REVISTA INGI – INDICAÇÃO GEOGRÁFICA E INOVAÇÃO
Vol. 10, n. 4, p. 3772-3788. Out/Nov/Dez (2026)
ISSN: 2594-8288
DOI: 10.51722/Ingi.v10.i4.12

OPEN ACCESS
www.api.org.br

AUTOMATIC CLASSIFICATION OF IFS INNOVATION PROJECTS ACCORDING TO THE 2030 AGENDA USING LARGE LANGUAGE MODELS

CLASSIFICAÇÃO AUTOMÁTICA DE PROJETOS DE INOVAÇÃO DO IFS SEGUNDO A AGENDA 2030 USANDO MODELOS DE LINGUAGEM DE GRANDE ESCALA

Allívia Rouse Carregosa Rabbani[1], Marta Jeidjane Borges Ribeiro[2], Mateus Costa dos Santos[3], Chirlaine Cristine Gonçalves[4], José Augusto Andrade Filho[5]

[1] Federal Institute of Bahia. Porto Seguro/BA. Brazil. Directorate of Innovation and Entrepreneurship – DInovE. Federal Institute of Sergipe. Aracaju/SE. Brazil. alliviarouse@hotmail.com. https://orcid.org/0000-0003-0564-7113
[2] Directorate of Innovation and Entrepreneurship – DInovE. Federal Institute of Sergipe. Aracaju/SE. Brazil. emestatistica@gmail.com. https://orcid.org/0000-0002-8459-1981
[3] Graduate Program in Environmental Sciences and Technologies. Federal Institute of Bahia. Porto Seguro/BA. Brazil. ti.mateuscosta@gmail.com. https://orcid.org/0009-0006-9340-9607
[4] Pro-Rectory of Institutional Development – Prodin. Federal Institute of Sergipe. Aracaju/SE. Brazil. chirlaine.cristine@ifs.edu.br. https://orcid.org/0000-0003-1321-3366
[5] Directorate of Innovation and Entrepreneurship – DInovE. Federal Institute of Sergipe. Aracaju/SE. Brazil. jose.andrade@ifs.edu.br. https://orcid.org/0000-0001-7496-7386

Abstract

The Federal Institutes of Education, Science and Technology (IFs) play a strategic role in promoting innovation and territorial development. The 2030 Agenda recognizes science, technology, and innovation as central instruments for achieving its goals. This article analyzes the titles of innovation projects approved by the Federal Institute of Sergipe (IFS) between 2014 and 2023, classifying them according to the Sustainable Development Goals (SDGs) through a large language model (LLM). The research is exploratory-descriptive in nature and adopts a mixed-methods approach, combining automated text processing with quantitative analysis. Results indicate a strong thematic concentration in SDG 9 (56%), followed by SDG 4 (9%), SDG 8 (7%), and SDG 7 (7%). Together, these four SDGs account for approximately 80% of the initiatives analyzed, reflecting the institutional mission of the IFs to integrate vocational training, applied research, and territorial innovation. The LLM-based classification predominantly yielded medium and low confidence levels, indicating inherent limitations in the exclusive use of project titles as analytical input. The findings suggest that LLMs represent promising tools for the institutional monitoring of innovation policies, although their application requires human validation and an expanded textual corpus for greater interpretive robustness.

Keywords: Monitoring; Federal Network; LLMs; SDGs.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Resumo

Os Institutos Federais de Educação, Ciência e Tecnologia (IFs) desempenham um papel estratégico na promoção da inovação e do desenvolvimento territorial. A Agenda 2030 reconhece a ciência, a tecnologia e a inovação como instrumentos centrais para o alcance de suas metas. Este artigo analisa os títulos dos projetos de inovação aprovados pelo IF de Sergipe entre 2014 e 2023, classificando-os segundo os Objetivos de Desenvolvimento Sustentável (ODS) por meio de um modelo de linguagem de grande escala (LLM). A pesquisa tem caráter exploratório-descritivo e adota uma abordagem mista, combinando o processamento automatizado de texto com a análise quantitativa. Os resultados indicam forte concentração temática dos projetos no ODS 9 (56%), seguida pelos ODS 4 (9%), 8 (7%) e 7 (7%). Em conjunto, esses ODS concentram aproximadamente 80% das iniciativas analisadas, refletindo a missão institucional dos IFs de articular formação profissional, pesquisa aplicada e inovação territorial. A classificação utilizando LLM apresentou predominância de níveis de confiança médios e baixos, indicando limitações inerentes ao uso exclusivo dos títulos como insumo analítico. Os resultados sugerem que os LLMs se configuram como ferramentas promissoras para o monitoramento institucional de políticas de inovação, embora sua aplicação requeira validação humana e ampliação do corpus textual para maior robustez interpretativa.

Palavras-chave: Monitoramento; Rede federal; LLMs; ODS

1 Introduction

The Federal Institutes of Education, Science and Technology (IFs) represent one of Brazil's most significant public policies in the realm of professional, scientific, and technological education. They pursue, in an inseparable manner, academic training, technological development, and social inclusion from a territorialized perspective, distinguishing themselves from traditional universities by emphasizing interaction with the productive sector and the communities in which they operate (Brasil, 2008; Pacheco, 2010; Rabbani *et al.*, 2025).

In this context, innovation is intrinsically linked to the capacity to generate new productive combinations capable of transforming economic structures and producing social value, while also driving contemporary policies. A country's innovative capacity is directly associated with its level of economic development and the quality of its educational and scientific systems (OCDE, 2018).

