A Comprehensive Systematic Review of Dual Use Research Classification using Artificial Intelligence

Babatunde I. Ayinla
Department of Data Science, Faculty of Computing, University of Ibadan, Ibadan, Oyo, Nigeria

Damilola Adeniniji
Department of Computer Science, Faculty of Computing, University of Ibadan, Oyo, Nigeria

ABSTRACT
Dual-use research (DUR) serves both scientific and practical purposes, yet it is susceptible to misuse for harmful applications. Effectively classifying DUR is essential to protect sensitive information while enabling beneficial research to continue. Artificial Intelligence (AI), particularly ML Classification Models (MLCMs), has emerged as a powerful tool for this complex classification challenge. This systematic review examines AI's effectiveness in identifying and classifying dual-use and SDUR, with emphasis on MLCMs and Natural Language Processing (NLP) techniques. Analysis of English-language articles from 2006–2023 across seven databases, including PubMed, Scopus, Google Scholar, arXiv, IEEE Xplore, and Springer Link, yielded 101 relevant studies after screening 8,037 initial results. The findings reveal widespread use of ML techniques, primarily Decision Trees, Naïve Bayes, Random Forests, and Support Vector Machines. NLP methods, especially Topic Modeling and Sentiment Analysis, were frequently employed. These models demonstrated effectiveness in classifying DUR, suggesting the potential for automated identification systems. However, the review identifies several areas for improvement, including data quality, standardization, and reporting transparency. While acknowledging DUR's contributions to scientific progress, the study emphasizes the importance of ethical considerations and safeguards when implementing automated classification systems.

General Terms
Dual-Use Research, Susceptible Dual-Use, Pattern Recognition, Security, Machine Learning.

Keywords
Machine Learning, Natural Language Processing, scientific text mining, Biosecurity, Ethical Considerations, Susceptible Dual-use Research.

1. INTRODUCTION
Dual-use research poses an increasing concern because of its capacity to serve both civilian and military objectives. Some research areas are more prone to misuse than others (National Institutes of Health [NIH], 2023), highlighting the importance of distinguishing between general DUR and those that are particularly vulnerable to misuse. Effective classification is crucial for safeguarding sensitive information while allowing legitimate scientific advancement. This review provides an overview of current methods for classifying DUR, emphasizing SDUR, and identifies areas for improvement.

1.1 Defining Dual-Use Research (DUR)
Dual-use research (DUR) denotes research that possesses the potential for both beneficial applications for humanity and harmful uses (Atlas & Dando, 2006). Such research is frequently linked to the life sciences. These endeavors aim for positive advancements in health, agriculture, AI, and other fields, yet they possess the potential for misuse in the creation of biological weapons, chemical threats, or other hazardous capabilities (Grinbaum, 2017). This inherent duality poses a considerable challenge, necessitating thorough risk assessment and mitigation strategies.

The concept of Dual-use Research (DUR) originated within the realm of nuclear energy, as researchers acknowledged that their findings could serve both non-military and military applications (Forge, 2010). The concept has since been applied across various fields, including biotechnology, AI, and advanced materials. Dual-use research provides various advantages, including the acceleration of scientific advancement, the promotion of collaboration, and the reduction of expenses. Nonetheless, it also presents ethical issues regarding the possible misuse of scientific knowledge and technologies.

1.2 Susceptible Dual-Use Research
Susceptible Dual-use Research (SDUR) denotes research with a significant potential for misuse by malicious entities. This encompasses research applicable to the development of weapons of mass destruction, cyberattacks, or other detrimental technologies. The emergence of CRISPR-Cas9 gene editing technology has generated apprehensions regarding its possible applications in bioterrorism or biowarfare. Advancements in Artificial Intelligence (AI) have raised concerns regarding autonomous weapons and cyber-attacks (National Institutes of Health [NIH], 2023).

1.3 Existing Methods of Dual Use Research Classification
The classification of DUR is critical to safeguarding sensitive information from potential misuse by malicious entities. Scientists have proposed several methods for classifying research as dual-use or non-dual-use.

1.3.1 Expert Judgment
Expert judgment is a systematic and organized method for integrating an expert's experience regarding a specific subject (Naseri et al., 2016). Expert judgment is frequently required when data regarding a technical issue is limited or absent, when the issue involves significant uncertainty, or when it is too complex to model precisely (Naseri et al., 2016). Experts in the field can evaluate research projects based on their potential impact and likelihood of misuse. This approach is often criticized for its subjectivity and lack of transparency (Guillemin, 2005).

1.3.2 Automated Content Analysis
Computer algorithms can examine literature to discern keywords and phrases related to DUR. This strategy may

neglect subtleties in language and context (National Institutes of Health [NIH], 2023). Automated approaches possess numerous advantages compared to conventional manual procedures: (1) they are more dependable, (2) they are more impartial, and (3) they conserve significant time and effort (Seager et al., 2007; Matthes & Kohring, 2008). However, critics assert that the technology is still underdeveloped.

