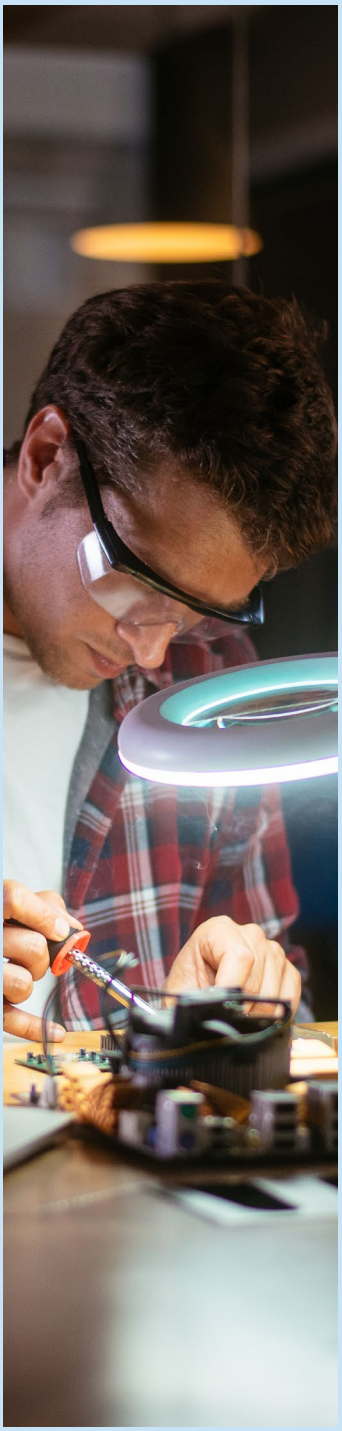
VIEW POINT



NATURAL LANGUAGE GENERATION'S CONTRIBUTION IN LIFE-SCIENCE INDUSTRY

Infosys®
Navigate your next

Table of the content

Enhancing Pharma Efficiency in NLG 3
Types of NLG 3
NLG Tools (Market & Open Source) 4
Parameters for Selecting Suitable NLG Tools 4
NLG in PBI Dashboards 6
Limitation of NLG tools integrating with PBI dashboard 6
NLG Use-case in Pharma Industry 7

External Document © 2023 Infosys Limited

Enhancing Pharma Efficiency in NLG

AI has been transforming the healthcare and pharma industries by introducing new ways of improving patient care, enhancing operational efficiencies, and streamlining decision-making processes. One of the key applications of AI usage in healthcare is the application of Natural Language Generation (NLG) technology to interpret the results of medical scans like CT, X-ray, and MRI.

Traditionally, radiologists and other medical professionals have had to manually review these scans to detect any abnormalities or disorders. However, with the help of AI and NLG, this process can be automated, resulting in faster and more accurate diagnoses.

AI algorithms use computer vision and deep learning to analyze the images and other patient-related data from medical scans. These algorithms can detect and identify even the slightest anomalies in the images that may not be visible to the human eye. Once the analysis is complete, the findings are presented in the form of structured tables and images.

NLG technology is then used to convert these structured data into natural language text reports that can be easily understood by medical professionals. These reports highlight the key points that need attention and help doctors make informed decisions about the patient's care. The use of NLG in medical imaging analysis can significantly reduce the time and effort required for manual reviews, enabling healthcare professionals to diagnose and treat patients faster. Moreover, NLG can also help in standardizing the interpretation of medical scans, minimizing the chances of errors and inconsistencies in diagnosis.

The integration of NLG tools with Power BI for Pharma and Medical Devices requires careful planning and execution to ensure that the resulting insights are accurate, relevant, and compliant with industry regulations.

Types of NLG

Based on its extent, Natural Language Generation (NLG) in AI is classified into three types: -

Primary Level NLG : The most fundamental level of Natural Language Generation involves extracting and consolidating key data points to generate coherent and informative sentences. As an illustration, a basic weather report could be presented as follows: "The air humidity today is 78%."

Template-driven NLG : Template-driven NLG, also referred to as the advanced stage, operates by employing template-centric paragraphs to dynamically generate language. This approach heavily relies on meticulously coded rules featuring pre-determined text, designated placeholders, and specialized data representations. The generation process is guided by well-defined business rules, often implemented through iterative commands such as if/else statements. Stock market updates, sports score charts and basic business reports can be made using this type of NLG.

