AI DUE DILIGENCE IN M&A TRANSACTIONS: EMERGING LEGAL RISKS IN CROSS-BORDER DEALS

Dhruv Daruka & Sanjana Bhandarkar, Symbiosis Law School, Pune, Symbiosis International University

ABSTRACT

Traditional merger and acquisition due diligence was designed at a time when tangible assets and traditional forms of intellectual property made up most of the value. The increasing role of artificial intelligence as the primary value driver when considering a potential acquisition means that existing due diligence frameworks show structural flaws that expose an acquirer to significant amounts of unidentified liability arising from algorithms and protecting data supply that carries regulatory risk. Thus, the author contends that the unique nature of AI assets require a completely independent and distinct due diligence process, or framework, from standard forms of intellectual property, data, or other forms of technology assessments. Drawing from the EU AI Act, the General Data Protection Regulation, the Digital Personal Data Protection Act of 2023 (India), and NIST's Framework for Risk Management of AI, this article provides a model due diligence assessment matrix composed of seven elements: dataset verification, the governance of models, regulatory compliance mapping, assessments of explainability, assessment of bias, cybersecurity controls, and warranties in contracts. Ultimately, the author asserts that if acquirers do not have a comprehensive AI-specific due diligence process in place, they will find themselves with a category of hidden liabilities that cannot be identified through the normal due diligence process.

I. Introduction

Microsoft's acquisition of Activision Blizzard with an estimated value of $68.7 billion and Google's incorporation of DeepMind into their overall operating system are examples of significant changes in M&A rationale. Acquirers are increasingly looking for AI capabilities rather than just factories, brands or traditional forms of intellectual property.[1] Much of the value contained within these types of transactions is locked up in machine learning models, proprietary training datasets, algorithmic architectures and inference infrastructure that cannot be evaluated using traditional due diligence methods.[2]

This change is also significant from a legal perspective as AI systems present unique liability issues that do not fall within the standard due diligence performed on IP or data protection audits.[3] For instance, an artificial intelligence trained on large amounts of text data may have embedded copyright material without permission and will likely not represent the organization accurately. Similarly, an artificial intelligence recruitment tool could potentially produce years of discriminatory outcomes giving rise to legal risks which would not be evident from reviewing financial records. Likewise, medical devices may have been evaluated under regulatory approvals that cease to exist once the company is sold. The risks associated with these examples are not speculative; they represent latent liabilities which attach to the asset being acquired.[4]

The doctrinal issue is structural in nature. Norms regarding M&A due diligence typically address 'technology assets' through assessing software license agreements, patent portfolios, data ownership, and cybersecurity status. The conventional frameworks used to perform technology due diligence must evolve to capture the unique nature of artificial intelligence and how it will fit with existing laws governing liability for these types of transactions. The use of due diligence models to assess AI-heavy transaction risks has not been used to determine whether the training data corpora, the regulatory classification of algorithmic output, the explainability requirements for automated decision-making systems, and the compliance

[1] Tom Wheeler, Phil Verveer & Gene Kimmelman, New Digital Realities, New Oversight Solutions in the US (2020); Nico van Eijk, AI and M&A: Transactional Considerations in the Age of Machine Intelligence, 34 J. INT'L ECON. L. 1 (2021).
[2] Guild v. OpenAI, Inc., No. 23-cv-08292 (S.D.N.Y. filed Nov. 2023); Getty Images (US), Inc. v. Stability AI, Ltd., No. 23-135 (D. Del. filed Feb. 2023).
[3] Amazon Scraps Secret AI Recruiting Tool That Showed Bias Against Women, REUTERS (Oct. 10, 2018).
[4] Regulation (EU) 2024/1689 of the European Parliament and of the Council of 13 June 2024 laying down harmonised rules on artificial intelligence (EU AI Act) [2024] OJ L 1689.