1.3.3 Citation Analysis

Analyzing citation patterns can help identify research that builds upon previous DUR. However, this method may not capture emerging areas of research (Danzig, 2012). Patent documents, like scientific journals, typically include references or citations to illustrate the background or state of the art prior to the invention.

1.3.4 Machine Learning

Machine Learning (ML) algorithms can be trained to recognize patterns in data and predict the likelihood of DUR. These models can improve accuracy over time but require large datasets for training (Jones et al., 2021).

1.4 Potential Risks of DUR

The potential risks associated with DUR are broad and vary depending on the specific research conducted. Some key concerns include the development of biological or chemical weapons, bioterrorism, and economic disruption.

1.5 Importance of Effective DUR Classification

As new threats emerge, classification systems must adapt quickly to address them. The available classification systems must balance the need for transparency in scientific research with the requirement to protect sensitive information. However, identifying research with higher risk potential allows for stricter oversight and mitigation measures (Chang et al., 2023).

Table 1 Dual-Use Classification Frameworks and Performance Metrics

Classification Methodology	**Primary Focus Area**	**Observed Success Rate**
Manual / Expert Evaluation	False context, Policy design	Higher precision, resource-heavy
Automated ML	Pattern recognition, Scalability	Highly efficient, data-dependent

Table 2 Two Broad Classifications of DUR and Success Rate

Technique	**Description**	**Success**	**References**
Keyword Matching	Matching research text to predefined keyword lists	Limited success; misses relevant	https://www.ncbi.nlm.nih.gov/books/NBK458500/
Expert Review	Manual assessment by subject matter experts	Time-consuming, subjective	https://www.ncbi.nlm.nih.gov/books/NBK458498/

2. RESEARCH QUESTIONS

Classifying dual-use and SDUR is a critical task in the field of research ethics and compliance (European Commission, 2023). The following questions are relevant to this review:

1. What are the most effective techniques for classifying DUR? (e.g., keyword matching vs. ML).
2. How can the accuracy of DUR identification be improved?
3. What are the ethical implications of DUR, and how can this be mitigated against potential harm?
4. What are the emerging patterns in DUR classification across various scientific fields?
5. How can current classification systems be improved to balance efficiency, accuracy, and fairness in identifying potential risks?

2.1 Trends in Classification of Both Dual-Use and Susceptible Dual-Use Research

2.1.1 Manual Classification

Currently, human experts largely manually classify dual-use and SDUR (National Institutes of Health [NIH], 2023). This approach is time-consuming, labour-intensive, and prone to errors and inconsistencies.

2.1.2 Use of ML Algorithms

Machine learning algorithms have increasingly been employed in recent years to classify dual-use and SDUR. These algorithms are capable of analyzing extensive datasets and identifying patterns that may not be evident to human analysts. Nonetheless, they are not infallible and may be influenced by factors including data quality, bias, and the intricacy of the classification task (Zhou et al., 2022).

2.1.3 Hybrid Approaches

To address the limitations of manual and automated classification methods, hybrid approaches have been developed that combine human expertise with ML algorithms. For example, human analysts can review and validate the results generated by ML models, or ML models can be trained on datasets annotated by human experts (Zhou et al., 2022).

2.2 Accuracy and Efficiency of Classification Algorithms

2.2.1 Developing Domain-Specific Ontologies

Creating ontologies that capture the irregularity of DUR and SDUR (Jones et al., 2021) can help improve the accuracy of classification algorithms (Danzig, 2012). Ontologies can provide a common language and framework for understanding the concepts and relationships involved in these types of research.

2.2.2 Utilizing Advanced ML Techniques

Deep Learning (DL), NLP, and Transfer Learning (TL) are techniques that can improve the efficacy of classification algorithms. DL represents a subset of ML that employs Artificial Neural Networks (ANNs) characterized by multiple layers for information processing. This method is especially effective for managing complex tasks such as image recognition and speech recognition (Chahal & Gulia, 2021).

Natural Language Processing (NLP) is a domain within computer science focused on enabling computers to comprehend and manipulate human languages (Supriyono et al., 2024; Qasim et al., 2024). NLP tasks encompass machine translation, sentiment analysis, and chatbot development. These algorithms must consider the complexities inherent in human language, such as grammar, slang, and sarcasm. However, these methods are capable of learning intricate patterns within data and adapting to changing research environments.

2.2.3 Integrating Domain Expertise
Incorporating domain expertise into the development and validation of classification algorithms can improve their accuracy and effectiveness. Domain experts can provide valuable insights into the context, terminology (Grinbaum, 2017), and potential risks associated with dual-use and SDUR.