The territorial dimension of innovation is shaped by interactions among different social and institutional actors, including educational institutions that constitute local and regional innovation systems. Brazil provides incentives for cooperation among scientific institutions, enterprises, and government, stimulating technology transfer and applied knowledge production (Brasil, 2004). In this regard, the IFs' capillarity across the national territory places them in a strategic position to foster innovation processes in regions with lower scientific and technological density (Cassiolato; Lastres, 2005).

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Parallel to these institutional transformations, the international sustainable development agenda has assumed a central position in global public policy. In 2015, the member states of the United Nations adopted the 2030 Agenda for Sustainable Development, a global pact comprising 17 Sustainable Development Goals (SDGs) and 169 targets aimed at promoting fairer, more inclusive, and environmentally sustainable societies, explicitly recognizing science, technology, and innovation as fundamental instruments for addressing global challenges (ONU, 2015).

In Brazil, the Agenda was incorporated in 2016 through Decree No. 8,892/2016, which expanded the SDG list by incorporating SDG 18 (Ethno-Racial Equality). However, persistent challenges remain regarding the monitoring of institutional contributions to achieving the targets, particularly with respect to scientific and technological output (UNESP, 2025; Brasil, 2026; ISS, 2026).

Mapping scientific and technological output in light of the SDGs has thus become a growing agenda in academic literature (Sandrino; Jamuchewski, 2022; Fuchs *et al.*, 2023; Lemarchand *et al.*, 2023). Approaches based on machine learning and semantic analysis have enhanced the precision of such classifications, enabling the processing of large volumes of institutional records with greater scale and reproducibility (Angin *et al.*, 2022).

In this field, the advancement of Large Language Models (LLMs) has represented a qualitative leap in available methodological possibilities. These models are capable of understanding complex semantic relationships in technical and scientific texts, enabling thematic classification tasks with high levels of reliability. Studies indicate that automated processing can classify documents with precision levels comparable to those achieved by specialized human evaluators, making them promising tools for the institutional monitoring of Science, Technology and Innovation (S&T&I) policies at scale (Suadaa *et al.*, 2025; Morales-Hernández *et al.*, 2022; Duda *et al.*, 2025).

Despite these advances, gaps remain in the literature regarding the analysis of specific educational and research institutions' contributions to the 2030 Agenda. Studies have concentrated primarily on global bibliometric databases, privileging research universities and internationally circulated journals (Bautista-Puig *et al.*, 2021). In the Brazilian context, empirical analyses investigating the alignment of Federal Institutes' scientific and technological output with the SDGs, especially through LLM-based approaches, remain scarce (Duda *et al.*, 2025; Rabbani *et al.*, 2025).

Against this backdrop, the Federal Institute of Education, Science and Technology of Sergipe (IFS) constitutes a privileged context for investigating the relationship between institutional innovation and sustainable development, given its publicly available data repository (https://publicacoes.ifs.edu.br/). The IFS currently operates ten campuses distributed across the state of Sergipe, in addition to its Rectory headquartered in Aracaju (IFS, 2026). This territorial capillarity

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

places the institute in direct contact with the socioeconomic demands of a state marked by structural development challenges, with significant poverty and inequality indicators in several regions, particularly in the interior (IBGE, 2023).

The institute coordinates institutional projects through internal calls and public funding agencies (IFS, 2026), although without explicit orientation toward the SDGs. Identifying thematic patterns in output contributes to understanding which sustainability fields the IFS supports throughout its trajectory, while also subsidizing institutional policy monitoring. In light of this scenario, the research is guided by the following question: to what extent are the innovation projects approved by the IFS between 2014 and 2023 aligned with the SDGs, and to what degree are LLMs capable of automatically and reliably identifying this alignment based on project titles?

Based on this question, the study departs from the hypothesis that the IFS's innovative output is predominantly concentrated in SDGs related to technology and vocational training, directly reflecting the institutional mission of the Federal Institutes and the courses they offer, while LLMs demonstrate operational capacity for large-scale thematic classification, though with precision limitations when the available textual input is restricted to project titles.

Thus, the present study aims to analyze the innovation projects approved at the IFS between 2014 and 2023, using an LLM to classify them according to the 18 SDGs adopted by Brazil, assessing both the thematic profile of institutional output and the model's performance as a monitoring tool for science, technology, and innovation policies.

2 Material and Methods

This research is exploratory-descriptive in nature and adopts a quantitative approach supported by LLM-based automated processing. The LLM employed was Claude Sonnet 4.6, developed by Anthropic and accessed through the interface https://claude.ai (Anthropic, 2024). This is a general-purpose language model with semantic comprehension capabilities for technical and scientific texts in Portuguese, selected for its availability, performance on thematic classification tasks, and the ability to receive structured prompt instructions with reference libraries.

It is acknowledged that models of this nature may exhibit biases associated with pre-training data, prompt formulation, and stochastic variability in generated responses, that is, successive executions of the same prompt may produce slightly different classifications, compromising strict reproducibility. To mitigate this risk, all classifications were performed in a single, continuous session with a fixed, structured prompt applied uniformly across all projects, as detailed in section d.1.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Nevertheless, future studies are advised to adopt APIs with fixed generation parameters to ensure full reproducibility.

