International Journal of
Technology and
Systems
(IJTS)

Natural Language Generation (NLG) for Automated Report Generation

Journal of Technology and Systems
ISSN : 2788-6344 (Online)
Vol. 5, Issue No. 1, pp 48 – 59, 2023
www.carijournals.org

Natural Language Generation (NLG) for Automated Report Generation

[1*]Deepak Nanuru Yagamurthy, [2]Rajesh Azmeera, [3]Rhea Khanna

[1*]Technology Professional, Department of Information and Technology, 7 Eleven Inc, US

[2]Technology Professional, Department of Information and Technology, Stryker, US

[3]Technology Professional, Department of Research, Microsoft, US

https://orcid.org/0009-0009-9546-6615

Accepted: 16th Oct 2023 Received in Revised Form: 2nd Nov 2023 Published: 9th Nov 2023

Abstract

Automated report generation plays a vital role in numerous industries, enabling efficient and accurate communication of complex data. Natural Language Generation (NLG) is a powerful technique that leverages artificial intelligence (AI) to transform structured data into human-readable narratives. This white paper provides an overview of NLG and its applications in automated report generation. We explore the benefits, challenges, and best practices associated with NLG-based report generation, along with real-world examples. Furthermore, we discuss the current state of the technology, future prospects, and potential ethical considerations.

Keywords*:* *Natural Language Generation, Machine Learning, Artificial Intelligence and Quality Assurance*

1. INTRODUCTION

Source: Deep Learning Adaptive Computation and Machine Learning Series II, Natural Language Generation. SinceDirect, Gartner: Natural Language Generation (NLG) Market Overview, Building Natural Language Generation Systems.Cambridge University Press, Deep Learning Adaptive Computation and Machine Learning Series II

In today's data-driven world, organizations often face the challenge of interpreting and communicating vast amounts of structured data effectively. Traditional manual report generation processes can be time-consuming, error-prone, and lack consistency. Natural Language Generation (NLG) offers a solution by automating the transformation of structured data into coherent and human-like narratives. This white paper explores the potential of NLG for automated report generation across various domains and discusses the benefits, challenges, best practices, and prospects of this technology.

2. **Natural Language Generation (NLG): An Overview**

2.1. **Definition and Components**:

Natural Language Generation involves the generation of human-readable text by machines. It encompasses several components, including data analysis, data-to-text mapping, sentence planning, and surface realization. These components work together to convert structured data into narratives that convey meaningful information to end-users.

2.2. **NLG Techniques**:

NLG techniques can be rule-based, template-based, or machine learning-based. Rule-based NLG relies on predefined linguistic rules and templates to generate text. Template-based NLG utilizes prebuilt sentence templates with slots that are filled with data. Machine learning-based NLG leverages deep learning models, such as recurrent neural networks or transformers, to generate text based on learned patterns and structures in training data.

3. **Applications of NLG in Automated Report Generation**

3.1. **Financial Reporting**:

NLG automates the generation of financial reports, including earnings reports, shareholder letters, and investment summaries. It enables organizations to communicate financial insights and performance metrics in a concise and personalized manner. It enables the transformation of structured financial data, such as balance sheets, income statements, and cash flow statements, into human-readable narratives. In financial reporting, NLG finds applications in various areas. NLG also plays a crucial role in generating investment summaries, allowing asset managers and investors to receive personalized reports summarizing portfolio analysis, market trends, and performance metrics. Additionally, NLG facilitates the generation of shareholder letters, enabling companies to communicate financial highlights and strategies in a personalized and informative

manner. Compliance reporting is another area where NLG proves valuable, automating the process of generating regulatory reports, ensuring accuracy, and consistency in financial disclosures. By leveraging NLG technology, financial organizations can streamline their reporting processes, saving time and effort while delivering accurate and meaningful reports to stakeholders.

Figure1~ProductAnalysis

Here are some specific applications and examples:

3.1.1. Earnings Reports:

Earnings reports are crucial for both publicly traded companies and investors. NLG can automatically generate comprehensive earnings reports from raw financial data.