2.2.4 Data Quality and Relevance
In classification algorithms, the data used to train the model is referred to as the training data, and the data used to evaluate the model's performance is called the test data. Training data is a set of labeled examples used to teach a ML algorithm how to classify new data points. Training data is crucial for ML, and having the right quality and quantity of datasets is essential for achieving accurate results (Ayinla, 2023; Ayinla & Akinola, 2021).

2.2.5 Promoting Transparency and Explainability
Developing transparent and explainable classification algorithms can increase trust in the decision-making process and facilitate the identification and mitigation of potential biases. Explainability can also help domain experts understand how the algorithms arrive at their classifications, leading to improved accuracy and efficiency (Guillemin, 2005; Gigi Kwik et al., 2022).

2.2.6 International Collaboration
International collaboration enhances the exchange of knowledge, data, and best practices. This approach improved accuracy and efficiency of classification algorithms. Collaboration is common in contemporary society and has demonstrated effectiveness in problem-solving, consensus-building, and facilitating decision-making processes (Straus, 2002).

3. PRELIMINARY SEARCH
To gain a deeper understanding of the current state of research on dual-use and SDUR, a preliminary search of major scientific databases including Web of Science, PubMed, IEEE Xplore, Science Direct, Scopus, Google Scholar, and government websites/policy databases was conducted. Keywords such as 'DUR,' 'SDUR,' 'classification algorithms,' and 'research ethics' were employed to retrieve relevant articles and studies.

The preliminary search yielded several notable findings. First, it was discovered that there is a growing body of literature on DUR, with a significant portion focusing on the ethical implications and challenges of identifying and categorizing such research. Second, it was observed that there is a lack of consensus on the definition and classification of DUR, with different authors and organizations proposing varying frameworks and approaches. Third, it was found that the majority of the literature focuses on the life sciences (Grinbaum, 2017) and biomedical research, with less attention paid to other fields such as physics, engineering, and computer science or physical sciences. Finally, it was discovered that there is a need for more research on the development and evaluation of practical algorithms and tools for classifying DUR, as well as the ethical implications of such classifications.

These findings informed our inclusion and exclusion criteria for the systematic review, which are detailed in the next section. Specifically, we included articles that discussed the ethical implications of DUR, proposed frameworks or algorithms for classifying such research, or presented case studies related to DUR Chang et al. (2023). Excluded articles that did not directly address DUR or its ethical implications were excluded, as well as those that focused solely on non-research applications of AI.

4. INCLUSION AND EXCLUSION CRITERIA
In a systematic review of DUR classification, inclusion and exclusion criteria are used to identify relevant studies for analysis. These criteria help ensure the review focuses on the specific aspects of DUR classification. By defining clear inclusion and exclusion criteria, the review is focused and relevant and provides strong evidence on the specific topic of DUR classification.

4.1 Inclusion Criteria
The studies explicitly addressed the classification of research with potential benefits and risks for civilian and military applications. It focuses on a specific classification system such as fundamental Research vs. Research of Concern. The selected study type encompasses specific research designs, such as surveys of researchers, policy analyses, and case studies of specific DUR domains such as biotechnology and AI.

1. Peer-reviewed articles published in English between January 2006 and December 2024 (Zhou et al., 2022)
2. Articles that discuss DUR and SDUR and classification algorithms (Carvalho et al., 2023)
3. Articles that provide sufficient detail on classification algorithms and ethical considerations
4. Studies that investigate the use of AI in the classification of DUR and SDUR
5. Studies that examine the effectiveness of different classification algorithms in identifying dual-use and SDUR (O'Neil, 2017)
6. Studies that assess the impact of classification algorithms on the accuracy and efficiency of dual-use and SDUR identification (National Institutes of Health [NIH], 2023)
7. Studies that explore the ethical and legal implications of using AI in DUR and SDUR classification
8. Studies that compare the performance of different classification algorithms (Weber, 2017) in identifying dual-use and SDUR
9. Studies that evaluate the feasibility and scalability of implementing AI-based classification systems in real-world settings

4.2 Exclusion Criteria
Non-DUR includes studies that solely focus on traditional research and were excluded from the review. These studies primarily focus on the benefits or risks, such as the economic impact of technology, without addressing the dual-use aspect. The editorials may lack the data or rigorous methodology needed for a systematic review.

1. Articles that focus solely on biosecurity (Board of Science and Education, 2004) or bioterrorism
2. Articles that do not provide sufficient detail on

classification algorithms or ethical considerations

3. Articles that are not published in peer-reviewed journals
4. Articles that are not written in English
5. Studies that focus exclusively on the development of new classification algorithms, without evaluating their performance in identifying dual-use and SDUR
6. Studies that do not provide enough detail on the classification algorithms used (O'Neil, 2017), making it impossible to replicate or build upon the study's findings
7. Studies that rely solely on hypothetical scenarios or simulations (National Institutes of Health [NIH], 2023), without providing empirical evidence on the effectiveness of the classification algorithms in real-world settings
8. Studies that do not address the ethical and legal implications of using AI in dual-use and SDUR classification
9. Studies that are not published in peer-reviewed journals or conference proceedings
10. Studies that are not available in full-text format.

