

APPLICATION OF ARTIFICIAL INTELLIGENCE AND TECHNOLOGIES BIG DATA IN ESG MONITORING AND ANALYSIS

I.R. Berdikulova

SamIES, PhD

irodaberdikulova82@gmail.com

Abstract: This article examines the potential and limitations of artificial intelligence (AI) and Big Data technologies in ESG monitoring, analysis, and rating systems for companies. Specific AI applications in the ESG field are considered, including processing non-financial reports using NLP methods, satellite monitoring of environmental indicators, news sentiment analysis for the early detection of ESG risks, and predictive corporate governance models. Traditional and AI-based approaches to ESG ratings are compared in terms of data coverage, relevance, and objectivity of assessments. The international experience of leading ESG providers and fintech startups is analyzed. Prospects for implementing AI tools for ESG analysis in the financial and analytical sector of Uzbekistan are discussed.

Keywords: artificial intelligence, Big Data, ESG monitoring, NLP, machine learning, satellite data, alternative data, ESG rating, fintech.

Abstract: The article examines the capabilities and limitations of artificial intelligence (AI) and Big Data technologies in ESG monitoring, analysis, and corporate rating systems. Specific AI applications in the ESG domain are reviewed: processing of non-financial reports using NLP methods, satellite monitoring of environmental indicators, news sentiment analysis for early identification of ESG risks, and predictive models of corporate governance. Traditional and AI-driven approaches to ESG rating are compared in terms of data coverage, timeliness, and objectivity of assessments. International experience of leading ESG data providers and fintech startups is analyzed. Prospects for implementing AI-based ESG analysis tools in the financial analytics industry of Uzbekistan are examined.

Keywords: artificial intelligence, Big Data, ESG monitoring, NLP, machine learning, satellite data, alternative data, ESG rating, fin tech .

Introduction.

The quality and availability of ESG data has traditionally been a key bottleneck in sustainable investing: companies disclosed non-financial information in various formats, with varying levels of detail and verifiability, and the volume of data available for independent analysis was extremely limited. The technological revolution in AI and Big Data is fundamentally changing this situation, opening the possibility of analyzing unprecedented volumes of structured and unstructured data in real time.

According to consulting firm PwC (2022), the total market size for AI solutions for ESG analytics reached \$1.8 billion in 2022 and is projected to exceed \$9 billion by 2027, with a compound annual growth rate of 38%. This growth reflects both growing investor demand for

Journal: <https://ifsmrc.org/index.php/aijrm>

objective and timely ESG data and the decreasing cost of computing power and data storage, making AI solutions affordable for a wider range of users.

Traditional ESG ratings, based primarily on companies' self-disclosure (questionnaires, annual reports, sustainability reports), suffer from a number of systemic shortcomings: a significant time lag (data for the reporting year is published with a delay of 6-12 months); the risk of "greenwashing" due to the lack of independent verification; and limited coverage (small and medium-sized companies, as well as companies in jurisdictions with low disclosure requirements, are excluded from traditional rating systems). AI and Big Data technologies can overcome these limitations.

The purpose of this article is to systematize existing AI applications in the ESG field, assess their potential and limitations, analyze international experience, and develop recommendations for the use of these technologies in the context of the developing financial market of Uzbekistan.

Natural Language Processing (NLP), a branch of AI that analyzes and interprets text data, is widely used in the automation of ESG reporting. Key tasks include: automatic extraction of numerical ESG indicators from reports of any format (PDF, HTML, XBRL); classification and thematic modeling of ESG content; identification of discrepancies between stated objectives and actual results; and sentiment analysis of corporate communications on ESG topics.

Leading ESG providers are actively integrating NLP into their analytical processes. Sustainalytics (Morningstar) has developed a system for automatically monitoring over 70,000 news sources in 13 languages to identify ESG incidents in real time. MSCI uses NLP algorithms to speed up the processing of corporate reports: the analysis time for a single report has been reduced from several days (with manual processing) to just minutes. Fintech startup RepRisk uses NLP to process over 100,000 sources in 20 languages daily to identify reputational ESG risks.

The emergence of large-scale language models (LLMs) like GPT-4 and their specialized versions opens up entirely new possibilities for ESG analysis. In 2023, several financial institutions launched pilot projects using LLMs to analyze ESG reports, stress-test climate narratives, and generate sustainability summaries. Preliminary results demonstrate a significant increase in analyst productivity with comparable analysis quality. Remote sensing (RS) using satellite technology represents a fundamentally new source of objective ESG data that is independent of company self-disclosure. Modern commercial satellite systems enable monitoring a wide range of environmental indicators with a spatial resolution ranging from 30 centimeters to several meters and data update rates ranging from several hours to several days.