The methodological process was organized into four sequential stages: (a) data collection and organization, (b) processing and deduplication, (c) filtering by funding line, and (d) SDG alignment, subdivided into thematic classification (d.1) and bibliometric keyword analysis (d.2).

a. Data Collection and Organization: The dataset was obtained through consultation of the public repository of approved projects in the IFS Publication System (SISPUBLI; https://publicacoes.ifs.edu.br/), consolidated into an .xlsx spreadsheet containing 3,356 records. The following were considered innovation funding programs: the Institutional Program for Undergraduate Scholarships in Technology and Innovation (PIBITI); the Innovation in Teaching Scholarship Program (PROBEN); IFSTEC; COVID-19 response initiatives; the IFS Sustainable Program; Lab IFMaker; IF More Entrepreneur (in Portuguese, IF Mais Empreendedor); Voluntary Innovation Projects; and programs such as the Institutional Business Incubation and Creative Economy Program (PIIEEC), the Innovation and Development Program (PID-IFS), the Academic Competition Teams Support Program (CAD-IFS), the Applied Innovation Program (PIAI), and IFS award-winning projects. The temporal analysis was supplemented by a cross-tabulated chart (year × funding program) recording the number of projects for each combination.

b. Data Processing and Deduplication: Given the volume and administrative nature of the dataset, a systematic process for identifying and removing duplicates was carried out using Python and the pandas library. The deduplication identified two analytical categories: (i) exact duplicates, defined as records with identical titles or differences restricted to uppercase/lowercase, accents, spacing, or punctuation; and (ii) semantic duplicates, characterized by titles describing the same project with minor textual variations, such as acronym expansion, term reordering, or abbreviations. For each group of duplicates, the most complete and informative record was retained, prioritizing the presence of an institutional code and the greatest number of characters in the title. This process reduced the dataset from 3,356 to 2,778 unique projects (duplication rate of 17.01%), eliminating 553 exact duplicates and 18 semantic duplicates.

c. Filtering by Funding Line: From the consolidated list, projects linked to funding programs with an explicit focus on innovation and technology were selected. Fourteen search terms were defined, corresponding to the main institutional funding lines: PROBEN, PIBITI, COVID, Innovation, Pre-incubation, Lab IFMaker, IF More Entrepreneur, PIIEEC, Sustainable, IFSTEC, Award-winning, PID-IFS, CAD-IFS, and PIAI. Matching was performed using regular expressions, case-insensitive, applied to the "Program" field. This procedure yielded a corpus of 417 projects distributed across 13 distinct programs over a ten-year period.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

d. SDG Alignment: This stage was divided into two phases: (d.1) project classification; and (d.2) keyword extraction from titles. The classification system was calibrated using: (1) an explanatory library containing the SDGs, their targets, and indicators (ONU, 2015; Brasil, 2026); and (2) the academic keyword library from Australian universities (*SDSN Australia/Pacific SDG Keyword List*; SDSN, 2026).

d.1 SDG Classification: Each project was assigned: (i) a primary SDG, corresponding to the objective of greatest thematic alignment; (ii) up to two secondary SDGs, when complementary links were evident; and (iii) a confidence level (high, medium, or low) based on the thematic clarity of the title. Tiebreaking criteria were as follows: titles related to technology, automation, or digital systems without a defined thematic application were classified primarily under SDG 9 (Industry, Innovation and Infrastructure); titles referring to pedagogical methods, teacher training, or instructional material development were assigned to SDG 4 (Quality Education). The confidence levels were operationalized as follows: (1) High: the title contains a term directly associated with the classified SDG; (2) Medium: the title suggests thematic alignment but admits alternative interpretations; and (3) Low: the title is generic or ambiguous, making unequivocal SDG assignment difficult. The structured prompt provided to the model included: a summary description of the 18 SDGs, a keyword library per SDG (adapted from SDSN, 2026), classification instructions with explicit tiebreaking criteria, and a standardized tabular output format. Classification was conducted in batches of up to 50 titles per call, in a single, continuous session with no prompt variation between batches.

It is acknowledged that classification based exclusively on project titles constitutes an inherent methodological limitation of the available corpus. Titles are, by nature, concise and do not always reflect the thematic complexity of the initiatives they describe. This limitation calls for caution in interpreting the results, particularly in cases classified with low or medium confidence. The present analysis should be understood as an exploratory approximation rather than a definitive mapping of institutional SDG alignment. Future studies should incorporate abstracts, keywords, and full texts to enhance classification robustness.

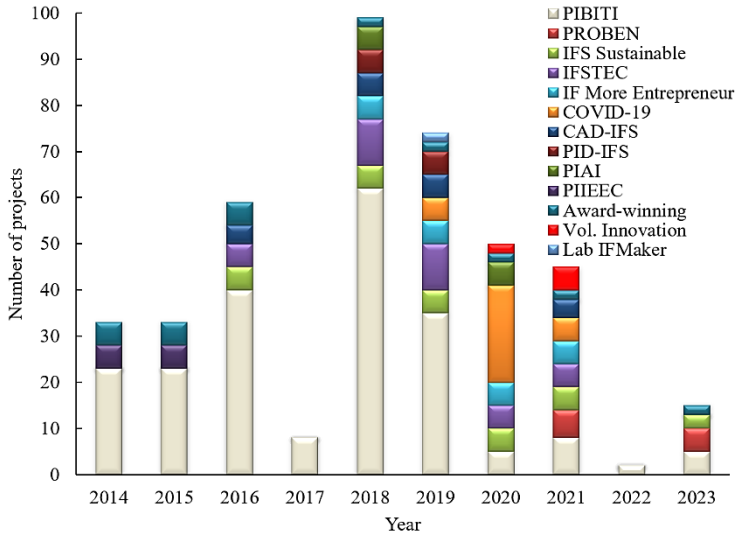
d.2 Bibliometric Keyword Analysis: To identify the most recurrent terms in the corpus, a keyword library organized by SDG was constructed, comprising descriptors validated by their frequency of occurrence in the analyzed titles. The count was performed through Unicode normalization of titles and substring matching, with duplicate control per title. Results were sorted in descending order by absolute frequency and converted to relative frequency (%) in relation to the total of 417 projects, enabling the identification of the dominant thematic axes of institutional innovative output.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