Advanced Level NLG : Advanced NLG tools exhibit greater versatility compared to basic and template-driven approaches. They leverage the power of Machine Learning to transform data into coherent narratives, featuring well-defined introductions, comprehensive elaborations, and conclusive remarks. It relies on deep learning neural networks that adeptly capture and internalize lexical, morphological and grammatical patterns inherent in written language, enabling them to deliver highly refined and contextually appropriate outputs.

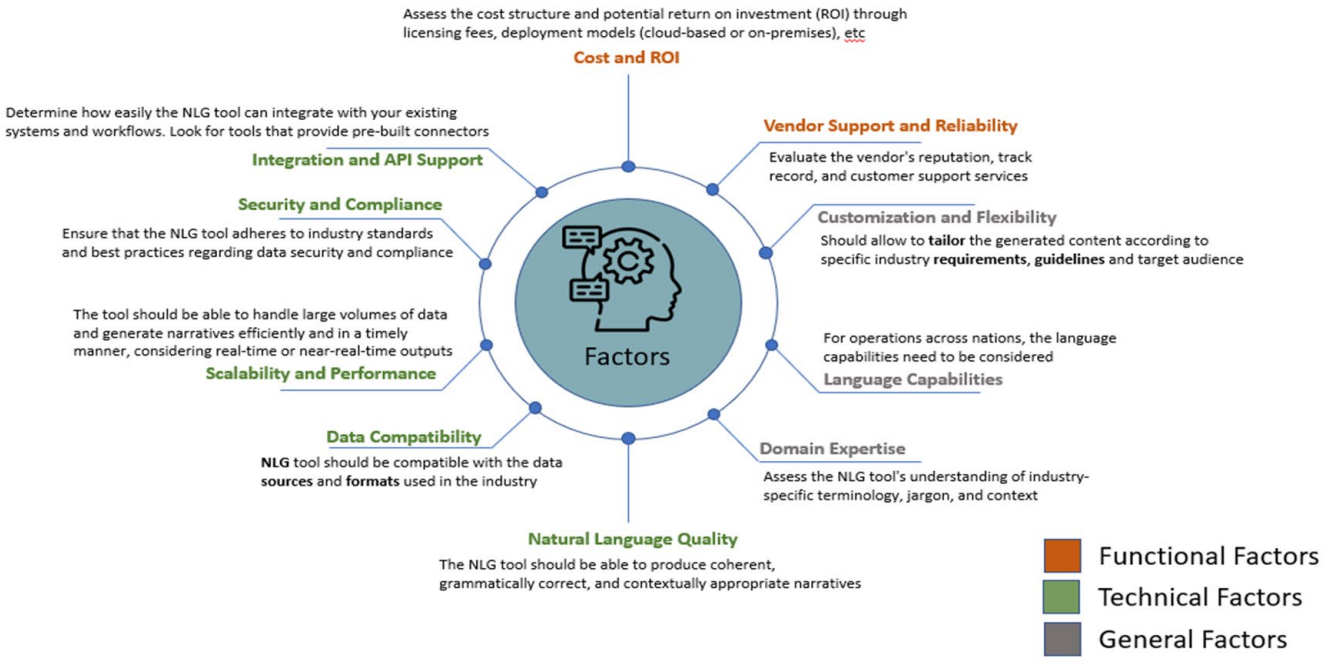
External Document © 2023 Infosys Limited

NLG Tools (Market & Open Source)

Tool	Type	License	Features	Strengths	Weaknesses
Text-to-Text Transfer Transformer (T5)	Free	Apache 2.0	Dialogue generation, text summarization, question answering	Can generate text in a variety of styles, including news articles, blog posts, and poems.	Can be slow to train and requires a lot of data.
Generative Pre-trained Transformer 3 (GPT-3)	Licensed	Closed source	Text summarization, question answering, code generation, creative writing	Can generate very realistic and creative text.	Very expensive to license.
Bard	Free	Apache 2.0	Text summarization, question answering, code generation, creative writing	Can generate text in a variety of styles, including news articles, blog posts, and poems.	Still under development and can sometimes generate inaccurate or misleading text.
RASA NLG	Licensed	Apache 2.0	Dialogue generation, text summarization, question answering	Can be used to generate text for a variety of purposes, including chatbots, customer support systems, and marketing materials.	Can be complex to set up and use.
Hugging Face Transformers	Free	Apache 2.0	Dialogue generation, text summarization, question answering	A library of pre-trained NLG models that can be used to generate text in a variety of styles.	Can be difficult to use for beginners.
Genie AI	Licensed	Closed source	Dialogue generation, text summarization, question answering	Can generate creative and informative text and is used by some major companies.	Can be expensive.