Example:

"In the last fiscal quarter, Company X reported a 15% increase in revenue, driven by strong sales in its software division. Operating expenses remained stable, resulting in a 20% growth in net income. The company's earnings per share (EPS) increased from $1.20 to $1.45, reflecting a 20.83% growth compared to the same quarter last year."

3.1.2. Investment Summaries:

Investment professionals rely on reports that summarize the performance of investment portfolios. NLG can create personalized investment summaries, highlighting key metrics and trends.

Example:

"Your investment portfolio performed well last quarter, with a 12% return on investment. The technology sector was the primary driver, contributing to a 25% increase in your portfolio's value. Notably, your investments in Company Y saw significant growth, resulting in a 30% ROI for that specific holding."

3.1.3. Shareholder Letters:

NLG can generate personalized shareholder letters that provide an overview of a company's financial health, strategic initiatives, and future outlook.

Example:

"Dear Shareholders,

I am pleased to report that Company Z achieved another year of growth and profitability. Our revenues increased by 8% compared to the previous year, and our net income rose by 12%. This success is attributed to our new product launches and expanded market presence. We are committed to continuing this trajectory and delivering long-term value to our shareholders."

3.1.4. Compliance Reporting:

NLG plays a significant role in automating compliance reporting, particularly in the financial industry. It ensures accurate and consistent generation of regulatory reports.

Example:

"In accordance with regulatory requirements, Company A has prepared the annual report detailing its financial activities. The report includes an analysis of financial statements, risk assessments, and adherence to regulatory guidelines. Our compliance review indicates that the company has maintained accurate records and adhered to all financial regulations throughout the fiscal year."

3.2. **Business Intelligence Reporting**:

NLG transforms complex business intelligence data into user-friendly reports, empowering decision-makers with clear and actionable insights. It facilitates the interpretation of analytics dashboards, performance metrics, and key performance indicators (KPIs). Business Intelligence (BI) reporting is essential for organizations to make data-driven decisions, but interpreting complex data and communicating insights can be challenging. NLG is transforming the field of BI reporting by automating the generation of human-readable narratives from data visualizations and analytics dashboards. NLG algorithms analyze the underlying data and translate it into natural language, enabling users to gain valuable insights without needing to dive into complex charts and graphs. NLG-powered BI reporting provides a concise summary of key performance indicators (KPIs), sales trends, market analysis, and other business metrics. It enables decision-makers to understand the story behind the data and take informed actions. With NLG, organizations can streamline their reporting processes, increase the accessibility of data-driven insights, and empower users at all levels to make data-informed decisions, leading to improved business performance.