3 Results and Discussion

Between 2014 and 2024, the IFS funded a total of 2,778 projects through its various programs, of which 417 were classified as innovation and technology projects. PIBITI was the most prominent program with 220 projects, followed by IFSTEC (82), IFS Sustainable (38), PROBEN (36), COVID-19 emergency initiatives (21), Voluntary Innovation Projects (14), IF More Entrepreneur (14), CAD-IFS (8), PIIEEC (6), PIAI (6), PID-IFS (5), award-winning projects (4), and Lab IFMaker (3). Notably, 2018 recorded the highest number of approved innovation projects (100). No innovation projects were identified as approved in 2024; thus, the effective analytical scope covers the period from 2014 to 2023 (Figure 1).

Figure 1 – Distribution of innovation projects funded through Innovation/Technology programs at the Federal Institute of Sergipe (IFS) during 2014–2023, based on analysis conducted by Claude Sonnet 4.6.



Source: Authors (2026)

Over the last decade, the largest concentration of IFS projects aligned to SDG 9 (n=233; 55.9%), although these were classified with medium (n=103) or low (n=119) confidence, indicating possible imprecision in the thematic assignment (Table 1). Other SDGs with notable participation included: SDG 4 (39; 9.4%), SDG 8 (31; 7.4%), SDG 7 (28; 6.7%), SDG 2 (27; 6.5%), SDG 6 (20; 4.8%), SDG 3 (18; 4.3%), SDG 12 (9; 2.2%), and SDG 11 (6; 1.4%). By contrast, SDG 16 had only two projects, while SDGs 5 and 10 each had one, evidencing low representativeness.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Table 1 – Distribution of projects funded through Innovation/Technology programs at the Federal Institute of Sergipe (IFS) during 2014–2023, classified according to the Sustainable Development Goals (SDGs), based on analysis by Claude Sonnet 4.6, with respective confidence levels.

SDG	SDG Name	No. Projects	% of Total	High Confidence	Med. Confidence	Low Confidence
SDG 9	Industry, Innovation and Infrastructure	233	55.9%	11	103	119
SDG 4	Quality Education	39	9.4%	8	5	26
SDG 8	Decent Work and Economic Growth	31	7.4%	5	1	25
SDG 7	Affordable and Clean Energy	28	6.7%	3	0	25
SDG 2	Zero Hunger and Sustainable Agriculture	27	6.5%	2	2	23
SDG 6	Clean Water and Sanitation	20	4.8%	5	2	13
SDG 3	Good Health and Well-Being	18	4.3%	0	3	15
SDG 12	Responsible Consumption and Production	9	2.2%	1	2	6
SDG 11	Sustainable Cities and Communities	6	1.4%	1	0	5
SDG 16	Peace, Justice and Strong Institutions	2	0.5%	0	0	2
SDG 5	Gender Equality	1	0.2%	0	1	0
SDG 10	Reduced Inequalities	1	0.2%	0	0	1
SDG 14	Life Below Water	1	0.2%	0	1	0
SDG 15	Life on Land	1	0.2%	0	0	1
TOTAL		**417**	**100%**			

Source: Authors (2026)

The analysis of confidence levels assigned by the classifier reveals a structural limitation that warrants attention proportional to its impact on the validity of results. In the case of SDG 9, which concentrates 55.9% of the projects and constitutes the study's primary finding, only 11 of the 233 cases were classified with high confidence, while 103 received medium confidence and 119 received low confidence. This means that more than half of this article's central finding rests on low-confidence classifications, which requires that quantitative conclusions be interpreted cautiously and treated as exploratory approximations rather than definitive mappings.

This pattern is not exclusive to SDG 9: only 41 projects (9.8% of the total) received high confidence ratings, while medium and low confidence together account for over 90% of all classifications. This distribution reflects an inherent limitation of the available corpus: project titles

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

are, by nature, concise and frequently ambiguous, failing to capture the full thematic complexity of the initiatives they describe. To overcome this limitation in future studies, the incorporation of abstracts, keywords, and, preferably, full project texts as classification input is strongly recommended.

It is further noted that not every good practice possesses structural depth, as SDG alignment depends on concrete conditions of implementation. Future studies should investigate the capacity of initiatives to produce substantive transformations, distinguishing structural actions from those representing only punctual or merely formal contributions, such as sporadic donations, the posting of informational materials without accompanying educational action, or mere compliance with minimum legal obligations (ISS, 2026). It is also acknowledged that the present study did not incorporate a formal human validation protocol for the model's classifications. The specialized literature recommends, for this type of task, procedures such as the manual verification of representative samples and the calculation of concordance coefficients, such as Cohen's Kappa, between the automatic classifier and independent human evaluators (Morales-Hernández *et al.*, 2022; Suadaa *et al.*, 2025).

The assertion, recurring in the literature, that LLMs achieve precision levels comparable to those of human specialists refers to specific corpus, task, and model configuration conditions, and cannot be directly transferred to the present context without empirical verification. Future studies should incorporate this validation protocol to confer greater robustness to the classifications and enable the empirical assessment of model performance on the IFS institutional corpus. In this regard, the processes for identifying sustainability components in higher education institution documents find in LLMs tools of high operational efficiency, but ones that require the creation of annotated datasets for training and testing, enabling models to perform identification in a more precise and auditable manner (Duda *et al.*, 2025; Suadaa *et al.*, 2025).