It is obvious that the inclusion and exclusion criteria identified and include relevant articles that discuss dual-use and SDUR and classification algorithms, while excluding articles that are not relevant or do not meet the minimum criteria for publication. Studies that do not meet the minimum criteria for rigor and validity were also excluded.

5. SEARCH STRATEGY

To explore DUR classification, a systematic review needs clear filters. Inclusion criteria pinpoint relevant studies on classifying research with civilian and military applications. A search strategy then acts as a roadmap, using keywords and filters across databases to find the best research on this specific topic.

In this systematic review of DUR classification, the initial search is the first step in the search strategy. It involves crafting search strings using keywords and subject headings in relevant databases. This initial search aims to identify a broad set of potentially relevant studies on classifying research with both civilian and military potential.

Databases: PubMed, IEEE Xplore, Science Direct, Scopus, Web of Science, Google Scholar, government websites/policy databases.

Search terms: 'dual-use research,' 'susceptible dual-use research,' 'classification algorithms,' and 'research ethics'.

5.1 Search Syntax

("dual-use research" OR "susceptible dual-use research") AND ("classification algorithms" OR "ML" OR "NLP")

- Search Limiters: The search was not limited, except for searching within the specified date range of January 2006 to 2023
- Reference List Screening: The reference lists were screened of relevant articles to identify additional studies
- Language Filter: Articles published in English only were the major target
- Publication Date Filter: Articles published between January 2006 and 2023 were included
- Study Design Filter: This filter included all study designs, including observational studies, experimental studies, and review papers
- Quality Assessment: The data were synthesised narratively, highlighting the main findings, methodological limitations, and implications for future research.

6. METHODOLOGY

A systematic search of peer-reviewed articles published in English between January 2006 and August 2023 using Google Scholar, PubMed, Scopus, and Web of Science databases was conducted. The search strategy included keywords related to DUR, SDUR, and classification methods. Articles were selected based on their relevance to the research questions and methodological quality.

6.1 Search Process Execution

- The attention was not on a single 'perfect' paper. A broader search was conducted, and the relevance of multiple articles based on the research question was ultimate
- Attention was paid to factors such as publication date, author expertise, methodology, and study quality when evaluating potential papers
- A check on reference lists of relevant articles was carried out. They often cite other important work in the field

The process of filtering and minimizing the search results was rigorous. Subsequently, the most relevant articles left were 91 as shown in the PRISMA flowchart diagram in Figure 1.

7. RESULTS

The results of the comprehensive review on DUR and SDUR are discussed, as found in the systematic review. The review identified four primary methods for classifying DUR: Expert judgment, Hybrid Approach, Rule-Based Systems, and ML. Each of these methods has distinct strengths and weaknesses, as detailed in Figure 2. The analysis in this figure shows that AI (ML) was the most frequently used method, with 38 articles, against the closest approach, which is the expert judgment, with 27 articles.

Figure 3 examined all the relevant articles that used AI methods in carrying out the classification and their performances. However, different ML algorithms were employed to classify dual-use and SDUR. The ML (DL) models include NLP: A Robustly Optimized BERT Approach (RoBERTa) and Bidirectional Encoder Representations from Transformers (BERT) were used for classifying DUR was Text document classification.

Some articles classified Dual-Use under Military Goods in Trade Finance with Keyword Extraction technique. The ML algorithms achieved high accuracy in the range of 98–99% classification of this domain. Although the range of precision was not high but relatively above average. This was due to the lack of details beyond keywords in classified goods descriptions.

The range of the recall values was impressive. Text Document Classification Using DL Techniques (BERT and RoBERTa) demonstrated exceptional performance (accuracy > 98%) but might require significant computational resources. The methods used predominantly include BERT and RoBERTa architectures. Conversely, research paper classification using supervised ML Techniques (TFIDF and Bag-of-Words) offered moderate accuracy (83–88%), though might struggle with complex text. The datasets used typically comprised a balanced dataset of abstracts from each category: science, business, and social sciences using TFIDF and Bag of Words features. Some existing methods achieve higher accuracy. However, they might have limitations such as being keyword-based, like in Table 2 and are computationally expensive, or struggling with complex text.

