

# *Why you should hire an automated analyst*

## Advanced Natural Language Generation in action

### *Executive summary*

Organizations on the journey to digital transformation are looking at opportunities to drive innovation that are practical and can be easily measured, however this often proves to be a complicated task. Advanced natural language generation (“Advanced NLG”), technology that transforms data into insightful language, is emerging as a practical application of Artificial Intelligence (“AI”), which can be used across the enterprise to drive operational efficiency, scale employee expertise, and accelerate decision-making. Akin to a top-tier analyst who interprets and communicates what is most interesting and important in the data, Advanced NLG is already being used to analyze and articulate data-driven information at incredible scale. This paper by PwC and Narrative Science will review how to appropriately evaluate an opportunity for Advanced NLG, with a practical framework for choosing an optimal starting point and aligning resources.

### *The reality of digital transformation*

#### **Despite significant investments, a gap in data & decision-making**

You are making the investment to become a data-driven enterprise, digitizing and centralizing data assets, hiring skilled analysts, and reporting information to relevant parties. In the current environment of "disrupt or be disrupted," you are most likely investigating opportunities to hire the best and brightest data-savvy and analytics-focused employees. Unwilling to be the next Fortune 500 enterprise to disappear you are exploring innovative technologies to reduce costs without necessarily reducing headcount. You are not alone in the pursuit of digital transformation.

In fact, 80% of executives agree that identifying opportunities to digitize their enterprise is a critical part of the innovation process<sup>1</sup> and they are willing to pay to make this happen. By 2019, companies around the world are expected to have spent a total of \$2.1 trillion on digital transformation.<sup>2</sup>

<sup>1</sup> 2017 PwC Global Digital IQ Survey

<sup>2</sup> IDC, IDC FutureScape: Worldwide IT Industry 2017 Predictions, IDC #US41883016, November 2016

So then why, despite a significant investment in time, talent, and resources, are your employees still spending countless hours manually analyzing, interpreting, and communicating data insights? After all of this effort in data aggregation and reporting, why are we still left with tables and charts that leave us questioning, "What does this mean for me? For my business unit? For my customers?" And why do these critical insights still take so long to produce?

Forward-looking enterprises are realizing the answers to these questions are not simply found by hiring the elusive data scientist or making further investments in data visualization tools alone. To fill the gap between data and decision-making, they are looking at opportunities within AI and automation, with over half of Global CEOs exploring the benefits of humans and machines working together.<sup>3</sup>

### Promises and perils of emerging technology

The challenge? Where to start. The promises of AI have historically been accompanied by hype that can prove to be costly and confusing, wrought with black-box applications ("How exactly did that model arrive at that decision?"), complicated deployment cycles ("Why do I need to drastically change or replace my existing infrastructure again?") and unclear ROI ("I'm going to need to justify this investment. When will that happen?"). Instead of starting with the most pressing needs of the business, many executives are lured by the "cool factor" of a given technology, failing to abide by a golden rule of digital transformation: Don't invest in emerging technology for the sake of innovation, invest to solve real problems.

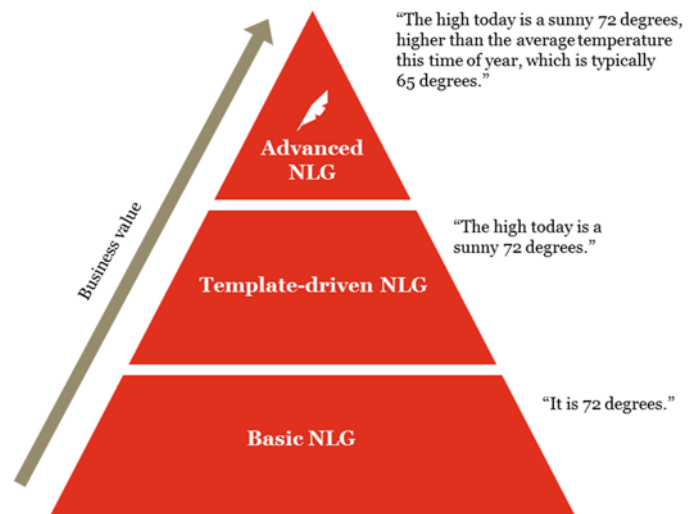
Advanced NLG is a practical application of AI that transforms data into intelligent, human-sounding language, and can be used across an enterprise to take advantage of its own data by turning it into a valuable asset. The difference between a human analyst and the Advanced NLG Automated Analyst? Advanced NLG performs these tasks automatically, consistently, and at scale - - enabling key employees' time to be freed up and spent on higher-value tasks or analyses.