The concentration of 80% of projects in SDGs 9, 4, 8, and 7 reflects the IFS's institutional mission of integrating vocational training, applied research, and territorial innovation for development (Brasil, 2008; Pacheco, 2010; Rabbani *et al.*, 2025), corroborating the study's initial hypothesis. The prominence of SDG 9 may stem from programs such as PIBITI, which are oriented toward engineering and prototyping. SDG 4 highlights the institution's tradition in professional education, while SDGs 8 and 7 reveal alignment with entrepreneurship and the energy transition, addressing local productive demands and national S&T&I directives. This distribution is not random but mirrors the institutional identity, funding instruments, and regional vocations of Sergipe.

The organization of keywords extracted from project titles reflects the richness of thematic information contained therein, functioning as a mirror of the institution's activities. Only keywords

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

with a frequency of occurrence greater than 0.1% were included in the table (Appendix A), highlighting their contribution to the 2030 Agenda, enabling strategic portfolio management and transparent communication with society (SDSN, 2026). This alignment valorizes the positive aspects of institutional output, such as the diversity of areas served and the coherence with territorial demands, strengthening the IFS's role as a driver of sustainable development.

This study confirmed that PIBITI is the primary funding mechanism for IFS innovation projects (Figure 1). The keyword analysis of project titles confirms the thematic concentration areas associated with the SDGs and reveals the influence of the institution's curricular structure on the profile of its innovative output (Sandrino; Jamuchewski, 2022). Overall, a predominance of SDGs 4, 7, 8, 9, and 12 is observed (Appendix A), a result consistent with the SDG-based classification of project titles (Table 1).

For SDG 4, prominent terms include teaching (5.8%), learning (2.6%), gamification/educational games (1.9%), and maker/DIY (1.9%), reflecting the presence of undergraduate teacher education programs (Physics, Mathematics, Chemistry, and Biological Sciences) and Graduate Programs in Professional and Technological Education, which promote active methodologies and pedagogical innovation. For SDG 7, dominant terms include solar energy (4.1%), photovoltaic (1.2%), and energy efficiency (0.7%), evidencing alignment with Electrical Technology and Renewable Energy Systems programs.

SDG 8 is associated with junior enterprise (3.6%), incubation/startup (1.4%), and business management (0.5%), reflecting programs in tourism, commerce, business administration, and logistics that foster entrepreneurship and business model innovation. SDG 2 features terms such as agroecology (1.2%), aquaponics (1.2%), irrigation (1.0%), and agricultural substrates (1.0%), associated with agricultural science programs including Agroecology, Agribusiness, Animal Science, Aquaculture, and Fishery Resources.

For SDG 6, terms include effluents (1.0%), sewage (0.7%), water monitoring (0.7%), water reuse (0.7%), and sanitation (0.7%). SDG 12 encompasses terms such as sustainability (3.8%), waste (2.9%), reuse (1.4%), and sugarcane ash (1.2%), indicating projects focused on waste management and sustainable technological solutions associated with Environmental Sanitation, Chemistry, Agroecology, and Agroindustry programs.

The prominence of SDG 9 is evidenced by terms such as low cost (6.2%), app/software (3.6%), prototype/prototyping (3.4%), concrete/mortar (3.1%), automation (2.4%), robotics/robot (2.4%), civil engineering (1.9%), BIM (1.4%), artificial intelligence/machine learning (1.4%), Arduino/microcontroller (1.2%), and sensor/instrumentation (1.0%). These results demonstrate a strong technological orientation in IFS innovation projects, primarily related to Engineering,

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Computer Science, and Automation programs. Complementarily, SDG 11 is associated with terms such as local/community tourism (0.7%), housing (0.5%), and smart cities (0.2%), indicating initiatives oriented toward urban planning, housing infrastructure, and territorial development.

Low incidence was observed for terms associated with SDGs 5 (Gender Equality), 10 (Reduced Inequalities), 14 (Life Below Water), 15 (Life on Land), and 18 (Ethno-Racial Equality). This thematic distribution can be interpreted through two main factors. First, the IFS's curricular structure has a relatively lower presence of programs directly oriented toward biodiversity conservation or social science approaches to structural inequalities, which tends to shape the profile of research and innovation projects developed. Second, the methodological scope adopted in this investigation focused exclusively on technological innovation funding programs, excluding research and extension programs, which traditionally exhibit higher incidence of themes related to the social, cultural, and environmental dimensions of sustainable development (Fleck *et al.*, 2025).

Regarding the implications for institutional management of innovation funding programs, the results reveal a strong concentration of projects in SDGs 4, 7, 8, 9, and 12, indicating that IFS output is directly associated with its curricular structure and the technological areas it offers. This pattern suggests convergence between the institution's vocational identity and the applied nature of the solutions developed in the analyzed projects.

In this regard, institutional management can strengthen this strategic positioning through the consolidation of structural applied programs in areas such as renewable energy, automation, sustainable construction, and the circular economy. At the same time, the results identify opportunities for institutional inductive policies aimed at broadening the thematic diversity of innovation. The incorporation of specific thematic lines in institutional funding calls directed at social inclusion, equality, and environmental conservation may stimulate interdisciplinary approaches and expand the institution's potential contribution to socioenvironmental challenges. In this context, strengthening partnerships with territorial actors, including social organizations, local governments, and the productive sector, may favor the emergence of innovation projects oriented toward public problems and regional demands (ISS, 2026).