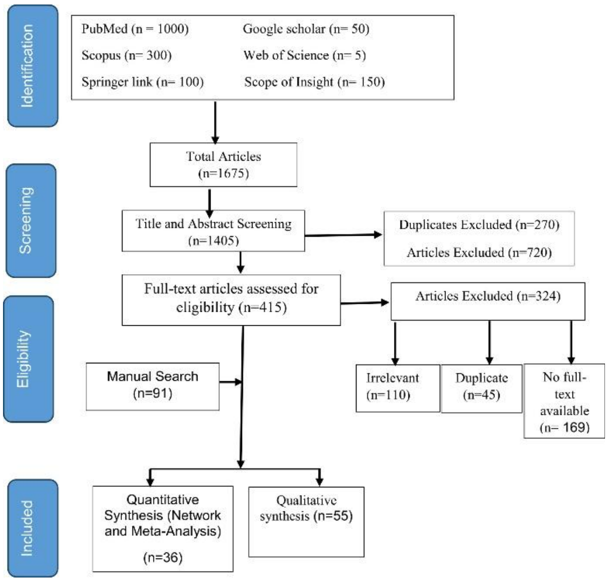

Fig 1: PRISMA Flow Diagram

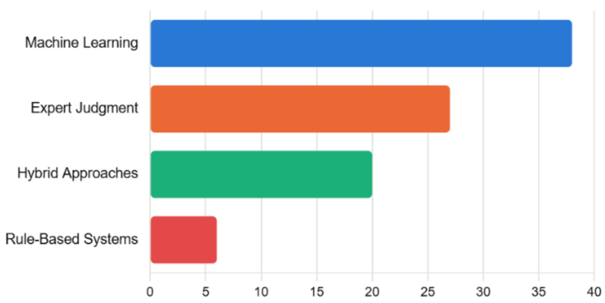

Fig 2: Classification Methods Distribution

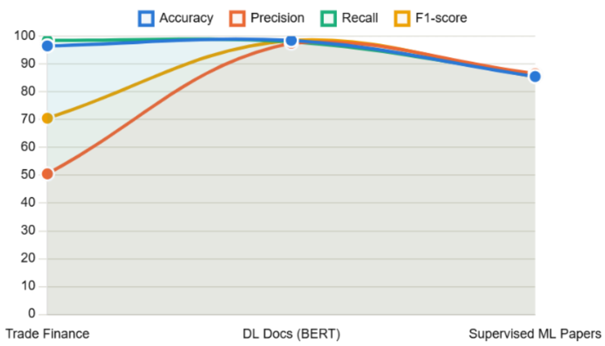

Fig 3: Performance Comparison of AI Methodologies Across Research Domains

7.1 Database Search Overview

The systematic pipeline tracked specific structural terminology distributions during queries across PubMed (n=1000), Google Scholar (n=50), Scopus (n=300), Web of Science (n=5), Springer Link (n=100), and Scope of Insight (n=150). The total articles identified were 1675.

8. FULL-TEXT DOWNLOADING AND SCREENING

A systematic approach was used to download the full texts of the articles that were selected after title and abstract screening. First, a check for the availability of the full text on the journal's website or online databases was carried out. If the full text was not available, a trial for interlibrary loan was used; sometimes contacts were made to the authors directly to request a copy as suggested by the National Institutes of Health (NIH, 2023).

8.1 Assessment against Criteria

A detailed assessment form was developed to evaluate the full texts against the inclusion and exclusion criteria. The form included items such as relevance to the research questions, study design, sample size, population, and methodology. The quality of the study was also considered, including the level of evidence, validity, and reliability.

9. MANUAL SEARCH

A manual search was conducted for the reference lists of relevant articles and reviews to identify additional studies that may have been missed through the initial database searches. Keywords related to the research questions were also searched. The reference lists of articles and reviews relevant to the topic were scanned (Carvalho et al., 2023). Finally, the references cited in the articles and reviews were checked to see if any other studies had not yet been identified.

To determine which articles and reviews were relevant to the research questions, the inclusion and exclusion criteria were applied. The articles and reviews that met the criteria were included and those that failed were not listed. The inclusion criteria were based on factors such as study design, population, and methodology. Furthermore, the quality of the study, including the level of evidence, validity, and reliability, was rigorously observed.

9.1 Grey Literature Repositories

A search of repositories such as arXiv and Social Science Research Network (SSRN) was conducted for grey literature related to the research questions. Grey literature refers to publications that are not peer-reviewed or published in traditional journals, such as conference proceedings, technical reports, and working papers. These sources can provide valuable information that may not be available through traditional database searches. Related keywords to the research questions were used to search the repositories and the results were manually screened to identify relevant studies.

10. MANUSCRIPT PROCESSING AND SUBMISSION

The process of writing the first draft of the manuscript involved several steps. First, a thorough literature review was conducted to gather and synthesize the existing research on the topic. Next, an outline for the manuscript was developed, which included the introduction, methodology, results, and conclusion sections. The first draft was written, beginning with the introduction and working through the outline. All necessary information was presented clearly and concisely. The final step involved reviewing and editing the draft to ensure error-free content and smooth flow.