## Meet your new automated analyst

### Natural Language Generation: Not all the same

Natural Language Generation ("NLG"), a subfield of AI which produces language as output on the basis of structured data input, is not a new concept. What is new, however, is the recent adoption of NLG into more enterprises. There are a plethora of ways the technology is being employed, primarily to improve human productivity, increase operational efficiency, focus analysts, and achieve regulatory compliance.

To understand where NLG can be applied within your organization, it's important to delineate the different types of this technology. On one side of the spectrum is basic NLG, or essentially mail merge, the act of restating numbers into language via light rules or functions. In the middle of the spectrum, is templated NLG, or the process of building sentences backed by the use of business rules, basic calculations, and templates with pre-built text. Templated NLG can be equated to a very junior analyst, who can send the same analysis snippet every week to accompany a spreadsheet or a dashboard, however is limited in his or her ability to draw from multiple data sources, perform custom analysis, and communicate in a personalized way, as he or she has no real understanding of the business domain.



Natural Language Generation is a subfield of Artificial Intelligence that produces language from data

<sup>3</sup> PwC, "Five Forces that will reshape the global landscape of anti-bribery and anti-corruption"

## Advanced NLG in action

### Global financial institution begins compliance reporting transformation

PwC worked with a major, global financial institution to examine potential ways to reduce the time it takes to prepare certain regulatory reports.

In one use-case, relating to the narrative generation associated with certain Know Your Customer (“KYC”) efforts around the identification of potentially high-risk customers, PwC worked in conjunction with the financial institution to reverse-engineer its current reporting process back to the individual data elements. While the narrative leveraged inputs from a number of disparate data sources, PwC employed its domain knowledge of both KYC requirements and the institution’s own systems to concatenate the information. PwC then configured and automated the report using Narrative Science’s Quill (Advanced NLG) engine.

The financial institution generates tens-of-thousands of these reports every year. The application of PwC’s domain knowledge and experience with Advanced NLG facilitated an estimated 25% reduction in the time it takes to prepare a final narrative, resulting in savings in excess of \$1 million per year. Beyond the financial savings however, key stakeholders cited the consistency with which the reports could be generated, as well as the speed to insight, as key benefits.

At the more sophisticated end of the spectrum is Advanced NLG. Instead of beginning with a template, Advanced NLG begins with the user's intent, or in other words, an understanding of what needs to be communicated. If the goal of a report is to identify drivers of sales over the last quarter, for example, the system would need to know:

1. **The audience:** If the VP of sales is receiving the narrative report, the system would generate an overall synopsis of key takeaways. If it was a regional manager, the output would be a deep-dive into that particular region.
2. **The analysis:** Given the goal of the report, the system would perform a driver analysis to identify the contributors or detractors of overall performance, a time-series analysis to measure different periodic dimensions in your key metrics, and would rank cohorts against each other, as well as against standard benchmarks.
3. **The data:** More often than not, data to fulfill the types of analyses above won't be in one centralized location. The system would just have to be pointed to each data source so that it can extract the relevant information needed to power the narrative.

Narrative Science's Quill engine is an example of an Advanced NLG system. Similar to a top performing analyst, Quill performs complex assessments to characterize events and identify relationships, understands what information is especially relevant, learns about certain domains, utilizes specific analytics and language patterns accordingly, and generates language with the consideration of appropriate sentence length, structure and word variability. The result is an intelligent narrative that can be produced at scale or customized to an audience of one.