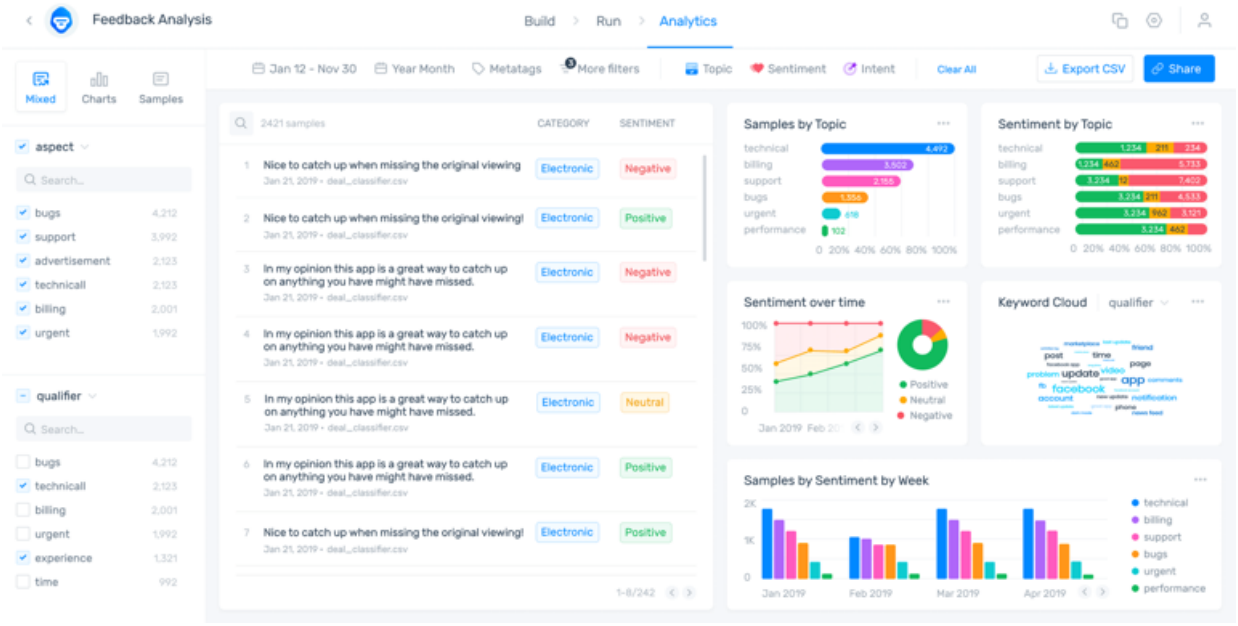

Figure2~DashboardSample

Here are some applications and examples:

3.2.1. Key Performance Indicator (KPI) Reports:

NLG can automatically generate reports that summarize KPI data and provide contextual narratives.

Example:

"In the last quarter, Company A saw a 15% increase in website traffic, primarily driven by an effective social media campaign. This surge in online engagement resulted in a 10% rise in sales, exceeding our targets. The improved customer retention rate, now at 85%, indicates a successful implementation of our customer loyalty program."

3.2.2. Market Analysis Reports:

NLG can analyze market data and generate reports that outline market trends, competitor analysis, and opportunities for growth.

Example:

"Our market analysis report reveals that the healthcare sector is experiencing rapid growth, with a 20% increase in demand for telehealth services. Competitor analysis shows that Company B is gaining market share due to its innovative patient engagement platform. To capitalize on this trend, we plan to invest in similar technology and expand our telehealth services."

3.2.3. Sales and Revenue Reports:

Sales and revenue reports are vital for understanding the financial health of a business. NLG can summarize this data, highlighting key achievements and areas for improvement.

Example:

"Last month, Company C achieved a record-breaking sales performance with a 25% increase in revenue. The sales team's efforts in targeting the enterprise segment were particularly successful, resulting in a 35% increase in sales within that sector. The revenue growth is attributed to both increased sales volume and improved pricing strategies."

3.2.4. Customer Feedback and Satisfaction Reports:

NLG can analyze customer feedback and satisfaction survey data to generate reports that provide insights into customer sentiment.

Example:

"Our customer feedback report reveals a consistent improvement in customer satisfaction over the past year. The implementation of a chatbot for customer support has resulted in a 30% decrease in response times and a 15% increase in overall customer satisfaction. Customers particularly appreciate the convenience of 24/7 support."

3.2.5. Budget and Expense Reports:

NLG can automate the generation of reports related to budgets and expenses, providing insights into cost management and allocation.

Example:

"The budget report indicates that we are operating within the allocated budget for all departments. However, a detailed expense analysis reveals that the marketing department exceeded its budget due to increase advertising spend. This expenditure was in line with our strategy to boost brand awareness, resulting in a 12% increase in customer inquiries."

3.3. **Healthcare Reporting**:

In healthcare, NLG can assist in generating clinical reports, patient summaries, and personalized treatment plans. It enables healthcare professionals to communicate medical information accurately and efficiently. In the healthcare industry, accurate and timely reporting is crucial for effective patient care, clinical decision-making, and compliance with regulations. NLG is playing a transformative role in automating healthcare reporting processes. NLG algorithms can analyze large volumes of structured healthcare data, including electronic health records (EHRs), medical imaging reports, and patient outcome data, and convert them into meaningful narratives. NLG-powered healthcare reporting enables the generation of clinical reports, patient summaries, and treatment plans that are clear, concise, and tailored to individual patients. Healthcare professionals can leverage NLG to communicate medical information accurately and efficiently, enhancing

collaboration among care teams and improving patient outcomes. NLG also facilitates the creation of clinical decision support systems, where real-time patient data is translated into actionable insights, aiding physicians in making informed treatment decisions. By automating healthcare reporting, NLG improves the efficiency, accuracy, and accessibility of critical medical information, ultimately enhancing the quality of care provided to patients.

