

Sumário

Agradecimentos.....	5
Prefácio.....	7
A orquestração da confiança na Era da Inteligência Artificial	7
DORA KAUFMAN	
Introdução.....	15
1. Inteligência Artificial como tecnologia de propósito geral: correspondência, conceitos e definições	21
1.1. Tecnologia de propósito geral (<i>general purpose technology</i>) e inteligência artificial.....	22
1.1.1. A Inteligência artificial como GPT do século XXI	24
1.2. Inteligência artificial: conceitos e definições.....	30
1.2.1. (in)Definição do conceito de IA.....	30
1.2.2. Implicações do conceito de IA no plano regulatório.....	34
1.2.3. <i>Providers</i> (desenvolvedores) e <i>deployers</i> (implementadores/aplicadores)	45
2. Da lógica e conhecimento ao aprendizado e criação	53
2.1. Sistemas especialistas: baseados em lógica e conhecimento.....	57
2.2. Sistemas baseados em aprendizado e geração: a aprendizagem de máquina (<i>machine learning</i>) e suas técnicas.....	62
2.2.1. <i>Machine learning</i> (ML)	63
2.2.2. Redes neurais profundas (<i>deep learning</i>).....	85
2.2.3. IA preditiva e IA generativa.....	95
2.3. <i>Narrow Narrow AI</i> , AGI e superinteligência.....	102

3.	Ética e Inteligência Artificial.....	107
3.1.	Inteligência artificial e agência moral.....	107
3.1.1.	Agência e responsabilidade moral.....	117
3.1.2.	Agência moral e responsabilidade jurídica.....	118
3.1.3.	Previsibilidade e dever de prevenção.....	122
3.2.	Vieses e algoritmos.....	125
3.2.1.	Vieses (<i>bias</i>) no contexto da IA.....	126
3.2.1.1.	No sentido de segmentação e distinção de tendência ou desvios.....	126
3.2.1.2.	Em relação às suas categorias principais.....	130
3.2.1.3.	Em relação aos seus aspectos negativos e prejudiciais.....	132
3.2.1.4.	Em relação aos seus efeitos positivos para equidade e justiça.....	139
3.2.2.	Desafios para a implementação dos princípios de justiça e equidade.....	146
3.2.3.	Intervenções nas diferentes etapas do processo de design dos sistemas de IA.....	162
4.	<i>Ethics of AI, Responsible AI e AI Safety</i> como Agendas para a governança e gestão de riscos.....	167
4.1.	<i>Ethics for AI</i>	167
4.2.	<i>Responsible AI</i>	175
4.3.	<i>AI Safety</i>	190
4.3.1.	O problema do controle (<i>control problem</i>).....	197
4.4.	Explicabilidade e interpretabilidade.....	204
4.4.1.	<i>Explainable AI (xAI)</i>	215
4.4.2.	Interpretabilidade.....	234
4.4.2.1.	Interpretabilidade e causalidade: o processo cognitivo, o processo social e o produto da explicação.....	236
4.4.2.2.	Semiótica e design da informação aplicados à interpretabilidade.....	242

4.4.2.3. Exemplo de <i>framework</i> para a aplicação dos conceitos de explicabilidade e interpretabilidade.....	244
5. A regulação, as diretrizes e os <i>frameworks</i> de governança e gestão de riscos da IA.....	253
5.1. Impactos da IA e sua regulação.....	259
5.1.1. A integração de forças para a efetividade da regulação da IA....	265
5.2. <i>EU AI Act</i> : aspectos de governança e de gestão de riscos.....	271
5.2.1. Subsunção de um sistema de IA ao <i>EU AI ACT</i>	272
5.2.2. Sistemas de IA de alto risco.....	273
5.2.3. IA e os Setores de Aplicação (SAs): implementadores, desenvolvedores ou usuários corporativos?.....	277
5.2.4. Requisitos para sistemas de IA de alto risco.....	279
5.2.5. Obrigações de governança e gestão de riscos: desenvolvedores e implementadores.....	286
5.2.6. Modelos de IA de uso geral (<i>general-purpose AI models</i>).....	291
5.2.7. Estrutura de governança estatal (nível europeu e nacional) multinível.....	294
5.3. <i>Frameworks</i> de governança e gestão de riscos.....	298
5.3.1. <i>Process-Based Governance Framework</i> (PBG) – Alan Turing Institute.....	300
5.3.2. ISO/IEC 42001: 2023.....	303
5.3.3. <i>NIST AI Risk Management Framework</i> (AI RMF).....	305
5.3.3.1. As dimensões do AI RMF.....	308
5.3.3.2. As quatro funções principais (<i>core functions</i>).....	309
5.3.4. Análise comparativa dos <i>frameworks</i> analisados.....	314
6. Governança corporativa para a gestão de riscos e o <i>compliance</i> ético e legal da IA.....	319
6.1. Governança Corporativa.....	319
6.2. Estrutura de governança corporativa e IA.....	324

6.2.1. Estrutura organizacional da governança da IA: operacionalizada pela estrutura geral ou por um subconjunto especializado?	326
6.2.2. Estrutura organizacional da governança corporativa da IA: dimensões e caminhos para a implantação.....	331
6.2.3. Programa permanente de governança corporativa da IA: operacionalização e sustentação	338
6.2.3.1. Comitê de Ética de IA.....	343
6.2.4. <i>Ethical, Legal, Technological Blueprint for AI (ELT-AI)</i>	350
6.2.5. Estudo de casos.....	361
6.2.6. Benefícios e desafios da governança.....	376
7. Conclusão	383
Referências	389