burden associated with the EU AI Act's risk-mapping structure are appropriate in determining whether to purchase an AI asset. As a result, acquirers making AI-heavy acquisitions have assessed potential liabilities related to their AI-heavy acquisitions to virtually none until these liabilities are discovered after the acquisition is completed. This article has five total parts. The II section describes current due diligence frameworks as inadequate for assessing AI assets and how these frameworks will not provide sufficient protection for AI asset acquirers as AI technology evolves. The III section describes the various emerging regulatory frameworks applicable to AI technologies that provide AI asset acquirers with the identification of transaction-specific risks associated with acquiring AI assets (e.g., EU AI Act, GDPR, DPDP Act 2023, and the NIST AI RMF). Section IV provides example case studies of the risks identified in the second and third sections. Section V contains a model AI Due Diligence Matrix for common items that acquirers should use when assessing the risk of acquiring AI assets. The sixth section contains recommendations for reform.

II. The Structural Inadequacy of Conventional Due Diligence

Although due diligence in mergers and acquisitions usually includes financial, commercial, technical, and legal aspects, technology assets are normally evaluated for legal due diligence under three areas: ownership of intellectual property; compliance with data protection laws (including privacy notices, consent processes, and history of breaches); and compliance with contracts (including licensing agreements and any open-source dependencies). The due diligence framework assumes that there is an asset to be evaluated and can therefore be evaluated from the perspective of a transactional ontology based on discrete, finite value- and liability-bordered assets and that those assets can be valued and assigned liabilities based upon their documentation.[5]

AI systems will confuse this transactional ontology in at least three ways: their value exists in a distributed manner and is a function of the emergent property of the interplay between the data used to train the model, the model's architecture, the process of fine-tuning the model, and the infrastructure used to derive inferences from the model. Due diligence that considers only a company's model codebase without also evaluating the data pipeline through which the data to train the model was gathered, how the data was labelled, and the fine-tuning corpus will

[5] Dennis J. White, Due Diligence in Mergers and Acquisitions: A Practitioner's Guide (2020).

incorrectly assess the value of the model as an asset.[6]

In addition, risk and liability that existed in the past still carry forward for AI systems. Unlike a patent that has a defined ownership history and known scope, a trained model encoding patterns of discrimination that may have been produced during training; or violating copyrights through the through parameters that trained; or have violated privacy due to unlawful data processing that occurred during training. These types of historical liabilities often cannot be identified after they have occurred through a post hoc examination.[7] This is demonstrated by the commercial failures of IBM Watson Health — these algorithmic systems at deployment appeared to be technically functional; however, when they were deployed in a context outside what the data upon which they were trained allowed them to be classified as functionally viable, and thus caused them to be commercially and legally noncompliant.[8]

Finally, AI systems can also be re-classified upon acquisition of the AI system. The EU regulates the classification of AI systems by level of risk, therefore imposing different obligations upon "deployers" and "providers" which exist when compared to the original developer of the AI system.[9] When there is a change of control that reclassifies an AI system — because the new owner deploys the AI system in a higher risk context than the original developer — a compliant AI system can be rendered a non-compliant AI system overnight without any changes to the AI system.

III. The Emerging Regulatory Architecture and Its Transaction Implications

A. The EU AI Act

The European Union's AI Act, which became operational in August 2024, implements a risk-based regulatory framework that includes differences in compliance obligations based on how high-risk, limited-risk, and low-risk systems fall into each risk category. There are three areas of concern for M&A practitioners regarding the Act.[10]

[6] Lyria Bennett Moses & Janet Chan, Algorithmic Prediction in Policing: Assumptions, Evaluation, and Accountability, 28 POLICING & SOC'Y 806 (2018).
[7] Julia Angwin et al., Machine Bias, PROPUBLICA (May 23, 2016) (documenting embedded bias in recidivism prediction algorithms).
[8] Casey Ross & Ike Swetlitz, IBM's Watson Supercomputer Recommended 'Unsafe and Incorrect' Cancer Treatments, Internal Documents Show, STAT NEWS (July 25, 2018).
[9] EU AI Act, supra note 4, arts. 3(3), 3(4), 25.
[10] EU AI Act, supra note 4, arts. 5, 6–7 (Annex III), 50, and recitals 1–10.