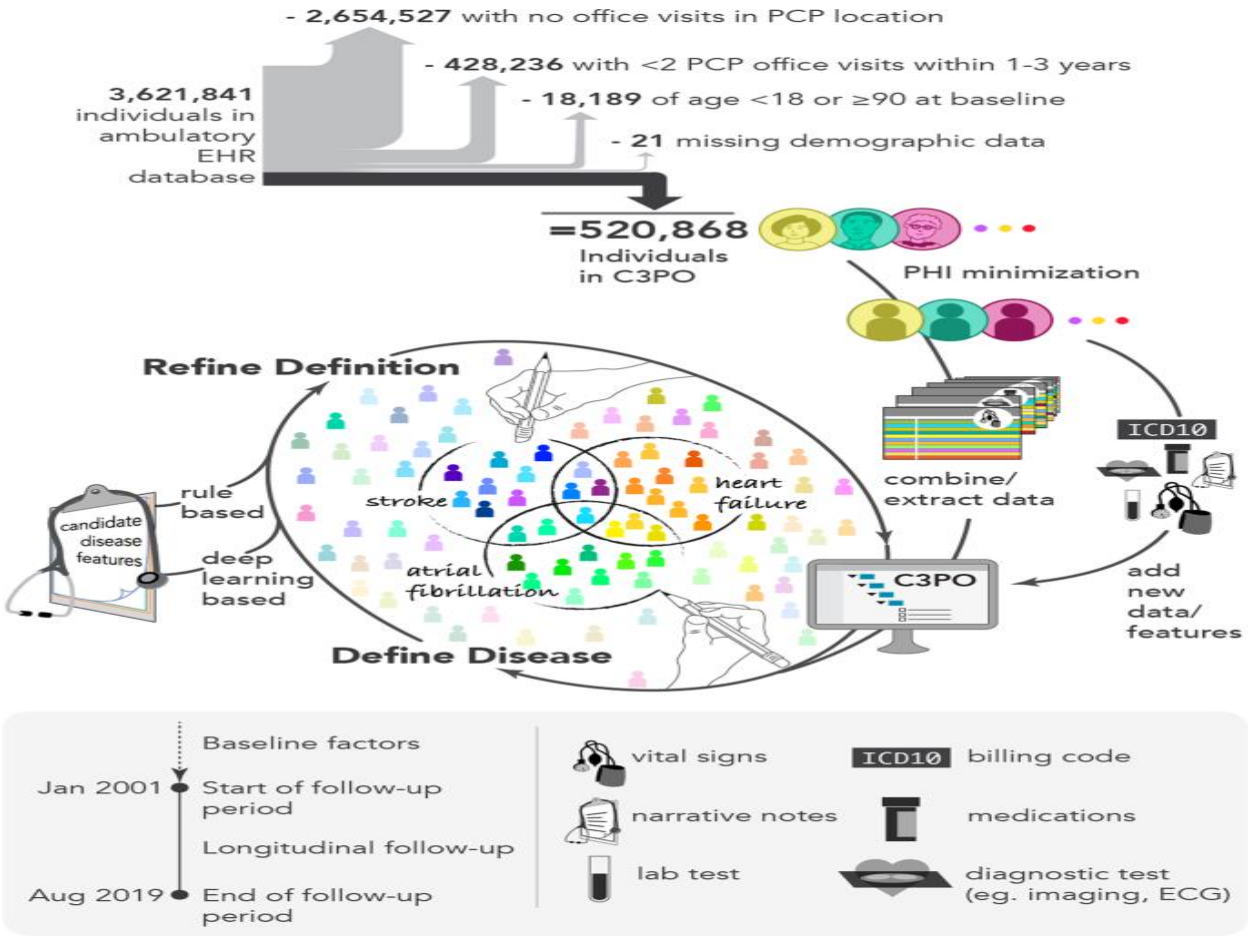

Figure 3~Sample Arch

Here are some applications and examples:

3.3.1. Patient History and Medical Reports:

NLG can be used to automatically generate patient history reports and medical summaries based on electronic health records (EHRs).