Parameters for Selecting Suitable NLG Tools

The key considerations for NLG tool integration in different areas of the Pharma industry involves several factors to consider:



External Document © 2023 Infosys Limited

NLG in LS Value Chain

Clinical Trial Data: Integration of NLG should consider the ability to examine and clarify complex clinical trial data, including efficacy and safety outcomes.

Drug Candidate Selection: NLG should assist in generating natural language reports to aid in the selection of promising drug candidates for further development.

Adverse Event Analysis: NLG can be used to automatically detect and summarize adverse events reported during clinical trials.

Regulatory Compliance: NLG-generated content must adhere to regulatory guidelines for drug labeling and medical documentation.

Drug Information Updates: NLG can assist in automatically generating updates for drug information, labels, and package inserts.

Medical Literature Summaries: NLG can summarize complex medical literature, making it easier for medical writers to extract essential information.

Adverse Event Reporting: NLG can be utilized to automatically generate adverse event reports for regulatory submission.

Signal Detection: NLG should assist in detecting signals of potential safety concerns from vast amounts of pharmacovigilance data.

Risk Assessment: NLG-generated reports can facilitate risk assessments of drug safety profiles.

Patient Query Resolution: NLG technology assists in patient query resolution by automatically generating accurate and relevant responses to patient inquiries or questions.

Mediating Patients and Providers: NLG plays a vital role in mediating between patients and healthcare providers by facilitating effective communication and understanding. It can assist in generating patient reports, summarizing medical records, and translating complex medical information into understandable language for patients.

24x7 Primary Support: NLG technology can support 24x7 primary support by automating the generation of responses to common queries or issues.

NLG adoption in LS Value Chain against tool parameters to indicate how much complexity (M-Moderate and S-Significant) is expected.

Parameters	Clinical Trial Data	Drug Candidate Selection	Adverse Event Analysis	Regulatory Compliance	Drug Information Updates	Medical Literature Summaries	Adverse Event Reporting	Signal Detection	Risk Assessment	Patient Query Resolution	Mediating Patients and Providers	24x7 Primary Support
Integration and API Support	M	M	M	M	M	M	M	M	M	M	M	M
Security and Compliance	M	M	M	M	M	M	M	M	M	M	M	M
Scalability and Performance	S	S	S	S	S	S	S	S	S	S	S	S
Data Compatibility	M	M	M	M	M	M	M	M	M	M	M	M
Natural Language Quality	S	S	S	S	S	S	S	S	S	S	S	S
Domain Expertise	M	M	M	M	M	M	M	M	M	M	M	M
Language Capabilities	S	S	S	S	S	S	S	S	S	S	S	S
Customization and Flexibility	M	M	M	M	M	M	M	M	M	M	M	M
Vendor Support and Reliability	M	M	M	M	M	M	M	M	M	M	M	M
Cost and ROI	M	M	M	M	M	M	M	M	M	M	M	M

Legend	Utility
M	Moderate
S	Significant

Note - This assessment is a general guideline, and the actual level of impact may vary depending on the specific requirements, industry regulations, and individual circumstances of each area in the Pharma industry.

External Document © 2023 Infosys Limited

NLG in PBI Dashboards

With NLG integrated into Power BI, it is possible to generate natural language explanations based on the data presented in the graph, where the use-cases (UCs) could be . - For a graph showing the trend in sales for a particular product, the NLG platform can be configured to generate a natural language explanation of the trend, such as "Sales for product X have increased steadily over the past six months, driven primarily by a surge in demand from new customers."

NLG integration in Power BI can improve the user experience by providing more context and interpretation of data visualizations. It helps in reduction of the time and effort required to analyze data, allowing users to focus on making informed decisions based on insights generated by NLG tools.