The first area of concern is that the Act's obligations are placed on two functional roles; "Providers" and "Deployers." An acquirer becomes a Provider of high-risk AI if they accept the role of Provider from their acquisition target and gain full responsibility for the obligations under Article 16 of the Act[11], which include: conformity assessment; technical documentation; post-market monitoring; and registration in the EU AI database. If the target has not completed any of these obligations prior to acquisition, the acquirer may incur substantial costs associated with remediating compliance until it can fulfil the compliance obligations. If these obligations were not met by the target prior to acquisition, the acquirer could potentially face civil liability to third parties for negligent provision of the technology or service through the use of the high-risk AI system.

Second, the Act imposes specific requirements regarding the governance of training data for high-risk AI. Under Article 10[12], in order to qualify for high-risk AI classification, the training data must be relevant to the purpose for which it is being used, representative of the population it will serve, and must have as few errors and bias as technically possible. As a result, due diligence must consider the legality of the training data that was collected and must evaluate whether it meets the criteria established in the Act. Thus, due diligence will involve a much deeper level of inquiry than what is typically expected in M&A personnel when conducting due diligence on a target for the acquisition of its AI system.

Thirdly, the AI applications prohibited under Article 5, such as biometric identification in real-time and social scoring models, create a risk of unlimited liability.[13] If an acquirer determines after they have closed to purchase the AI system of a target company that it falls within the prohibited category of applications, they must cease operation of the system immediately; there is no way to cure the prohibition.

B. GDPR and Training Data Legality

In addition, the complexity of due diligence surrounding the intersection between the GDPR and AI training has created some of the most challenging due diligence questions for current transactional practice. Article 6 of the GDPR states that all processing of personal data must follow a lawful basis as defined by the GDPR. This includes processing for the purpose of

[11] EU AI Act, supra note 4, art. 16.
[12] EU AI Act, supra note 4, art. 10.
[13] EU AI Act, supra note 4, art. 5.

training AIs to process personal data.[14] If a dataset used to train an AI was compiled through web scraping or compiling data without a clear lawful basis, there is a risk that the AI model trained on the dataset may be deemed to have been processed unlawfully; there is no possible way to cure the risk of having processed the dataset unlawfully after it has been processed.

Several enforcement actions taken against Clearview AI by a number of EU data protection agencies demonstrates the potential scope of regulatory risk: in total, the four data protection agencies have collectively imposed fines of more than €50 million for violations of the GDPR because the datasets to train Clearview AI's facial recognition algorithm were compiled without a lawful basis.[15] When an acquirer purchases a similarly situated AI system, they are acquiring not only the technology, but any historical processing violations as well.

C. Digital Personal Data Protection Act 2023

While India's Digital Personal Data Protection Act, 2023 is not entirely operational yet (there is still some subordinate legislation to complete) there are responsibilities based on the DPDP for cross-border compliance that will have implications for AI M&A transactions with respect to either Indian data principals or Indian Data Fiduciaries pursuant to the Act.[16] For example, and especially for due diligence purposes, unlike under the GDPR where data controllers bear the compliance obligations, in the DPDP, data fiduciaries determine the purpose and means of processing data; therefore, a post-acquisition restructuring, where the acquirer becomes a data fiduciary for personal data collected from India, may require the acquirer to obtain new consent from data subjects, locally store the data within India, and notify the Data Protection Board in India.[17]

Because that the Central Government has discretion regarding limiting cross-border transfers from India to specific countries, the cross-border transfer provisions could create new transaction risk for an acquirer as it attempts to consolidate Indian data assets into a global infrastructure.[18]

[14] Regulation (EU) 2016/679 of the European Parliament and of the Council (GDPR) [2016] OJ L 119/1, art. 6.
[15] See Garante per la protezione dei dati personali v. Clearview AI (2022); ICO Penalty Notice re Clearview AI (2022).
[16] Digital Personal Data Protection Act 2023 (India), No. 22 of 2023.
[17] DPDP Act 2023, supra note 16, s. 2(i) (definition of 'data fiduciary'), s. 6 (consent).
[18] DPDP Act 2023, supra note 16, s. 16 (transfer of personal data outside India).

