



# New gold: Opinion mining in Africa

## INTRODUCTION

---

Mining and Africa have a synonymous but complex relationship. The continent's wealth of mineral resources (US Geological Survey, 2012) has motivated nations to go to war and people seeking employment to migrate in masses. While delivering wages, infrastructure and urban development, mining has simultaneously led to the fragmentation of families, the exploitation of workers, the infamy of racial stratification and revolt.

In the digital age, a new version of the mining industry has been born. Data has now become the raw material delivering wealth, employment and opportunity (The Economist, 2017). The processing of minerals is being replaced by the processing of information. Billion-dollar valuations of companies mastering this new art have led to a new gold rush which, similar to other times in history, hit the west coast of the United States and then spread across the world.

Africa has enjoyed a share of the spoils. Over the past five years an increasing number of African information entrepreneurs (Imaralu, 2013) has built innovative internet-driven solutions which demand new skills and create new jobs inspired by the success of their peers through lucrative foreign acquisitions (Thomas, 2014; Thomas, 2015; Timm, 2017).

---

### Sentiment analysis

One example of these solutions that African innovators (Ventureburn, 2017) have adopted is "opinion mining" (BrandsEye, 2017) or "sentiment analysis" (Wikipedia, 2017a):

*Opinion mining (also known as sentiment analysis) refers to the use of natural language processing, text analysis and computational linguistics to identify and extract subjective information ... Generally speaking, opinion data analysis aims to determine the attitude of a speaker or a writer on some topic ...*

The discipline was initiated in the 1960s when academics began to explore the potential of using the qualitative research method of content analysis to infer human behaviour from a conversation by making use of computers (Gottschalk & Gleser, 1969; Stone, Dunphy & Smith, 1966). Fifty years later, advancements in computing power, the magnitude of online conversations available through social media and the internet as well as the ongoing development of the academic enquiry have given birth to a new industry of software and service providers (Pang & Lee, 2008). Sentiment analysis tools and technology are founded on natural language processors and machine learning. Search engines gather data about specific words or themes. Algorithms scan, filter and "learn" the use of words and phrases in online conversations and other forms of relevant digital content to determine whether the "sentiment" of those conversations or content is neutral, positive or negative. This is important because what individuals say about certain topics is an indicator of how they might act towards it – the foundation of qualitative research. More advanced tools offer deeper insights into the themes and topics connected to that sentiment and the demographic and geographic details of those expressing such opinions.

Sentiment analysis solutions that rely on machines alone to process conversation struggle to deliver accurate results. This is due to the way human beings use language. We frequently use sarcasm, wit and idioms to express our opinions. We mix the languages we use. We use words and phrases in ways that make it difficult for machines to pick up the meaning and context of what we are saying. Accuracy is arguably the most important measure of quality when wanting to understand whether a citizen, voter, customer or competitor is positively orientated to an organisation, or not, and what the reasons for that orientation might be. To improve accuracy,



## NEW WORLD OF BUSINESS

By Colin Habberton

Founder: Relativ Group

the best ingredient is to add humans to the process of analysing sentiment. The combined competencies of crowds of humans (understanding the quality of what lies between the lines of conversations) and machines (processing the quantity of conversation at scale) creates a potent blend of accuracy and insight.

Considering its past, a number of factors place Africa in a unique position to take advantage of the information economy. Opinion mining and its applications have the potential to become an industry equal to its traditional namesake regarding employment and income opportunity. In welcome contrast, it holds promise in creating a workforce that, unlike the mining of the past, finds value in the location, culture and context of its workers. The discovery of this new gold is made possible thanks to an unlikely inheritance.

### A dark history with a silver lining

The descent of colonial powers brought a wave of immigration from Europe over the past 500 years. The shipping routes around the Cape to the Far East contributed to the expansion of European settlement across Africa initiated by the Portuguese “discoverers”, followed by the exploits of the Dutch East India Company and its eventual demise at the hands of the British. The discovery of diamonds, gold and other minerals from the late 1800s led to the establishment of companies that would define the nature of the African resource-based economy for over a century. Mining and the expansion of commercial agriculture created a significant demand for labourers, and the indigenous and immigrant populations were drawn to seeking wages by working for companies owned by European settlers.

In many parts of Africa and the rest of the world, mines sparked the establishment of towns and cities aided by the influx of immigrants from different countries and cultures. With the variety of people from different places came a range of languages. Today, African cities are becoming increasingly connected to the internet through the proliferation of affordable devices and data plans, some sponsored by the cities as a service to their citizens (Project Isizwe, 2017).

