: uma esquema para o futuro), maio de 2023, https://query.prod.cms.rt.microsoft.com/cms/api/am/binary/RW14Gtw.

[xvi] Instituto Nacional de Padrões e Tecnologia (NIST), "AI Risk Management Framework" (*Estrutura de gestão de risco de IA*), https://www.nist.gov/itl/ai-risk-management-framework; Comissão de Proteção de Dados Pessoais (PDPC), "Singapore's Approach to AI Governance" (Abordagem de Singapura sobre a Governança de IA), https://www.pdpc.gov.sg/Help-and-Resources/2020/01/Model-AI- Governance-Framework, CIPL, "Organizational Accountability" (Responsabilidade organizacional), https://www.informationpolicycentre.com/organizational-accountability.html.

[xvii] Ver o relatório do CIPL sobre tecnologias de proteção à privacidade (a ser publicado, outono de 2023).

[xviii] Incentivizing Accountability: How Data Protection Authorities and Law Makers Can Encourage Accountability (Incentivando a responsabilidade: como legisladores e autoridades de proteção de dados podem incentivar a responsabilidade), CIPL, 23 de julho de 2018, disponível em https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_paper_2_-_incentivising_accountability_-how_data_protection_authorities_and_law_makers_can_encourage_accountability.pdf..

[xix] A Autoridade da Concorrência e Mercados (CMA), o Gabinete do Comissário de Informação (ICO), o Regulador britânico das comunicações (Ofcom) e a Autoridade de Conduta Financeira (FCA) participam do DRCF. Ver https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum. Para saber mais sobre o campo de atuação de transparência algorítmica do DRCF, ver https://www.gov.uk/government/publications/transparency-in-the-procurement-of-algorithmic-systems- findings-from-our-workshops.

[xx] Nos Países Baixos, a Autoridade dos Consumidores e dos Mercados (ACM), Autoridade Holandesa para Proteção de Dados (AP), a Autoridade Holandesa para Mercados Financeiros (AFM) e a Autoridade Holandesa de Mídia (CvdM) lançaram o Fórum de Cooperação para Regulação Digital (SDT). Ver https://www.acm.nl/en/about-acm/cooperation/national-cooperation/digital-regulation-cooperation-platform-sdt. O Centro Francês de Especialização em Regulamentação de Plataformas Digitais (PEReN) foi formado sob a autoridade dos Ministérios da Economia, Cultura e Tecnologia Digital. Com base na experiência em ciência de dados, é uma fonte de conhecimento técnico e apoio aos reguladores digitais da França. Ver https://www.peren.gouv.fr/en/. O Fórum de Reguladores de Plataforma Digital da Austrália reúne a Comissão Australiana de Concorrência e Consumidores (ACCC), a Autoridade Australiana de Comunicações e Mídia (ACMA), o Comissário de Segurança Eletrônica (eSafety) e o Escritório do Comissário Australiano de Informações (OAIC). Ver https://www.accc.gov.au/about-us/media/media-

updates/communique-digital-platforms-regulators-forum. A Irlanda criou a Rede de Reguladores Econômicos, que reúne sete reguladores. Ver https://www.econreg.ie/about/our-members/.

[xxi] Ver Christopher Hodges e CIPL, "Organizational Accountability in Data Protection Enforcement" (Responsabilidade organizacional na aplicação da proteção de dados), outubro de 2021, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_organizational_accountability_in_data_protection_enforcement_-_how_regulators_consider_accountability_in_their_enforcement_decisions_6_oct_2021_.pdf

[xxii] Ver o artigo do CIPL "Regulatory Sandboxes in Data Protection – Constructive Engagement and Innovative Regulation in Practice" (Sandboxes regulatórias na proteção de dados – engajamento construtivo e regulamentação inovadora na prática) 8 de março de 2019 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protectionconstructive_engagement_and_innovative_regulation_in_practice 8_march_201 9_.pdf.