Example:

"Patient X's medical history report includes information on past illnesses, surgeries, and medications. Notably, the patient has a history of hypertension and was prescribed medication for

blood pressure control. The recent lab results show stable blood pressure levels, indicating an effective treatment plan."

3.3.2. Discharge Summaries:

Upon a patient's discharge from a healthcare facility, NLG can assist in creating detailed summaries of the treatment and care provided during the stay.

Example:

"Mr. Y was discharged following a three-day hospitalization. He was treated for pneumonia with antibiotics and respiratory therapy. Mr. Y's condition significantly improved during his stay, and he is now prescribed a two-week course of antibiotics to complete his treatment."

3.3.3. Public Health Reporting:

NLG is also applied in generating public health reports. For instance, during an outbreak, NLG can automatically create reports summarizing the number of cases, geographic spread, and recommendations for containment.

Example:

"Our public health report on the recent flu outbreak in Region Z indicates a total of 150 reported cases. Most cases have been clustered in urban areas. We recommend increased vaccination campaigns and public awareness to mitigate further spread."

3.3.4. Clinical Trial Reports:

In the context of clinical trials, NLG can be used to produce summaries of trial results, ensuring easy comprehension by both researchers and regulatory authorities.

Example:

"The clinical trial report for Drug Candidate A shows promising results. Patients in the experimental group experienced a 30% reduction in symptoms compared to the control group. This significant improvement supports the drug's efficacy in treating the condition."

3.3.5. Quality Improvement Reports:

Healthcare institutions use NLG to create quality improvement reports, which outline areas of improvement and recommended actions.

Example:

"Our quality improvement report identifies a delay in emergency room wait times as an area of concern. To address this issue, we recommend increasing staffing during peak hours and implementing a new triage system to prioritize critical cases."

3.3.6. Compliance and Regulatory Reports:

Healthcare organizations rely on NLG to produce compliance reports, which ensure adherence to healthcare regulations.

Example:

"The compliance report confirms that our facility is fully compliant with all state and federal healthcare regulations. We have maintained accurate records, ensured staff certifications, and conducted regular safety drills as required by law."

3.4. Research and Data Analysis Reporting:

NLG supports the generation of research reports, market analysis summaries, and scientific publications. It facilitates the dissemination of complex findings and insights to a broader audience.

4. **Benefits and Challenges of NLG in Automated Report Generation**

4.1. Benefits:

Benefit	Description
Increased efficiency	NLG automates the report generation process, saving time and effort compared to manual methods.
Consistency and standardization	NLG ensures consistency in reporting, reducing human errors and discrepancies.
Personalization	NLG enables customization of reports based on user preferences and requirements.
Scalability	NLG can generate reports at scale, accommodating large volumes of data.

4.2. Challenges:

Challenge	Description
Data quality and structure	NLG requires high-quality, well-structured data to generate accurate and meaningful narratives.
Contextual understanding	Capturing contextual nuances and generating coherent and contextually appropriate narratives can be challenging.
Domain-specific language	NLG systems need domain-specific knowledge to generate accurate and domain-specific reports.
Ethical considerations	NLG-powered reports should address ethical concerns such as bias, transparency, and data privacy.

5. **Best Practices for NLG-based Report Generation**

5.1. Data Preparation and Structuring:

- Ensure data quality and integrity through data cleaning and validation processes.

- Structure data in a way that facilitates effective data-to-text mapping.

- Consistent data schema simplifies the mapping process and improves the overall quality and coherence of the generated reports.

- Standardize categorical data by applying consistent labels or encoding methods to ensure uniformity and avoid ambiguity.

- Clearly document the chosen approach to maintain transparency and ensure the appropriate interpretation of the reports.

- Integrate domain-specific rules and guidelines into the NLG system to capture the nuances and intricacies of the subject matter accurately.

5.2. Language Generation Models:

- Select and fine-tune language generation models suitable for the specific report generation task.

- Train models on high-quality and diverse datasets to improve the quality of generated reports.
- Incorporate ethical considerations related to transparency, accountability, and privacy when using language generation models for report generation.

