



***TRS**

Tecnologia, Redes e Sociedade

Informação (gestão | espaços | ubíqua e inteligente)

Relatório Interno TRS 02/2026

Título

A simbiose necessária e o imperativo da infocomunicação

Autor(es)

Luís Borges Gouveia, UFP

Mês, Ano

Maio, 2026

Local de presença Web <http://tecnologiaredesesociedade.wordpress.com>

Repositório de trabalho científico *trs <http://bdigital.ufp.pt/handle/10284/3787>

Universidade Fernando Pessoa

Praça 9 de Abril, 349

4249-004 Porto, Portugal

Tabela de Conteúdos

Resumo.....	3
Abstract	3
1. Introdução: o contexto da sociedade da informação	4
2. A digitalização intensiva e a proliferação da IA generativa	4
3. Vulnerabilidades epistémicas e riscos na cibersegurança do dia-a-dia.....	5
4. O imperativo do fator humano: o modelo Human-in-the-Loop (HITL).....	6
5. Considerações finais.....	7
Transparência.....	7
Referências.....	7

A simbiose necessária e o imperativo da infocomunicação

O papel central do ser humano na governação de conteúdos da Inteligência Artificial Generativa

Luís Borges Gouveia, lmbg@ufp.edu.pt

Resumo

A transição para um quadro de Sociedade da Informação madura é pautada por uma digitalização contínua que reconfigura as práticas humanas e os modos de interação. O surgimento da Inteligência Artificial Generativa (IAG) potencia esta mutação, transformando os computadores em instrumentos universais de síntese informacional de grande escala. Contudo, a proliferação descontrolada de dados sintéticos gera riscos acrescidos para a segurança da informação, para ética e, desde logo, fiabilidade. Adotando uma perspetiva sistémica e infocomunicacional, é proposta a discussão dos impactos da IAG na atividade humana e é defendido que o conceito de Human-in-the-Loop (HITL) e a governação digital continuada são imperativos metodológicos e de cidadania indispensáveis para a validação de conteúdos e sustentabilidade do ecossistema informacional.

Palavras-chave: Inteligência Artificial Generativa; Sociedade da Informação; Validação Humana; Cibersegurança; Infocomunicação.

Abstract

The transition to a mature Information Society framework is driven by continuous digitalization that reconfigures human practices and modes of interaction. The emergence of Generative Artificial Intelligence (GenAI) amplifies this mutation, transforming computers into universal instruments for large-scale information synthesis. However, the uncontrolled proliferation of synthetic data creates heightened risks for information security, ethics, and, consequently, reliability. Adopting a systemic and infocommunication perspective, this report proposes a discussion on the impacts of GAI on human activity, arguing that the Human-in-the-Loop (HITL) concept and continuous digital governance are indispensable methodological and citizenship imperatives for content validation and the sustainability of the informational ecosystem.

Keywords: Generative Artificial Intelligence; Information Society; Human validation; cybersecurity; infocommunication.

1. Introdução: o contexto da sociedade da informação

A transição para a Web 3.0 (do mundo dos conteúdos para um contexto semântico) e a consolidação da economia digital culminaram numa viragem tecnológica marcada pela ubiquidade da Inteligência Artificial Generativa (IAG). Modelos de Linguagem de Grande Escala (LLMs) e sistemas de difusão de imagem transformaram o utilizador comum de um mero consumidor de conteúdos para um curador de ecossistemas automatizados. No entanto, a velocidade desta expansão superou a capacidade da sociedade de estabelecer mecanismos de filtragem analítica, exigindo um balanço crítico que coloque a tónica no fator humano e na infocomunicação (Floridi & Chiriatti, 2020).

A Sociedade da Informação e do Conhecimento caracteriza-se pelo recurso predominante ao digital para suportar a interação entre indivíduos e instituições, numa lógica de práticas em permanente construção (Gouveia & Gaio, 2004). Neste cenário de mudança contínua, as Tecnologias de Informação e Comunicação (TIC) deixaram de ser ferramentas de suporte para passarem a redefinir o nosso tempo, promovendo novos questionamentos sobre como aprendemos, comunicamos e nos organizamos (Gouveia, 2024b). A validação humana permanece, assim, um dos únicos pilares capazes de garantir a integridade factual e ética no espaço digital.