Specific ESG applications of satellite data include: monitoring methane emissions from oil and gas facilities (GHGSat and MethaneSAT satellites allow for the identification of emission sources down to the individual facility level—this technology is already used by investors to independently verify oil company emissions statements); monitoring deforestation and compliance with zero-deforestation requirements in agricultural and logging supply chains; monitoring nighttime light pollution as an indirect indicator of economic activity in hard-to-reach regions; and assessing the condition of agricultural land and water resources through spectral analysis.

Since 2017, the French startup Kayrros has been providing investors with satellite data on methane emissions from oil and gas companies, significantly supplementing and verifying the companies' official data. In a number of cases, independent satellite measurements have revealed actual emissions that exceed declared levels by several times—such discrepancies inevitably lead to a reassessment of ESG ratings and, consequently, the market value of the companies in question.

Alternative data—non-traditional sources of information not included in standard financial reporting—open up new opportunities for proactive (predictive) assessment of ESG risks. The most significant alternative sources for ESG analysis include: labor relations data from employee feedback platforms (Glassdoor, Indeed), which allow for real-time assessment of the S-component; consumer reviews and social media data, which are indicators of reputational risks and product liability; patent activity data, which is a proxy for assessing technological readiness for the low-carbon transition; and government procurement and litigation data, which are indicators of regulatory and governance risks.

Machine learning (ML) enables the identification of complex nonlinear relationships between alternative data and future ESG events. A 2022 study by Two Sigma Investments demonstrated that ML models trained on alternative data (employee reviews, news feeds, geolocation data) are capable of predicting major ESG incidents (industrial accidents, corruption scandals, environmental violations) with a lead time of 3-6 months and an accuracy of up to 72%, significantly exceeding the capabilities of traditional rating approaches.

Another important area is the use of AI to standardize and comparably evaluate ESG data. The problem of "ratings incomparability"—significant discrepancies in ESG ratings for the same companies from different providers (Berg, Kölbel & Rigobon, 2022, found an average correlation of only 0.54 between ratings from different agencies)—can be partially addressed by developing standardized aggregation algorithms that are applied consistently to data from different sources.

Uzbekistan's financial analytics industry is in its early stages of technological development: most analytical processes in banks and investment firms remain manual or semi-automated. However, a number of prerequisites for the implementation of AI-based ESG analysis tools have already been established. First, Uzbekistan has a growing pool of IT specialists: in 2023, the number of people employed in the ICT sector exceeded 40,000, and IT services exports reached \$300 million. Second, a number of government programs (Digital Uzbekistan, the IT Development Strategy until 2030) create institutional incentives for the adoption of advanced technologies.

Practical applications of AI for ESG analysis in Uzbekistan include: developing an NLP tool for automatically extracting ESG indicators from annual reports of Uzbek companies in Russian and Uzbek—this would significantly expand the scope of ESG analytics without a proportional increase in labor costs; integrating satellite data to monitor water consumption in the agricultural sector—relevant for assessing the S-component of cotton and textile enterprises; and creating a system for monitoring the ESG reputation of traditional companies based on analysis of local news feeds and social media. The Capital Market Development Agency (CMDA) is recommended to consider creating a centralized digital ESG platform for collecting, verifying, and publicly disclosing non-financial data from public companies, modeled on similar platforms on the Hong Kong Stock Exchange (HKEX ESG Data Platform) and the Singapore Stock Exchange (SGX). Such a platform would not only expedite the ESG disclosure process but also create a structured database for AI analysis.

Conclusion.

AI and Big Data technologies are transforming ESG analysis, overcoming key limitations of traditional approaches: time lag, reliance on self-disclosure, and limited coverage. NLP tools automate corporate reporting, satellite data provides independent verification of environmental indicators, and alternative data combined with machine learning models enables forecasting ESG events several months in advance. However, the use of AI in ESG analysis is associated with a number of limitations and risks: the "black box" problem in complex ML models complicates the

interpretation of results; biases in training data can reproduce and amplify systematic errors in traditional ratings; and language and cultural barriers limit the applicability of models trained primarily on English-language data in other language contexts.

List of references:

1. Uzbekcosmos Agency. Program for the Development of Earth Remote Sensing in Uzbekistan for 2022–2026. Tashkent, 2022.
2. Ministry of Digital Technologies of the Republic of Uzbekistan. Strategy "Digital Uzbekistan - 2030". Tashkent, 2020.
3. Two Sigma Investments. Alternative Data for ESG Prediction. — New York: Two Sigma, 2022.
4. MSCI. Using AI and Machine Learning in ESG Research. — New York: MSCI, 2022.
5. RepRisk AG. RepRisk ESG Data Science. — Zurich: RepRisk, 2023.
6. Kayrros SAS. Methane Watch: Satellite-Based Methane Monitoring. — Paris: Kayrros, 2023