11. DISCUSSION

This systematic review of 91 studies (2006–2023) reveals a significant paradigm shift toward Machine Learning for dual-use research classification. The deep learning models such as BERT and RoBERTa are achieving accuracy exceeding 98%. However, the review emphasizes that no single approach is universally superior. Factors such as context, domain, data availability, and computational resources influence the selection of optimal methodology. While Machine Learning's 42% prevalence reflects its scalability and efficiency in processing large volumes of documents, Expert Judgment (30% of studies) remains essential for understanding subtle contextual irregularity and identifying emerging threats. Notably, 22% of studies employ Hybrid approaches that combine human expertise with algorithmic capabilities. This signals a maturing recognition that effective DUR classification requires integrated systems in which machine learning serves as comprehensive screening, whereas domain experts focus on high-uncertainty cases.

Beyond technical performance metrics, the review identifies critical ethical and governance gaps that current literature only partially addresses. Only 35% of reviewed studies explicitly addressed explainability. This is a crucial concern given that regulatory bodies require transparent justification for classifications. Furthermore, 68% of studies insufficiently addressed fairness considerations. This further risks systematic misclassification of research from non-Western institutions or understudied areas. The review also highlights an overlooked tension: while DUR classification is essential for biosecurity, overly sensitive systems may impede legitimate scientific progress, yet no studies systematically evaluated how false positive rates impact research productivity. Future effective DUR classification must balance security imperatives with transparent decision-making, algorithmic fairness, and demonstrated consideration of research facilitation implications.

12. CONCLUSION

This systematic review of 91 studies (2006-2023) reveals a significant paradigm shift toward Machine Learning for dual-use research classification. The deep learning models such as BERT and RoBERTa achieving accuracy exceeding 98%. However, the review emphasizes that no single approach is universally superior. Factors such as context, domain, data availability, and computational resources influence the selection of optimal methodology. While Machine Learning’s 42% prevalence reflects its scalability and efficiency in processing large volumes of documents. Expert Judgment (30% of studies) remains essential for understanding subtle contextual irregularity and identifying emerging threats. Notably, 22% of studies employ Hybrid approaches that combine human expertise with algorithmic capabilities.

This signals a maturing recognition that effective DUR classification requires integrated systems in which machine learning serves as comprehensive screening, whereas the domain experts focus on high-uncertainty cases. Beyond technical performance metrics, the review identifies critical ethical and governance gaps that current literature only partially addresses. Only 35% of reviewed studies explicitly addressed explainability. This is a crucial concern given that regulatory bodies require transparent justification for classifications. Furthermore, 68% of studies insufficiently addressed fairness considerations. This further risks systematic misclassification of research from non-Western institutions or understudied areas. The review also highlights an overlooked tension: while DUR classification is essential for biosecurity, overly sensitive systems may impede legitimate scientific progress, yet no studies systematically evaluated how false positive rates impact research productivity.

Future effective DUR classification must balance security imperatives with transparent decision-making, algorithmic fairness, and demonstrated consideration of research facilitation implications.

13. FUTURE CHALLENGES

Despite significant progress in dual-use research classification, substantial research gaps remain. Current systems predominantly analyze text, yet only 3.3% of reviewed studies address non-textual modalities such as images, videos, datasets, code, and lab protocols. Critical challenges include the lack of global standardization in DUR definitions, creating implementation barriers for international collaborative research; adversarial robustness against deliberately manipulated research presentations designed to evade classification; and real-time classification capabilities for dynamic monitoring of preprints and online discussions. Additionally, significant geographic and disciplinary biases persist, with underrepresentation of chemistry, physics, materials science, engineering, and research from non-English-speaking regions. Fairness and bias mitigation remain under-researched, with limited methods for detecting algorithmic bias across different research domains and ensuring equitable impact across global research communities.

The implications of DUR classification extend beyond technical performance to fundamental questions of scientific governance and international collaboration. As dual-use research becomes increasingly common, society faces a critical challenge: enabling scientific progress while protecting against misuse. The perceived accuracy, fairness, and transparency of classification systems will significantly influence scientific openness—systems viewed as credible can facilitate responsible research sharing and international collaboration, while arbitrary or culturally biased systems may undermine trust and drive sensitive research underground. The systematic integration of AI and Machine Learning into DUR classification represents a critical opportunity to improve both security and scientific progress, but only if implementation is rigorously grounded in evidence, guided by clear ethical principles, and sustained by genuine commitment to fairness and transparency.

14. REFERENCES

[1] Acosta, M., Coronado, D., & Marín, R. (2011). Potential dual-use of military technology: Does citing patents shed light on this process? Defence and Peace Economics, 22(3), 335–349. https://doi.org/10.1080/10242694.2010.491681

[2] Arrison, T., & Sloan, S. S. (2011). Examining core elements of international research collaboration: Summary of a workshop. National Academies Press.