- Incorporate user feedback and conduct objective evaluations to identify areas for improvement and refine the language generation models accordingly.

5.3. Post-Processing and Quality Assurance:

- Implement post-processing steps to refine and enhance the generated reports.

- Establish quality assurance measures to validate the accuracy and coherence of the generated narratives.
- Implement feedback loops that allow users or stakeholders to provide input on the quality and relevance of the generated reports and incorporate their feedback for continuous improvement.

6. **Real-world Examples of NLG-powered Reporting Systems**

6.1. Automated Financial Report Generation:

Companies like **Narrative Science Quill** and **Arria NLG** leverage NLG to automate the generation of financial reports for clients, transforming raw financial data into narrative-based reports.

6.2. Business Intelligence Dashboards:

NLG is used in business intelligence tools like **Power BI** and **Tableau** to automatically generate textual summaries and insights based on the underlying data visualizations.

6.3. Clinical Decision Support Systems:

NLG assists in generating patient-specific clinical reports and treatment plans in healthcare settings, aiding healthcare professionals in making informed decisions.
- Epic Systems, a leading healthcare software company, developed the Deterioration Index, which uses NLG to analyze patient data and generate real-time reports.
- IBM Watson for Oncology is an NLG-powered CDSS designed to support oncologists in treatment planning for cancer patients.

- Cerner, a healthcare technology company, offers HealtheRegistries, an NLG-powered CDSS that assists in population health management and quality reporting.

7. **Current State of NLG Technology**

7.1. NLG Tools and Platforms:

Various NLG tools and platforms are available, including OpenAI's GPT-3, Arria NLG Studio, and Automated Insights' Wordsmith. These tools provide APIs and interfaces to facilitate NLG integration into existing systems.

7.2. Limitations and Areas for Improvement:

NLG technology still faces challenges in generating truly human-like narratives, understanding context, and handling complex linguistic phenomena. Research is ongoing to improve NLG models' contextual understanding and enhance their ability to handle domain-specific language.

8. **Future Prospects and Potential Ethical Considerations**

The future of NLG in automated report generation holds great potential. Advancements in AI and NLG techniques are likely to result in more accurate, coherent, and context-aware report generation systems. However, ethical considerations surrounding bias, transparency, and data privacy must be carefully addressed to ensure responsible and fair usage of NLG technology.

9. **Conclusion**

Natural Language Generation (NLG) offers significant benefits for automated report generation across diverse industries. By automating the transformation of structured data into human-readable narratives, NLG enables efficient communication of complex.

10. **Declarations:**

- **Ethical Approval : Not Applicable**
- **Funding: Not Applicable**
- **Availability of data and materials: Not Applicable**

11. References:

[1]. Natural Language Generation. SinceDirect. [Online]. Available at:
https://www.sciencedirect.com/topics/social-sciences/natural-language-generation

[2]. National Library of Medicine. NCBI. [Online]. Available at:
https://www.ncbi.nlm.nih.gov/pmc/articles/PMC61265/

[3]. Gartner: Natural Language Generation (NLG) Market Overview. [Online]. Available at:
https://www.gartner.com/en/documents/3942031

[4]. Ehud Reiter and Robert Dale. [Press]. Building Natural Language Generation Systems.Cambridge University Press.

[5]. Christopher D. Manning and Hinrich Schütze. [Book]. Automated Insights: Narrative Science for Business Intelligence.

[6]. Daniel T. Hindle. [Book]. Natural Language Generation for Business.

[7]. Ian Goodfellow, Yoshua Bengio, Aaron Courville, [Book]. Deep Learning (Adaptive Computation and Machine Learning series 1.

[8]. Ian Goodfellow, Yoshua Bengio, Aaron Courville [Book]. Deep Learning Adaptive Computation and Machine Learning Series II

[9]. John Paul Mueller, Luca Massaron, [Book]. Deep Learning for Dummies

[10]. Image Recognition with Deep Neural Networks and its Use Cases, [Online]. Available at: https://www.altexsoft.com/blog/image-recognition-neural-networks-use-cases/