[xxiii] Sandbox Regulatória do ICO https://ico.org.uk/for-organisations/regulatory-sandbox/.

[xxiv] Sandbox regulatória de dados IMDA, https://www.imda.gov.sg/how-we-can-help/data-innovation/data- regulatory-sandbox.

[xxv] Sandbox regulatória de IA da Datatilsynet, disponível em https://www.datatilsynet.no/en/regulations-and-tools/sandbox-for-artificial-intelligence/.

[xxvi] Sandbox sobre privacidade desde a concepção e padrão em projetos de Inteligência Artificial, Superintendência de Indústria e Comércio Colombiana, disponível em https://globalprivacyassembly.org/wp-content/uploads/2021/07/B6.-SIC-Colombia-Sandbox-on-privacy-by-design-and-by-default-in-AI-projects.pdf

[xxvii] CNIL, "Digital health and EdTech: the CNIL publishes the results of its first 'sandboxes'", julho de 2023, https://www.cnil.fr/en/digital-health-and-edtech-cnil-publishes-results-its-first-sandboxes.

[xxviii] No momento em que este artigo foi escrito, os decisores políticos da UE ainda não tinham decidido se tornariam voluntária ou obrigatória a criação de sandboxes de IA pelos Estados-Membros. Ver Luca Bertuzzi, "EU Council sets path for innovation measures in AI Act's Negotiations" (Conselho da UE define caminho para medidas de inovação nas negociações da Lei de IA), *Euractiv,* 10 de julho de 2023. https://www.euractiv.com/section/artificial-intelligence/news/eu- council-sets-path-for-innovation-measures-in-ai-acts-negotiations/. Ver também "First Regulatory Sandbox on Artificial Intelligence Presented" (Primeira sandbox regulatória sobre inteligência artificial apresentada) *Digibyte,* 27 de julho de 2022, https://digital-strategy.ec.europa.eu/en/news/first- regulatory-sandbox-artificial-intelligence-presented.

[xxix] Ver "Introducing Open Loop, a global program bridging tech and policy innovation" (Apresentando o Open Loop, um programa global reconciliando tecnologia e inovação em políticas), disponível em https://ai.facebook.com/blog/introducing-open-loop-a-global-program-bridging-tech-and-policy-innovation /; AI Impact Assessment: A Policy Prototyping Experiment (Avaliação do impacto da IA: um experimento de prototipagem de políticas), https://papers.ssrn.com/sol3/Delivery.cfm/SSRN_ID3772500_code715910.pdf?abstractid=3772500&mirid=1; and https://openloop.org/programs/open-loop-eu-ai-act-program/.

[xxx] IMDA, "Policy Prototyping" (Prototipagem de políticas) Policy Prototyping | IMDA - Infocomm Media Development Authority. O TTC Labs da Meta é parceiro do programa IMDA.

[xxxi] CIPL, "Ten Principles for a Revised US Privacy Framework" (Dez princípios para uma estrutura de privacidade dos EUA revisados), março de 2019, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_principles_for_a_revised_us_privacy_framework.pdf.

[xxxii] CIPL, "International Data Flows: Cross Border Privacy Rules, Privacy Recognition for Processors, and Global CBPR and PRP: Frequently Asked Questions" (Fluxos de dados internacionais: regras de privacidade transfronteiriças, reconhecimento de privacidade para processadores, e CBPR e PRP globais: perguntas frequentes), junho de 2023, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_cpbr_and_prp_faq_jun23.pdf.

[xxxiii] OCDE, "Recommendation of the Council on Artificial Intelligence" (Recomendação do Conselho de Inteligência Artificial), maio de 2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449

[xxxiv] O Acordo de Parceria Econômica Digital (DEPA) é um acordo entre Nova Zelândia, Singapura e Chile. Ver https://www.mfat.govt.nz/en/trade/free-trade-agreements/free-trade-agreements-in-force/digital-economy partnership-agreement-depa/.

[xxxv] Parceria Global em Inteligência Artificial, https://gpai.ai/.