Additionally, the implementation of institutional monitoring and evaluation instruments for innovation output guided by the SDGs is recommended. Indicator-based tracking systems can enhance the strategic management capacity of educational institutions (Angin *et al.*, 2022; Fuchs *et al.*, 2023; Lemarchand *et al.*, 2023) by enabling the identification of trends, gaps, and alignment opportunities between institutional research agendas and global sustainable development commitments.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

4 Conclusion

The application of an LLM indicated potential for classifying institutional projects according to the SDGs based on their titles, positioning itself as a promising tool for monitoring institutional innovation policies. The results indicate that, over the last decade, the IFS's innovative output exhibits a strong thematic concentration in technological areas, with a particular emphasis on SDG 9 (Industry, Innovation and Infrastructure). This pattern suggests that the institutional innovation agenda is strongly associated with the institution's curricular structure and the technological programs it offers, especially in engineering, automation, computing, and energy. However, the results also highlight relevant methodological limitations, as classification based exclusively on project titles may generate interpretive uncertainties. Therefore, the incorporation of additional metadata and the implementation of human validation protocols are recommended in order to increase analytical precision and the robustness of the obtained classifications.

5 Acknowledgments

This study was carried out with the support of the Sergipe State Foundation for Research and Technological Innovation Support (FAPITEC/SE), through financial resources from the State Fund for Scientific and Technological Development (FUNTEC), managed by the State Secretariat for Economic Development and Science and Technology (SEDETEC), under Calls FAPITEC/SE/FUNTEC No. 18/2023 and No. 12/2025; and with support from the Bahia State Research Foundation (FAPESB/BA).

6 References

ANGIN, M. *et al.* A roberta approach for automated processing of sustainability reports. **Sustainability**, v. 14, n. 23, p. 16139, 2022. Available at: https://doi.org/10.3390/su142316139. Accessed: Mar. 20, 2025.

ANTHROPIC. **Claude: a next-generation AI assistant**. San Francisco: Anthropic, 2024. Available at: https://www.anthropic.com. Accessed: Mar. 20, 2025.

BAUTISTA-PUIG, N. *et al.* Unveiling the Research Landscape of Sustainable Development Goals and Their Inclusion in Higher Education Institutions and Research Centers: Major Trends in 2000–2017. **Frontiers in Sustainability**, v. 2, n. 1, e13, 2021. Available at: https://doi.org/10.3389/frsus.2021.620743. Accessed: Mar. 20, 2025.

BRASIL. **Lei n.º 10.973, de 2 de dezembro de 2004**. Dispõe sobre incentivos à inovação e à pesquisa científica e tecnológica no ambiente produtivo e dá outras providências. Brasília:

Presidência da República, 2004. Available at: https://www.planalto.gov.br/ccivil_03/_ato2004-2006/2004/lei/l10.973.htm. Accessed: Mar. 20, 2025.

BRASIL. **Lei n.º 11.892, de 29 de dezembro de 2008**. Institui a Rede Federal de Educação Profissional, Científica e Tecnológica, cria os Institutos Federais de Educação, Ciência e Tecnologia, e dá outras providências. Brasília: Presidência da República, 2008. Available at: https://www.planalto.gov.br/ccivil_03/_ato2007-2010/2008/lei/l11892.htm. Accessed: Mar. 20, 2025.

BRASIL. **ODS 18: Igualdade Étnico-Racial**. Programas e Projetos. Ministério da Igualdade Racial, [2026]. Available at: https://www.gov.br/igualdaderacial/pt-br/assuntos/programas-e-projetos/ods18. Accessed: Mar. 28, 2026.

CASSIOLATO, J. E.; LASTRES, H. M. M. Sistemas de inovação e desenvolvimento: as implicações de política. **São Paulo em Perspectiva**, v. 19, n. 1, p. 34–45, 2005. Available at: https://doi.org/10.1590/S0102-88392005000100003. Accessed: Mar. 20, 2025.

DUDA, R. F. *et al.* Inteligência Artificial Generativa na Gestão de IES: Contribuições para o Mapeamento dos Objetivos de Desenvolvimento Sustentável. **Modelos Educativos Innovadores en Educación Superior**, v. 3, n. 1, p. 224-240, 2025. Available at: https://doi.org/10.33132/27449394.2497. Accessed: Mar. 20, 2025.

FLECK, L. *et al.* Programa Itaipu Parquetec: extensão universitária como vetor para a sustentabilidade territorial. **Revista Cultivando o Saber**, v. 18, n. 1, p. 13-26, 2025. Available at: https://cultivandosaber.fag.edu.br/index.php/cultivando/article/view/1333. Accessed: Mar. 20, 2025.

FUCHS, P. G. *et al.* Sustainable Development Goals (SDGs) in Latin-American Universities. **Sustainability**, v. 15, n. 11, p. 1–23, 2023. Available at: https://doi.org/10.3390/su15118556. Accessed: Mar. 20, 2025.

IBGE. **Censo Demográfico 2022: resultados do universo**. Rio de Janeiro: IBGE, 2023. Available at: https://www.ibge.gov.br/estatisticas/sociais/populacao/22827-censo-demografico-2022.html. Accessed: Mar. 20, 2025.

IFS – INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DE SERGIPE. **Notícias**. Aracaju: IFS, 2026. Available at: https://www.ifs.edu.br/ultimas-noticias.html. Accessed: Feb. 15, 2026.

ISS. **Guia de metas locais: ações viáveis no território**. Brasília: Instituto Selo Social, 2026. 62 p. Available at: https://www.calameo.com/read/007362143c189175a40b2. Accessed: Mar. 28, 2026.