D. NIST AI RMF and OECD AI Principles

Although NIST's AI Risk Management Framework is not strictly required by law in the United States, it is becoming an increasingly important reference for enterprise AI governance, if only on a de facto basis, and is referenced in the contractual representations and warranties that are often associated with technology-based M&A transactions.[19] It is about four main functions: Govern, Map, Measure and Manage, and provides a practical framework for the due diligence teams to work through. The OECD principles on AI also include principles of accountability, transparency, and robustness that are increasingly subject to verification in cross-border mergers by the international acquirers.[20]

IV. ILLUSTRATIVE CASE STUDIES

A. Microsoft and OpenAI

The billion-dollar Microsoft investment in OpenAI, not a full buyout, is a good example of the risks of a business that is based on AI. The exclusive deployment and use of OpenAI models as cloud services would give Microsoft a substantial reputation risk with regards to the training data practices of OpenAI, and in particular, the possibility of future litigation based on the EU AI Act on the use of copyrighted material in OpenAI models' training corpus, which was not yet a material aspect of the investment but was a material aspect when the Act came into force.[21]

B. Google DeepMind

The retrospective regulatory analysis has been quite extensive with Google's acquisition of DeepMind in 2014 for $500 million. The fact that there are liabilities to be identified arising from an AI system's data practices (in this case patient data from the Royal Free NHS Trust) only became apparent years after acquisition.[22] It emphasizes that the processing risk was not identified even after conducting the most basic due diligence during acquisition, so that a liability can only be discovered years after the acquisition of the AI system.

[19] National Institute of Standards and Technology, Artificial Intelligence Risk Management Framework (AI RMF 1.0) (2023).
[20] OECD Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449 (2019).
[21] See The New York Times Company v. Microsoft Corporation, No. 23-cv-11195 (S.D.N.Y. filed Dec. 2023).
[22] ICO, Royal Free London NHS Foundation Trust: Undertaking (2017).

C. IBM Watson Health

IBM Watson Health is a good example of a commercial failure of AI asset valuation. IBM's healthcare AI acquisitions are based on the idea that Watson's natural language processing could be commercialized in clinical decision support.IBM's three healthcare AI acquisitions Merge Healthcare, Phytel and Explorys have been based on the notion that Watson's NLP technology could be commercialised in clinical decision support. The model's out-of-distribution performance has not and cannot be tested by the usual due diligence.[23] The out-of-distribution performance of the model was not and cannot be assessed by conventional due diligence.

D. Clearview AI

The enforcement record of Clearview AI demonstrates that the liability for training data is clearly an ideal case study. The legality of training cannot simply be assumed based upon the technical traits of the system. It is possible to create a high performing (i.e. highly accurate) system with data processed illegally, and will create a regulatory liability for any jurisdiction that deploys the high performing system.[24]

V. Proposed Matrix for AI Due Diligence

The matrix below provides a framework to analyse AI-specific due diligence according to seven operational pillars. The framework can be used as a supplement to traditional legal due diligence not as a replacement.

Pillar 1: Data Set Verification Conducting all due diligence teams must be able to: (1) establish the legal basis for the collection of each part of the training dataset; (2) confirm if data scraped off the web was collected in accordance with the relevant terms and conditions of the applicable data protection law; (3) confirm if the data collected from licenses have any restrictions on commercial usage, transfer, or fine-tuning; (4) trace the provenance of synthetic data collected; and (5) confirm if data subject rights (such as erasure) can be implemented after collection.[25]

Pillar 2: Model Governance Review for assessment purposes of model governance would

[23] Eliza Strickland, How IBM Watson Overpromised and Underdelivered on AI Health Care, IEEE SPECTRUM (Apr. 2, 2019).
[24] sources cited supra note 15
[25] GDPR, supra note 14, art. 17 (right to erasure).

cover: (i) model version control and audit logs; (ii) documentation of model training, hyperparameters, and benchmarks for testing; (iii) records of red teaming and adversarial testing; (iv) any logs concerning incidents or safety events related to the model; and (v) availability of any model cards or equivalents.