Choose an NLG tool: First, you need to choose an NLG tool that meets your requirements. Some popular NLG tools that can be integrated with Power BI include Arria NLG, Narrative Science, and Yseop.

Create an account: Once you have chosen an NLG tool, you need to create an account and obtain an API key.

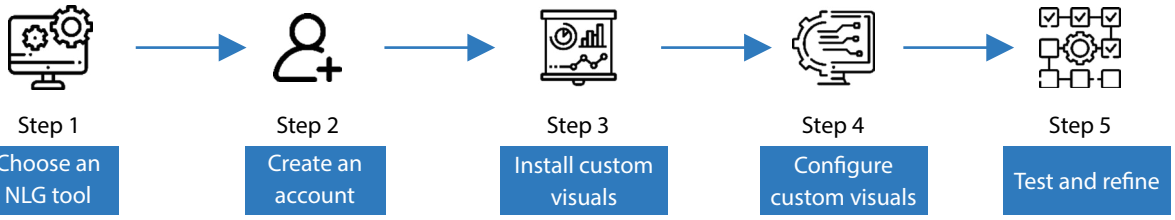
Install custom visual: To integrate NLG into your Power BI dashboard, you will need to use a custom visual that connects to the NLG tool's API. There are several custom visuals available that support NLG, such as Synoptic Panel, Smart Narrative, and Narrative BI.

Configure custom visual: After installing the custom visual, you will need to configure it to connect to the NLG tool's API. This typically involves entering your API key and configuring the visual to generate narratives based on the data displayed in the dashboard.

Test and refine: Once the custom visual is configured, you can test it by generating narratives based on the data displayed in the dashboard. You may need to refine the configuration and settings to ensure that the narratives are accurate and relevant to the data.

Our recommendation for enterprise class Generative AI applications

The organizations exploring generative AI either expect or are already experiencing tangible results. Microsoft Azure OpenAI platform or AWS Code Whisperer can generate code from English prompts boosting employees' productivity. Enterprises can use Infosys 4D framework to build privacy controls for adoption of generative AI.



Limitation of NLG tools integrating with PBI dashboard

There are limitations that may arise when integrating NLG tools with a Power BI dashboard. Some of these include:

Limited customizability: Some NLG tools may have limited options for customizing the generated narratives, which could limit the ability of users to tailor the reports to their specific needs

Data quality: The accuracy and relevance of the generated narratives can be impacted by the quality of the underlying data. If the data is incomplete, inaccurate, or outdated, the generated narratives may be misleading or irrelevant

Integration costs: Some NLG tools may require a paid subscription or license to integrate with Power BI. This could result in additional cost to the organization

Language adaptation: NLG tools may have limitations in adapting to different languages or dialects. The accuracy and fluency of the generated narratives could be impacted when dealing with languages other than the primary language supported by the tool.

Real-time updates: NLG tools may not provide real-time updates as the data in the Power BI dashboard changes. This can result in outdated or incorrect narratives being generated if the data is constantly evolving.

External Document © 2023 Infosys Limited

For example, if a Power BI dashboard generates a report that uses an NLG tool to generate insights on sales performance, the generated narrative may not accurately reflect the underlying data if the data is incomplete or inaccurate. This could result in the report leading to incorrect conclusions or actions being taken based on inaccurate information.