2. A digitalização intensiva e a proliferação da IA generativa

A exploração contemporânea do ambiente digital caracteriza-se pela automação da criatividade, pela representação de dados e pela interoperabilidade de sistemas que geram e distribuem fluxos informacionais de forma massiva (Cao et al., 2023; Gouveia et al., 2024). A IA generativa opera através do mapeamento probabilístico de vastos conjuntos de dados, permitindo a geração instantânea de texto, código, imagem e áudio.

Segundo uma perspetiva integradora da ciência da informação, este fenómeno altera os fluxos tradicionais de comunicação ao introduzir vetores opostos na sociedade:

- *Hiper-personalização e oportunidade: capacidade de moldar narrativas em tempo real para nichos específicos, descentralizando a capacidade criativa em contextos educativos, empresariais e de governação local;*
- *Democratização tecnológica: redução das barreiras técnicas para a criação de artefactos complexos e manipulação multimodo;*
- *Saturação informacional e entropia: produção massiva de conteúdos que desafia os motores de busca tradicionais e introduz o risco de degradação do espaço público digital devido à ausência de contexto cultural e ético nos algoritmos (Marcus, 2020).*

3. Vulnerabilidades epistémicas e riscos na cibersegurança do dia-a-dia

Apesar da aparente robustez, os sistemas de IA carecem de compreensão ontológica, consciência deontológica e de uma verdadeira ancoragem na realidade (grounding). Esta limitação técnica manifesta-se em três vetores críticos:

1. Alucinações tactuals (hallucinations): fenómeno em que o modelo gera afirmações falsas ou fictícias com elevado grau de assertividade gramatical (Ji et al., 2023);
2. Enviesamento algorítmico (bias): reprodução e amplificação de preconceitos históricos presentes nos dados de treino (Bender et al., 2021);
3. Diluição da propriedade intelectual e confiança: ambiguidade jurídica na reutilização de dados protegidos e impacto direto na perda de confiança dos cidadãos que consomem essa informação.

Na ótica da segurança da informação, o risco não se limita à falha do algoritmo, mas estende-se ao impacto da desinformação automatizada na tomada de decisão quotidiana (Gouveia, 2024c). À medida que os ecossistemas digitais passam a ser inundados por dados sintéticos, emerge o perigo da degradação progressiva do conhecimento. Quando modelos de IA passam a ser treinados com conteúdos gerados por outras IA, assiste-se a um colapso na qualidade da informação disponível (Shumailov et al., 2024). Garantir a integridade dos

dados tornou-se, por isso, uma atividade fulcral que exige uma postura ativa de cibersegurança e governança (Gouveia, 2024a; Gouveia et al., 2009).

4. O imperativo do fator humano: o modelo Human-in-the-Loop (HITL)

A arquitetura de segurança da informação e a governação digital no contexto da IAG não podem prescindir do julgamento humano. O conceito de Human-in-the-Loop (HITL) estabelece que a validação humana deve intervir em diferentes fases do ciclo de vida do conteúdo para mitigar os riscos de automação descontrolada (Amodei et al., 2016). O desenvolvimento tecnológico deve, obrigatoriamente, “salvaguardar a dimensão humana” (Gouveia et al., 2024), atuando a tecnologia como um extensor das capacidades humanas e não como o seu substituto acrítico.

Matriz de intervenção e Salvaguarda Humana

Dimensão	Atuação Humana	Objetivo Principal
Epistémica factual	/ Revisão por peritos, curadoria e <i>fact-checking</i> rigoroso.	Eliminar alucinações e garantir a verdade empírica (Zellers et al., 2019).
Ética deontológica	e Avaliação de impacto social, discurso de ódio e preconceito.	Alinhamento do conteúdo com os direitos humanos e valores sociais (Weidinger et al., 2021).
Estética contextual	/ Ajuste de tom, ironia e especificidades culturais.	Garantir a relevância e a empatia atendendo à realidade local e territorial (Gouveia & Gaio, 2004).

A validação humana atua como o filtro crítico que transforma o output estatístico da máquina em conhecimento estruturado e seguro para o consumo público. A componente humana na segurança da informação é a última e mais importante linha de defesa contra a erosão da verdade e a desumanização dos processos comunicacionais.