LEMARCHAND, P. *et al.* An evaluation of a computational technique for measuring the embeddedness of sustainability in the curriculum aligned to AASHE-STARS and the United Nations Sustainable Development Goals. **Frontiers in Sustainability**, v. 4, n. 1, p. 1-18, 2023. Available at: https://doi.org/10.3389/frsus.2023.997509. Accessed: Mar. 20, 2025.

MORALES-HERNÁNDEZ, R. C. *et al.* A comparison of multi-label text classification models in research articles labeled with sustainable development goals. **IEEE Access**, v. 10, n. 1, p. 123534-123548, 2022. Available at: https://doi.org/10.1109/ACCESS.2022.3223094. Accessed: Mar. 20, 2025.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

OCDE. **Oslo Manual 2018**: guidelines for collecting, reporting and using data on innovation. 4th ed. Paris: OECD Publishing, 2018.

ONU. **Transformando nosso mundo: a agenda 2030 para o desenvolvimento sustentável**. New York: United Nations, 2015. Available at: https://brasil.un.org/pt-br/91863-agenda-2030-para-o-desenvolvimento-sustent%C3%A1vel. Accessed: Mar. 20, 2025.

PACHECO, E. (org.). **Institutos Federais: uma revolução na educação profissional e tecnológica**. Brasília: Fundação Santillana; São Paulo: Moderna, 2010. Available at: https://memoria.ifrn.edu.br/bitstream/handle/1044/1013/Os%20institutos%20federais%20-%20Ebook.pdf. Accessed: Mar. 20, 2025.

RABBANI, A. R. C. *et al.* Os ODS como catalisadores de inovação nos institutos federais. In: **Semana Nacional de Ciência e Tecnologia**, 2025, Aracaju. Anais [...]. Aracaju: IFS, 2025. Available at: https://periodicos.ifs.edu.br/periodicos/SNCT/issue/archive.

SANDRINO, B.; JAMUCHEWSKI, L. B. M. ODS e atividades de pesquisa em cursos superiores de engenharias no decorrer do tempo de vida de uma instituição de ensino paranaense. **Faculdade Sant'Ana em Revista**, v. 6, n. 1, p. 47-59, 2022. Available at: https://iessa.edu.br/revista/index.php/fsr/article/view/2151. Accessed: Mar. 20, 2025.

SDSN. **Compiled Keywords for SDG Mapping**. UN Sustainable Development Solutions Network (SDSN) Australia/Pacific, [2017]. Available at: https://ap-unsdsn.org/wp-content/uploads/2017/04/Compiled-Keywords-for-SDG-Mapping_Final_17-05-10.xlsx. Accessed: Mar. 28, 2026.

SUADAA, L. A. *et al.* Automatic Classification of Multilanguage Scientific Papers to the Sustainable Development Goals Using Transfer Learning. **Jurnal RESTI (Rekayasa Sistem dan Teknologi Informasi)**, v. 9, n. 3, p. 604–612, 2025. Available at: https://doi.org/10.29207/resti.v9i3.6560. Accessed: Mar. 20, 2025.

UNESP. **Guia Agenda 2030:** integrando ODS, educação e sociedade. São Paulo: Universidade Estadual Paulista, 2025. 158 p. Available at: https://www2.unesp.br/portal#!/unesp2030. Accessed: Mar. 28, 2026.

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

Appendix A

Appendix A – Keywords identified in the titles of projects funded through Innovation/Technology programs at the Federal Institute of Sergipe (IFS) during 2014–2023, classified according to the Sustainable Development Goals (SDGs) and their frequency of occurrence, based on analysis conducted by Claude Sonnet 4.6.

SDG	**Name**	**Keywords**
SDG 1	Eradication of Poverty	Cooperativism (0.7%); Small farmer (0.5%); Settlement (0.2%); Rural communities (0.2%)
SDG 2	Zero Hunger and Sustainable Agriculture	Agroecology (1.2%); Aquaponics (1.2%); Irrigation (1.0%); Agricultural substrate (1.0%); Agriculture (0.7%); Functional foods (0.7%); Biofertilizer (0.7%); Food fermentation (0.7%); Horticulture (0.7%); Pest management (0.7%); Family farming (0.5%); Aquaculture/Fish farming (0.5%); Seedling production (0.5%); Agrobiodiversity (0.2%); Agricultural cooperativism (0.2%); Sustainable cultivation (0.2%); Heirloom seeds (0.2%); Silviculture/Forestry (0.2%)
SDG 3	Good Health and Well-Being	COVID-19 (2.6%); Disinfection/Sanitizer (1.4%); PPE/Personal protection (1.4%); Alcohol and sanitizing agents (1.2%); Herbal medicine/Medicinal plant (1.0%); Coliforms/Microbiology (0.7%); Infectious diseases (0.5%); Sunscreen (0.5%); UV rods/Ultraviolet (0.5%); Diagnostic imaging (0.2%); Hygiene/Clean hands (0.2%); Occupational health (0.2%); Vaccines/Immunology (0.2%)
SDG 4	Quality Education	Teaching (5.8%); Learning (2.6%); Gamification/Educational game (1.9%); Maker/DIY (1.9%); Augmented/Virtual reality (1.4%); Education (1.2%); Chemistry teaching (0.7%); Active methodology (0.7%); Distance learning/Virtual environment (0.5%); Mathematics teaching (0.2%); Teacher training (0.2%); Scientific literacy (0.2%); Teaching materials (0.2%); Educational robotics (0.2%); STEM/STEAM (0.2%); Educational simulation (0.2%)
SDG 5	Gender Equality	Gender (0.5%); Masculinity (0.2%); Gender-based violence (0.2%)
SDG 6	Clean Water and Sanitation	Effluent (1.0%); Sewage (0.7%); Water monitoring (0.7%); Water reuse (0.7%); Sanitation (0.7%); Automated irrigation (0.5%); Water quality (0.5%); Water treatment (0.5%); Wetland (0.5%); Adsorption/Contaminant removal (0.2%); Rainwater harvesting (0.2%); Water waste (0.2%); Desalination (0.2%); Urban sludge (0.2%); Porous stone/Filter (0.2%); Well/Groundwater (0.2%)
SDG 7	Affordable and Clean Energy	Solar energy (4.1%); Photovoltaic (1.2%); Biodiesel (0.7%); Energy efficiency (0.7%); On-grid photovoltaic system (0.7%); Aquaponics with solar energy (0.5%); Hybrid energy (0.5%); Solar vehicle/train (0.5%); Wind turbine/Aeolian (0.2%); Biofuel (0.2%); Biohydrogen (0.2%); Energy demand (0.2%); Wireless energy (0.2%); Current inverter (0.2%); Hydraulic microturbine (0.2%); Solar dryer/Cooker (0.2%); Solar tracker (0.2%)