NLG Use-case in Pharma Industry

Sr. No	Topic	Application Area	Problem	Solution	Implementation Steps
1	Generating patient-friendly summaries of clinical trial data	Clinical trials	Clinical trial data can be complex and difficult for patients to understand	NLG can be used to automatically generate patient-friendly summaries of clinical trial data. These summaries can be written in plain language and tailored to the specific needs of each patient.	• NLG algorithms are competent on a corpus of clinical trial data. • The algorithms are then used to generate patient-friendly summaries of the data. • The summaries are reviewed by clinical experts to ensure that they are accurate and informative. •The summaries are then made available to patients
2	Creating educational content for healthcare professionals	Healthcare education	Complex medical concepts can be difficult for healthcare professionals to understand	NLG can be used to create educational content for healthcare professionals, such as drug information leaflets, patient education materials, and continuing medical education (CME) courses. This content can be written in plain language and designed to the particular needs of each healthcare professional.	• The training data for NLG algorithms is a corpus of regulatory papers. • Mock regulatory documents are then created using the algorithms. • To make sure that the drafted documents adhere to the most recent regulations, regulatory specialists review them. • The regulatory agencies are then given the finished paperwork
3	Automating the drafting of regulatory documents	Regulatory compliance	Regulatory documents, such as clinical study reports (CSRs) and marketing authorization applications (MAAs), can be time-consuming and expensive to draft	NLG can be used to automate the drafting of regulatory documents. This can save time and resources, and ensure that regulatory documents are compliant with the latest guidelines.	• A corpus of regulatory papers is used to train NLG algorithms. • The algorithms are subsequently applied to create draft regulatory papers. • To ensure that the drafted papers adhere to the most recent regulations, regulatory specialists review them. • Finally, the regulatory agencies receive the completed documentation for approval
4	Generating personalized marketing materials	Marketing	Marketing materials can be ineffective if they are not targeted to the precise needs of the target audience.	NLG can be used to generate personalized marketing materials for patients and healthcare professionals. This helps in improving the effectiveness of marketing campaigns, and to target messages to specific audiences.	• A corpus of regulatory papers is used to train NLG algorithms. • After that, the algorithms are utilized to develop drafted regulatory documents. • Regulatory specialists analyze the drafted documents to ensure that they comply with the most recent requirements. • Following that, the final paperwork is submitted to the regulatory authorities

External Document © 2023 Infosys Limited

5	Developing new drug discovery tools	Drug discovery	The drug discovery process is long and expensive.	NLG can be used to develop new drug discovery tools that can help to identify potential new drug targets and to design new drug molecules. This has the potential to accelerate the drug discovery process and to bring new treatments to patients sooner.	• NLG algorithms are trained on a corpus of gene expression data and chemical compound data. • The algorithms are then used to identify potential new drug targets. • The targets are then evaluated for their potential to be used in drug discovery. • The most promising targets are then used to design new drug molecules

About the Author

Anshuman Dubey

Senior Consultant – Infosys Consulting – Life Science | Data Transformation

Anshuman Dubey is an experienced Senior Business Consultant with over 16 years of expertise in data products, data mart , data governance, data modelling and security, data visualization, and data consulting. He has a successful track record of implementing critical projects in various commercial and operational areas of the life sciences and healthcare industries. He is well-versed in both agile methodologies and waterfall methods and can collaborate effectively with cross-functional teams.

Reviewer

Shanmugam Chinna Lakshmanan

Principal, ICLS – Infosys Consulting | Data Transformation

Shanmugam Lakshmanan (Shanny) is a seasoned professional having 18+ years of experience in data products, data mart , data visualization, data governance, data security, data compliance, data quality, data modelling, data analytics and data consulting. He is managing the Data and Analytics Practice in Infosys Consulting for healthcare and life sciences clients.

References

https://www.sciencedirect.com/science/article/pii/S1532046418302016

https://www.analyticssteps.com/blogs/overview-natural-language-generation-nlg

https://nexocode.com/blog/posts/natural-language-processing-healthcare/

https://emerj.com/ai-sector-overviews/natural-language-processing-in-pharma-current-applications/

https://www.linguamatics.com/solutions

https://blog.gramener.com/nlp-use-cases-pharma-healthcare/

https://bmcmedresmethodol.biomedcentral.com/articles/10.1186/s12874-021-01347-1

https://aaltodoc.aalto.fi/bitstream/handle/123456789/119968/isbn9789526411316.pdf

https://efaidnbmnnnibpcajpcglclefindmkaj/https://arxiv.org/pdf/2110.15803.pdf

For more information, contact askus@infosys.com

Infosys®
Navigate your next

© 2023 Infosys Limited, Bengaluru, India. All Rights Reserved. Infosys believes the information in this document is accurate as of its publication date; such information is subject to change without notice. Infosys acknowledges the proprietary rights of other companies to the trademarks, product names and such other intellectual property rights mentioned in this document. Except as expressly permitted, neither this documentation nor any part of it may be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical, printing, photocopying, recording or otherwise, without the prior permission of Infosys Limited and/ or any named intellectual property rights holders under this document.