### Identifying relevant use cases

Anywhere in the enterprise where there is a need to communicate insight from structured data, there is an opportunity to leverage Advanced NLG. Typically, these opportunities fall into two categories: 1) Replicating an existing narrative report, or 2) Developing data-driven content where no content existed previously.

If exploring Advanced NLG for the first time, it's usually best to begin by replicating, or reverse-engineering, an existing report, as you already have an understanding of the audience's needs, analysis required, and data available. It also gives you the ability to measure the time and resources required prior to automation, so that you can more easily measure ROI. Consider the metrics needed to determine the impact of the NLG implementation and outline these ahead of time. For your first implementation, it is best to select a report where the use of NLG can show a clear time or financial savings.

To get started with Advanced NLG for the purpose of report replication, start by asking these questions:

**1. Does my team spend a significant amount of time creating a standardized report?**

*Measure the time required to aggregate the data, analyze the insights, interpret the results, and communicate to your audiences.*

**2. Is this report steeped in structured data?**

*NLG is intended to be used with structured data, or in other words, data within databases, spreadsheets, .csv files, or machine-readable formats like JSON.*

**3. Does this report need to be generated at a high frequency and/or volume?**

*If the report is only written once a year (frequency) or for a handful of people (volume), it may or may not be ripe for automation. NLG often shines when scale is required, either in the number of reports required or the number of times they need to be generated as the underlying data changes.*

**4. Are similar reports being drafted by individuals with differing levels of expertise in a multitude of locations or geographies?**

*Unlike a team of dispersed analysts with varying skillsets, who each have different writing styles and approaches to performing analyses, Advanced NLG enables standardization and consistency that is not possible with humans alone.*

If you've answered yes to more than one of these questions, you are ready to get started with Advanced NLG.



## **Ensuring there is ROI for your NLG project**

Now that you have an understanding of the considerations for Advanced NLG, it is critical to think of the pressing business issues you are facing across your enterprise to ensure you are attributing business value to your NLG project.

### **Operational efficiency, achieved**

Data-rich functions like a company's Financial Planning & Analysis ("FP&A") department are prime locations for Advanced NLG implementations. Manual and time-consuming reports such as P&L analyses can take highly skilled FP&A analysts away from participating in more strategic activities. In fact, only 24% of time spent on finance related tasks are insight-generating activities.<sup>4</sup> Per a PwC Finance Trends report, top performers operate at lower cost not by reducing service levels but by standardizing and simplifying their core processes and systems -- enabling them to free up resources to focus on business partnering.<sup>5</sup> By automating data-driven reports in a way that is relevant to each intended audience, finance employees can work on improving the bottom line, not purely reporting on it; and they can complete these critical reports in minutes or hours instead of weeks or months, enabling them to act on insights earlier.

### **A faster, more focused starting point**

Compliance and Internal Audit departments are other areas ripe for automation. Fifty-three percent of respondents to PwC's 2016 Financial Services Global Economic Crime Survey reported that spending on fighting economic crime was increasing. However, spending more has not meant less economic crime.<sup>6</sup> Instead of devoting more and more resources at increasing headcount and the need for training, compliance teams can automate portions of the analysis and reporting of both internal due diligence and regulatory reports. Through automatically surfacing the most relevant insights in natural language, investigative teams can better hone their focus areas with an accelerated starting point so they can then dig deeper into the root causes of suspicious

activity. In a pilot study PwC utilized Advanced NLG to generate the quantitative components of regulatory reports, such as Suspicious Activity Reports ("SAR Narratives"). By automating approximately 80% of the SAR Narrative, analysts can spend more time adding qualitative insights for a more thorough report or digging into particular observations.