Pillar 3: Regulatory Compliance Mapping the EU AI Act mandates that acquisitions map all acquired AI systems to their risk category and assess whether conformity assessments have been conducted.[26] Under GDPR, high risk DPIA assessments must be reviewed. Under the DPDP Act 2023, it must be verified whether the entity is a data fiduciary, consent records have been maintained, and whether there is a DPO. Any CFIUS requirements will need to be considered.[27]

Pillar 4: Explainability Audits Where AI systems feature in decision-making with legal effect or similar significance, i.e., for decisions concerning credit scoring, insurance underwriting, hiring, or clinical triaging, explainability obligations under Article 22 GDPR apply. When fully enacted, the EU AI Act would also apply.[28]

Pillar 5: Bias Assessments Historical bias within the training data or model outputs may constitute discrimination that is illegal under employment law, consumer protection law, or financial services regulation. The due diligence process would need to include a review of bias testing procedures, the identification of protected characteristics, an analysis of adverse impacts, and any regulatory communications about algorithmic discrimination.[29]

Pillar 6: Cybersecurity Controls AI systems are uniquely vulnerable to specific forms of attack, such as model inversion, data poisoning, prompt injection, and membership inference, which necessitate an assessment specific to the AI's vulnerabilities rather than merely traditional penetration testing.[30] The NIST AI RMF's Manage function provides a systematic methodology for performing this assessment.

Pillar 7: Contractual Warranties and Indemnities Agreements in purchasing AI-rich targets would need to include: (i) representations about the legality of the training data; (ii) warranties

[26] EU AI Act, supra note 4, arts. 43–44.
[27] Foreign Investment Risk Review Modernization Act of 2018 (FIRRMA), Pub. L. No. 115-232.
[28] GDPR, supra note 14, art. 22.
[29] Sandra G. Mayson, Bias In, Bias Out, 128 YALE L.J. 2218 (2019).
[30] Nicholas Carlini et al., Extracting Training Data from Large Language Models (2021).

of compliance with relevant AI regulations; (iii) indemnifications in connection with pre-acquisition algorithmic discrimination claims; (iv) an escrow arrangement in line with the regulatory risk; and (v) material adverse change clauses due to AI regulations.[31]

VI. Conclusion and Reform Proposals

The purchase of AI-based assets via traditional due diligence procedures results in a structural liability gap, which means that substantial issues will arise post-acquisition when the EU AI Act, the DPDP Act 2023, and other regulatory frameworks take effect in full. The sale of IBM Watson Health assets, the chain of actions taken against Clearview AI, and lawsuits regarding large language model training data are all examples of how AI-related liability does not arise from document analysis alone.

Three reforms are suggested by this paper. To begin, regulatory agencies overseeing mergers and acquisitions should design AI-specific guidance for M&A transactions that highlights the types of AI-related risks requiring closer analysis. These regulatory bodies include the Competition Commission of India, the European Commission's DG COMP, and the United Kingdom's Competition and Markets Authority. In addition, the IBA's M&A Committee should develop a standard AI Due Diligence Protocol along the lines of its Environmental and Cybersecurity Protocols. Lastly, acquisitions involving an AI component that exceeds a certain percentage of deal value should be contractually required to provide AI governance assurances. The AI Due Diligence Matrix presented above provides a basic outline for practitioners who can no longer afford to regard their algorithmic assets as yet another kind of intellectual property.

[31] Glenn D. West & Kim M. Shah, Debunking the Myth of the MAC: Materiality Thresholds in M&A Transactions, 62 BUS. LAW. 1087 (2007).