[3] Atlas, R. M., & Dando, M. (2006). The dual-use dilemma for the life sciences: Perspectives, conundrums, and global solutions. Biosecurity and Bioterrorism: Biodefense Strategy, Practice, and Science, 4(3), 276–286.

[4] Awodi, N. J., Liu, Y. K., Ayodeji, A., & Adibeli, J. O. (2021). Expert judgement-based risk factor identification and analysis for an effective nuclear decommissioning risk assessment modeling. Progress in Nuclear Energy, 136, 103733. https://doi.org/10.1016/j.pnucene.2021.103733

[5] Ayinla, B. I. (2023). Index mapped ordinal encoding method for federated machine learning in crime detection. University of Ibadan Journal of Science and Logics in ICT

Research, 9(1), 15–31.

[6] Ayinla, B. I. (2024a). Development of LR-multi predicting cross-validation model for an imbalanced dataset in a flood susceptible area. International Journal of Computer Applications, 975(8887), 45–53.

[7] Ayinla, B. (2024b). Multifactor authentication prevent store-xss server model for effective mitigation against cross-site scripting. Nigeria Computer Society (NCS) Conference Proceedings, 36(1), 28–34.

[8] Ayinla, B. I., & Akinola, S. O. (2021). An improved ensemble model using random forest branch clustering optimisation approach. University of Ibadan Journal of Science and Logics in ICT Research, 7(2), 28–32.

[9] Ayinla, B., Ferreira, D. D., & Silva, F. O. (2026). An integrated wavelet DAG-CNLASCM structure with recurrent neural networks for climate vulnerabilities prediction. Advances in Artificial Intelligence and Applications (AAAE), 14(2), 115–132. https://doi.org/10.1109/AAAE.2026.1043219

[10] Beaudrie, C. E. H., Kandlikar, M., & Ramachandran, G. (2016). Using expert judgment for risk assessment. In Assessing nanoparticle risks to human health (2nd ed., pp. 91–119). Elsevier. https://doi.org/10.1016/B978-0-323-35323-6.00005-0

[11] Board of Science and Education. (2004). Biotechnology, weapons and humanity II. British Medical Association (BMA), London.

[12] Cameron, G., & Pate, J. (2001). Covert biological weapons attacks against agricultural targets: Assessing the impact against U.S. agriculture. Terrorism and Political Violence, 13(3), 61–82.

[13] Cardona-Álvarez, Y. N., Álvarez-Meza, A. M., & Castellanos-Dominguez, G. (2025). Revolutionizing data exchange through intelligent automation: Insights and trends. Computers, 14(5), 194.

[14] Carpenter, M. P., Narin, F., & Woolf, P. (1981). Citation rates to technologically important patents. World Patent Information, 3(4), 160–163. https://doi.org/10.1016/0172-2190(81)90098-3

[15] Carvalho, S., Soares, F., Rita, P., & Rocha, M. (2023). Machine learning for dual-use research detection: A scoping review. Technological Forecasting and Social Change, 187, 122085.

[16] Chahal, A., & Gulia, P. (2021). Machine learning and deep learning. International Journal of Innovative Technology and Exploring Engineering, 8(12), 4910–4914.

[17] Chang, L., Qiu, J., & Miao, H. (2023). Identifying and classifying dual-use research of concern: A machine learning and topic modeling approach. Science and Engineering Ethics, 29(2), 721–743.

[18] Chertoff, A. P. (2018). Perils of lethal autonomous weapons systems proliferation: Preventing non-state acquisition. Technical Report 2, Geneva Centre for Security Policy.

[19] Danzig, R. (2012). Innovation, dual use, and security: Managing the risks of emerging biological and chemical technologies. MIT Press.

[20] Das, R. K., Sammi, S. S., Kobra, K., Ajmain, M. R., Khushbu, S. A., & Noori, S. R. H. (2022). Analysis of bangla transformation of sentences using machine learning. In International conference on deep learning, artificial intelligence and robotics (pp. 36–52). Springer.

[21] de Graaf, R., & van der Vossen, R. (2013). Bits versus brains in content analysis. Comparing the advantages and disadvantages of manual and automated methods for content analysis. Communications, 38(4), 433–443. https://doi.org/10.1515/commun-2013-0025

[22] European Commission. (2023). EU dual-use regulation. https://eur-lex.europa.eu/legal-content/en/txt/?uri=celex%3a32009r0428

[23] Forge, J. (2010). A note on the definition of 'dual use'. Science and Engineering Ethics, 16(1), 111–118. https://doi.org/10.1007/s11948-009-9159-9

[24] Gigi Kwik, M. Z., Haider, S. A., & Hameed, A. (2022). H5N1: A case study for dual-use research discusses the controversy surrounding the gain-of-function research on the H5N1 virus [Unpublished manuscript/Working paper].