5. Considerações finais

A exploração do digital através da Inteligência Artificial Generativa representa um marco evolutivo inquestionável, potenciando a eficiência à escala global e oferecendo oportunidades extraordinárias para redefinir os negócios, a educação e a governação. Contudo, o sucesso desta transição depende de não colocarmos a ênfase apenas no artefacto tecnológico, mas sim no viver, no aprender e no interagir das pessoas.

A validação humana não deve ser encarada como um entrave à velocidade do desenvolvimento, mas sim como uma componente obrigatória de controlo de qualidade e salvaguarda ética. O futuro da comunicação digital depende, fundamentalmente, de uma simbiose equilibrada: a capacidade computacional e preditiva da máquina guiada, corrigida e validada pelo rigor ético, discernimento e responsabilidade do ser humano.

Transparência

O presente texto foi produzido com auxílio da IA Generativa, tendo sido revisto e idealizado pelo seu autor. A produção do texto base, foi realizada explorando um LLM genérico: o Gemini 1.5 pro (Google, <https://gemini.google.com>), a 21 de maio, como elemento de demonstração do potencial deste tipo de tecnologias, os seus limites e a forma como tirar proveito da IA como ferramenta e não como substituto cognitivo do ser humano.

Referências

- Amodei, D., Olah, C., Steinhardt, J., Christiano, P., Schulman, J., & Mané, D. (2016). Concrete problems in AI safety. *arXiv preprint arXiv:1606.06565*.
- Bender, E. M., Gebru, T., McMillan-Major, A., & Shmitchell, S. (2021). On the dangers of stochastic parrots: Can language models be too big? *Proceedings of the 2021 ACM FAccT Conference*, 610-623.
- Cao, Y., Li, S., Liu, Y., Yan, Z., Dai, Y., Yu, P. S., & Sun, L. (2023). A comprehensive survey of AI-generated content (AIGC): A history of generative AI from GAN to ChatGPT. *arXiv preprint arXiv:2303.04226*.

Floridi, L., & Chiriatti, M. (2020). GPT-3: Its nature, scope, limits, and consequences. *Minds and Machines*, 30(4), 681-694.

Gouveia, L. B. (2024a). *Conceitos e práticas para a cibersegurança como atividade do dia-a-dia*. Faculdade de Letras da Universidade do Porto / CITCEM.

Gouveia, L. B. (2024b). *O Digital e as tecnologias que estão a redefinir o nosso tempo e a promover novos questionamentos*. Atas de Investigação CITCEM.

Gouveia, L. B. (2024c). *IA, cibersegurança e seres humanos*. Atas do Encontro sobre Cultura Digital e Sociedade.

Gouveia, L. B., & Gaio, S. (Orgs.). (2004). *Sociedade da Informação: balanço e implicações*. Edições Universidade Fernando Pessoa.

Gouveia, L. B., Gouveia, F. R., Carvalho, C., Moreira, J., Ruivo, J., Neves, N., & Gomes, P. (2009). *Modelos de Governança na Sociedade da Informação e do Conhecimento*. Repositório Institucional UFP.

Gouveia, L. B., Marques, M. B., & Santos, M. (2024). Inteligência artificial como um caminho para a inteligência artificial generativa: um percurso que salvaguarde a dimensão humana. *Revista de Ciências da Informação e Comunicação*, CITCEM.

Ji, Z., Lee, N., Frieske, R., Yu, T., Su, J., Xu, B., Ishii, E., Bang, Y. J., Madotto, A., & Fung, P. (2023). Survey of hallucination in natural language processing. *ACM Computing Surveys*, 55(12), 1-38.

Marcus, G. (2020). The next decade in AI: four steps towards robust artificial intelligence. *arXiv preprint arXiv:2002.06177*.

Shumailov, I., Shumaylov, Y., Zhao, Y., Gal, Y., Papernot, N., & Anderson, R. (2024). AI models spit out creative nonsense when trained on AI-generated data. *Nature*, 631, 755-759.

Weidinger, L., Mellor, J., Rauh, M., Griffin, C., Uesato, J., Huang, P. S., ... & Gabriel, I. (2021). Ethical and social risks of harm from Language Models. *arXiv preprint arXiv:2112.04359*.

Zellers, R., Holtzman, A., Rashkin, H., Bisk, Y., Farhadi, A., Roesner, F., & Choi, Y. (2019). Defending against neural fake news. *Advances in Neural Information Processing Systems*, 32, 9054-9065.