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

SDG	Name	Keywords
SDG 8	Decent Work and Economic Growth	Junior enterprise (3.6%); Tourism/Ecotourism (1.9%); Incubation/Startup (1.4%); Advisory/Consulting (1.2%); Agroecological experience (1.2%); Cooperativism (0.7%); Entrepreneurship (0.7%); Digital marketing (0.7%); Business management (0.5%); Growth Hacking (0.5%); Geographic indication/Crafts (0.5%); Business/Product (0.5%); New digital entrepreneurs (0.5%); Local/Industrial economy (0.2%)
SDG 9	Industry, Innovation and Infrastructure	Low cost (6.2%); App/Software (3.6%); Prototype/Prototyping (3.4%); Concrete/Mortar (3.1%); Automation (2.4%); Robotics/Robot (2.4%); Civil engineering (1.9%); BIM (1.4%); Artificial intelligence/ML (1.4%); Arduino/Microcontroller (1.2%); Sensor/Instrumentation (1.0%); Technological innovation (0.7%); RFID/Telemetry (0.7%); Remote monitoring (0.5%); Computational simulation (0.5%); Drone (0.2%); IoT/Internet of Things (0.2%)
SDG 10	Reduced Inequalities	Disability/PwD (1.0%); Reduced mobility (0.7%); Accessibility (0.5%); Exoskeleton (0.5%); Visual impairment (0.2%); Social inclusion (0.2%); Assistive technology (0.2%); Smart glasses (0.2%)
SDG 11	Sustainable Cities and Communities	Local/Community tourism (0.7%); Housing (0.5%); Smart cities (0.2%); Environmental comfort (0.2%); Social housing (0.2%); Temporary housing (0.2%); Rural infrastructure (0.2%); Building modeling (BIM) (0.2%); Urban landscape (0.2%); Historical heritage (0.2%); Metropolitan region (0.2%); Urban signage (0.2%)
SDG 12	Responsible Consumption and Production	Sustainability (3.8%); Waste (2.9%); Reuse (1.4%); Sugarcane ash (1.2%); Biosurfactant (0.7%); Cassava waste (0.7%); Bioplastic (0.5%); Composting (0.5%); Ecological alternative product (0.5%); Recycling (0.5%); Whey/Dairy waste (0.5%); Residual frying oil (0.5%); Sustainable material (concrete) (0.2%); Recycled tire/Rubber (0.2%); Construction waste (0.2%); Glass waste (0.2%); Reutilization (0.2%)
SDG 13	Climate Action	Environmental sustainability (0.7%); Caatinga/Semi-arid (0.5%); Sustainability barometer (0.2%); Climate/Climate change (0.2%); Environmental education (0.2%); Environmental management (0.2%); Environmental monitoring (0.2%)
SDG 14	Life Below Water	Hydroponics/Aquaponics (2.2%); Metal biosorption (0.5%); Aquaculture water quality (0.5%); Aquaculture/Fish farming (0.2%); Red crab (0.2%); Marine shrimp (0.2%); Mud crab (0.2%); Shellfish/Mollusks (0.2%); Aquatic probiotic (0.2%); River/Fluvial ecosystem (0.2%); Fish oil/Aquatic biofuel (0.2%)
SDG 15	Life on Land	Land use/Agroecology (1.2%); Flora/Plants (0.7%); Plant health/IPM (0.5%); Agrobiodiversity (0.2%); Biodiversity (0.2%); Caatinga (0.2%); Medicinal plants/Herbal medicine (0.2%); Heirloom seeds (0.2%); Semi-arid (0.2%)
SDG 16	Peace, Justice and Strong Institutions	Technological prospecting (0.5%); Citizenship (0.2%); Public management (0.2%); Process mapping (0.2%); Intellectual property (0.2%); Data security (0.2%)

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual

SDG	Name	Keywords
SDG 17	Partnerships for the Goals	Information/Extension network (0.7%); Network training (0.5%); Technological prospecting (0.5%); Community tourism (network) (0.5%); Rural internet access (0.2%); Certification/Competencies (0.2%); Junior enterprise (network) (0.2%); Geographic indication (0.2%); Institutional partnerships (0.2%); Intellectual property (network) (0.2%)
SDG 18	Ethno-Racial Equality	No keywords with occurrence equal to or greater than 0.1% were identified.

Source: Authors (2026)

Submetido: 24 de julho de 2026. Aceito: 04 de setembro de 2026. Aracaju/SE
API-Associação Acadêmica de Propriedade Intelectual