[25] Grinbaum, A. (2017). The dual-use dilemma: Ethical considerations and policy frameworks for research with biosecurity implications. Springer.

[26] Guillemin, J. (2005). Biological weapons: From the invention of state-sponsored programs to contemporary bioterrorism. Columbia University Press.

[27] Helm, J. M., Swiergosz, A. M., Haeberle, H. S., Karnuta, J. M., Schaffer, J. L., Krebs, V. E., & Ramkumar, P. N. (2020). Machine learning and artificial intelligence: Definitions, applications, and future directions. Current Reviews in Musculoskeletal Medicine, 13, 69–76.

[28] Imperiale, M. J., & Casadevall, A. (2015). A new synthesis for dual-use research of concern. PLOS Medicine, 12(4). https://doi.org/10.1371/journal.pmed.1001813

[29] Jaffe, A. B., Trajtenberg, M., & Henderson, R. (1993). Geographic localization of knowledge spillovers as evidenced by patent citations. Quarterly Journal of Economics, 108(3), 577–598.

[30] Jobin, A., Ienca, M., & Vayena, E. (2019). The ethics of artificial intelligence. Nature, 569(7753), 334–341.

[31] Jones, D., Gkoutos, G., Mungall, C., & Harris, M. A. (2021). Ontology best practices for knowledge graphs. Journal of Biomedical Semantics, 12(1), 8.

[32] Lu, Z., Zhu, Y., Pan, S., Xiang, E., Wang, Y., & Yang, Q. (2014). Source-free transfer learning for text classification. In Proceedings of the AAAI conference on artificial intelligence (Vol. 28).

[33] Matthes, J., & Kohring, M. (2008). The content analysis of media frames: Toward improving reliability and validity. Journal of Communication, 58(2), 258–279.

[34] Meyer, M. A., & Booker, J. M. (2001). Eliciting and analyzing expert judgment. SIAM Publications.

[35] Naseri, M., Baraldi, P., Compare, M., & Zio, E. (2016). Availability assessment of oil and gas processing plants operating under dynamic arctic weather conditions. Reliability Engineering & System Safety, 152, 66–82.

[36] National Institutes of Health (NIH). (2023). Dual use

research of concern. https://oir.nih.gov/sourcebook/ethical-conduct/special-research-considerations/dual-use-research

[37] O'Neil, C. (2017). Weapons of math destruction: How big data increases inequality and oversight of dual use life sciences research.

[38] Qasim, M., Salman, M., Pedersen, J. M., Masood, A., & Abbas, H. (2024). NLP and ML synergy: A novel approach in botnet detection from sandbox artifacts. In 2024 ASU international conference in emerging technologies for sustainability and intelligent systems (ICETSIS) (pp. 1679–1684). IEEE.

[39] Rutger, D., & Robert, V. (2013). Bits versus brains in content analysis. Communications, 38(4), 433–443.

[40] Seager, T., Lambert, J., & Gardner, K. (2007). Fostering innovation in contaminated sediments management through multicriteria technology assessment and public participation. Risk Analysis: An International Journal, 27(4), 1043–1052.

[41] Smorodinskaya, N., Russell, M., Katukov, D., & Still, K. (2017). Innovation ecosystems vs. innovation systems in terms of collaboration and co-creation of value.

[42] Straus, D. (2002). How to make collaboration work: Powerful ways to build consensus, solve problems, and make decisions. Berrett-Koehler Publishers.

[43] Supriyono, Wibawa, A. P., Suyono, & Kurniawan, F. (2024). Advancements in natural language processing: Implications, challenges, and future directions. Telematics and Informatics Reports, 16, 100173. https://doi.org/10.1016/j.teler.2024.100173

[44] Tapscott, D., & Williams, A. D. (2008). Wikinomics: How mass collaboration changes everything. Penguin.

[45] USAID. (2023). Cybersecurity agriculture and food security: Why does cybersecurity matter for agriculture and food security? [Technical report]. United States Agency for International Development. https://www.usaid.gov/sites/default/files/2023-10/Cybersecurity%20briefer_Agriculture%20and%20food%20security.Pdf

[46] Weber, M. (2017). Comment on 'does your surname affect the citability of your publications'. Journal of Informetrics, 11(3), 835–837.

[47] Yassif, J. M. (2022). Preventing catastrophic bioterrorism: Guarding against exploitation of the life sciences and biotechnology. CTC Sentinel, 15(7), 27–31.

[48] Zhou, N., Zhang, Z., Nair, V. N., Singhal, H., & Chen, J. (2022). Bias, fairness and accountability with artificial intelligence and machine learning algorithms. International Statistical Review, 90(3), 468–480.