

Natural Language Generation



Table of Contents

Why NLG?

p.2

NLG Going Forward

p.4

Delivering Financial News, Not Numbers

p.5

Why NLG?

The volume of data available to anyone with an internet connection is enormous. But without interpretation and actionable insight, this data can be confusing, overwhelming, and obscure. This is where Agrud Technologies, and NATURAL LANGUAGE GENERATION (NLG) come in.

Agrud seeks to harness the incredible potential of Natural Language Generation, a subfield of artificial intelligence, to produce content and analytics on an unprecedented scale. As well as providing broad market analysis for a mass audience, Agrud is striving to create content tailored for individuals. The possibilities of using NLG to tailor information and news to small audiences have been explored in other sectors. For example, Urbs Media in collaboration with Press Association developed a system called RADAR to provide an incredibly diverse range of local news stories derived from government and local authority data sets. In the financial world, however, the potential of NLG has not been fully realised.

Traditionally financial news organisations, such as YAHOO finance and Reuters, have relied on a combination of structured data, and content written and curated by human professionals and experts. But due to the enormous volume of available data, and the limited scope of human reporting, this leaves a great deal of data unhandled, unanalysed, and unseen. Structured Data, such as graphs and databases, offer dense information, but with no analysis or comprehension, so those without sufficient expertise are left in the dark. With NLG this limitation does not exist.

The automated handling of data in NLG is not only incredibly valuable to the consumer, but also to writers and content creators. Traditional media companies face increasingly difficult challenges posed by the digitization of the publishing industry. The availability of data has increased to a volume which has become prohibitive to its handling and analysis, especially in an environment which pressures the reduction of resources in content creation. Natural Language Generation can do the heavy lifting in this case, extracting information from enormous databases in a fraction of a second based on criteria which flag it as important or unusual. The point of NLG is not necessarily to automate 100% of

everything, but to re-configure the workflow of writers, and automate the most tedious and costly data-driven reports. In other words, to do what computers are good at.

Using NLG to transform raw data into written reports is not a new concept, but with the rise of applied NLG taking the business world by storm, it is becoming a necessity. Forbes declared NLG to be one of 2017's hottest trends, while Kristian Hammond, Narrative Science's co-founder, estimates that 90% of news could be algorithmically generated by the mid-2020s, much of it without human intervention.

A major strength of NLG over structured data is the power of written language. While charts and graphs can communicate a few variables at a time, a single word can convey a host of factors. For example, describing a company's quarter as "strong" can account for its stock price, profitability, its capital investments, its acquisitions, its talent and many other attributes. However, creating a platform which can emulate the depth, personality and variability of a human writer can be a tough obstacle to overcome.

NLG Going Forward

NLG can be as simple or as complex as its developer requires it to be. Generally speaking though, with more complexity comes more analysis and interpretation. Useful natural language generation software should not simply plug nouns and numbers into canned phrases or perform simple if-then conditions to change individual words. Agrud's NLG development is built around a layered structure in which depth can be added throughout, and extended after development. By using a series of interchangeable words, within interchangeable phrases, within interchangeable sentences, a very high degree of variability, expression, and analysis can be achieved. There is, of course, a limit to this method of development as every layer of depth increases development time and resources exponentially. As such, new NLG methods are being developed which reduce greatly the need for human annotation and in creating the NLG corpus.

End to End Natural Language Generation uses machine learning and neural networking to create NLG outputs with almost no human interference in the process. This is achieved through a process known as 'training'. This approach has had some degree of success, although only in projects of fairly limited scope. Creating more complex automated NLG systems using neural networking and machine learning can be problematic. One only needs to look at the latest generation of neural network, and deep learning based chatbots to see the problems which can occur with incorrect word associations, repetitive loops, and syntactical errors. Looking into the future, Agrud hopes to combine traditional NLG methodology, with its carefully constructed templates and predictable outputs, and Neural Network based NLG to create a platform which is as flexible as it is robust.

Delivering Financial News not Numbers

Agrud has produced a range NLG algorithms designed to turn the raw data you're used to into readable, engaging updates and articles, deliverable in a variety of ways depending on end user needs, including embedding into the client's API.

An End of Market Update provides a short daily article based on the major index movements of that day. The algorithm first analyses overall market performance, summarising it in a single sentence, and goes on to provide point/percentage index changes, volume deviation and top gainers/losers.

As well as daily updates, there are other algorithms which trigger when certain conditions are met. For example, when a shareholder of a company increases their holding significantly, a brief report is generated providing information on the transaction, the company, and on the individual where possible. So too when a the volume of sales in a company are driven to unusual highs or lows.

Lastly there are intraday updates available which can be delivered as frequently or infrequently as required. These give hour to hour updates, detailing the major market movements that you need to know about.

Agrud's ventures into Natural Language Generation have just started, and are already producing results. The proven benefits of automated content present a bright future for NLG in business and finance.

Intended Use Cases

The potential uses for Natural Language Generation are as broad as human speech itself. Initially, however, Agrud are targeting the financial sector, and news media.

For the financial sector Agrud offers automated data analysis, in written English.