### **Standardized, consistent communication**

Business intelligence ("BI") & analytics teams have a solid understanding of existing data and relevant analysis to perform, but they often face one consistent problem: explaining the insights they've gained to their business-oriented counterparts across the organization. Different audiences require different degrees of explanation, and even when looking at the same chart or graph, various audiences can interpret them differently. Advanced NLG can be integrated directly within BI and analytics tools, with narratives that dynamically explain the relevant insights within data and visualizations. Instead of manually annotating each chart and graph with a summary, teams can save time and provide standardized and consistent communication through automatically generated annotations. Per Gartner, by 2019, natural-language generation will be a standard feature of 90% of modern BI and analytics platforms.<sup>7</sup>



<sup>4</sup> PwC Finance Effectiveness Benchmark Report 2017

<sup>5</sup> PwC Finance Effectiveness Benchmark Report 2017

<sup>6</sup> PwC Global Economic Crime Survey 2016

<sup>7</sup> Gartner Predicts 2017

## Advanced NLG in action

### PwC Transforms Anti-Bribery, Anti-Corruption (ABAC) with data-driven narratives

ABAC is a compliance area that continues to receive significant attention, particularly in heavily regulated industries, but often struggles with resourcing and funding. While performing periodic health checks or during the course of an ABAC-related investigation, internal audit, compliance, and legal professionals often perform analyses on financial data underpinned by risk models to identify potential instances of impropriety or to monitor transactional data. Historically, documenting and reporting the results of these analyses requires significant time, resulting in an expensive process that may lack consistency and quality due to a number of factors, including differences in writing styles, levels of experience, variance in analysis approach, and time constraints.

To address these issues, the PwC Forensic Services team, in collaboration with Narrative Science, used Quill to develop a narrative based on risk models, indicators, and underlying data. The automated narrative allows compliance, audit, and forensic professionals to spend less time drafting, or developing ad-hoc data queries, and more time interpreting data-driven insights and taking actions on those insights. Forensic accountants estimate that they reduced the time to insight by as much as 50% in some cases. Reports that typically require hours of staff time are now generated in minutes in an automated, consistent approach that sharpens the investigator's focus as to where to hone in on potential corruption or fraud risk areas.

## Next steps

### Establishing a practical framework

As next steps, it is critical to align on the goals for the project, the skillsets required to be successful, and a timeline to work towards. Here is a practical framework to get started on your Advanced NLG project:

- **Destination:** What is the goal of both the NLG project overall (e.g., improve operational efficiency, achieve consistency, draft new content) and the initial narrative report type (e.g., anti-corruption health check, tangible asset valuation report, electronic discovery custodian report, journal entry summary)? Identify project measurement metrics; for example, how much time is currently spent creating this report? By outlining the goal of the initial report type, you will better understand the resources needed on your end to ensure you have access to the relevant domain specialists, data, and distribution components to ensure the solution is successful.
- **Domain expertise:** Implementing an Advanced NLG solution is similar to training a new, top-tier analyst. As such, you need subject matter expertise in the business area of your narrative report. Automating a SAR Narrative? You should have a person well-versed in Anti-Money Laundering regulations and FinCEN requirements. Producing a Finance P&L report? Ensure the FP&A professional has relayed the necessary reporting time periods.
- **Data:** Access to clean, complete, and structured data is important for any NLG implementation. You will need a person on your team that is knowledgeable about where the data resides and how best to access it.
- **Distribution:** Narratives are only transformative if they are actually read, so the output of Advanced NLG systems need to be seamlessly integrated into your existing workflow to ensure maximum adoption. Fifty-nine percent of executives believe that lack of integration of new and existing technologies and data pose a barrier to achieving expected results from digital technology initiatives.<sup>8</sup> Ensure your digital transformation efforts are successful by choosing technology that is easily extensible.

---

<sup>8</sup> 2017 PwC Global Digital IQ Survey

## ***Contact information***

**Daniel S. Krittman**

PwC

Principal

Direct: +1 (646) 818 7062

Email: [daniel.krittman@pwc.com](mailto:daniel.krittman@pwc.com)

**Mary Grace Glascott**

Narrative Science

Director of Product Marketing

Direct: +1 (312) 548 9781

Email: [mglascott@narrativescience.com](mailto:mglascott@narrativescience.com)



© 2017 PwC. All rights reserved. PwC refers to the US member firm or one of its subsidiaries or affiliates, and may sometimes refer to the PwC network. Each member firm is a separate legal entity. Please see [www.pwc.com/structure](http://www.pwc.com/structure) for further details.

This content is for general information purposes only, and should not be used as a substitute for consultation with professional advisors.