### Opinion mining in Africa

The combination of digital access, web literacy, language proficiency and the relatively lower cost of living in Africa creates a wealth of skilled workers to process opinion data. The most widely spoken languages on the continent are English, Arabic, French and Portuguese in addition to Spanish, German and 1 500 endemic African languages including Swahili, Zulu and Xhosa (Matshego, 2016). This range of languages, in addition to those languages spoken by immigrant communities from China and India in Africa, provides opinion mining companies with access to affordable, skilled workers that cover languages spoken by major sections of the world’s population. The internet makes it possible for opinion miners to work wherever they can connect without the cost and inconvenience of commuting to an office for set hours. Crucially, the earning expectations of African opinion miners are less than those based in the global north, presenting an opportunity for ongoing job creation.

---

The combination of digital access, web literacy, language proficiency and the relatively lower cost of living in Africa creates a wealth of skilled workers to process opinion data.

---

### New gold

We find ourselves immersed in the information age. But what lies ahead? The Terminator movie series released over 30 years ago (Wikipedia, 2017b) tells the story of a man and a robot that return from 2029 to destroy humanity’s ability to stand in the way of a future ruled by hostile machines by taking control of world’s computer systems. Tracing the development of our current equipment, we increasingly rely on machines to wake us up the morning, provide our electricity, prepare our food, assist us with our work, and monitor our fitness and our rest. Driverless cars, drone delivery and intelligent understanding of our preferences define how we move, and what we receive, read and consume. Taking one step back, we are moving rapidly into an age driven by machines, with less fiction than the science that created them.



## NEW WORLD OF BUSINESS

By Colin Habberton

Founder: Relativ Group

The implications for humanity are immense. We need to recognise that machines – like natural language processors – have, and most likely always will have, limitations in interpreting and understanding human action. The educators of today need to start preparing workers in terms of the role that machines will play in their future. It is not purely the use of the machines that needs to be taught. More importantly, we need to understand how they work and what their purpose is. A number of tools (Kolodny, 2016) and training courses (Code.org, 2017) are already available to begin this process, even from the age of three. Learning to code is a start.

Learning a new language will enable us to understand the people from a different country or culture. Likewise, humanity as a whole will need to learn the language of machines to avoid becoming marginalised by those same machines that we depend on every day.

To tame the beasts we have built, we need to understand the way in which these machines work and to confidently know what to do when they do not work. Like opinion mining, the true value delivered by machines is dependent on the crowd of human beings that have harnessed these beasts and refined their view of the world in which we live to give us the information we need – the new gold – to guide us to the destinations we desire.

### REFERENCES

- BrandsEye. 2017. *We are the world's leading opinion mining company*. Retrieved from <https://www.brandseye.com/>.
- Code.org. 2017. *Every student in every school should have the opportunity to learn computer science*. Retrieved from <https://code.org/>.
- Gottschalk, L.A. & Gleser, G.C. 1969. *The measurement of psychological states through the content analysis of verbal behavior*. University of California Press.
- Imaralu, D. 2013. *10 Africans who founded million-dollar internet companies*. Ventures. Retrieved from <http://venturesafrica.com/10-africans-who-founded-million-dollar-internet-companies/>.
- Kolodny, L. 2016. *Primo Toys rolls out Cubetto, a wooden robot that teaches kids to code*. Tech Crunch. Retrieved from <https://techcrunch.com/2016/10/31/primo-toys-rolls-out-cubetto-a-wooden-robot-that-teaches-kids-to-code/>.
- Matshogo, L. 2016. *How many languages of Africa are there?* Africa.com. Retrieved from <https://www.africa.com/many-african-languages>.
- Pang, B. & Lee, L. 2008. Opinion mining and sentiment analysis. *Foundations and Trends in Information Retrieval*, 2(1-2), 1-135.
- Project Isizwe. 2017. *Free WiFi for South Africa*. Retrieved from <http://www.projectisizwe.org/>.
- Stone, P.J., Dunphy, D.C. & Smith, M.S. 1966. *The general inquirer: A computer approach to content analysis*. Cambridge, Massachusetts: MIT Press.
- The Economist. 2017. *Data is Giving Rise to a New Economy: How is it shaping up?* Retrieved from <https://www.economist.com/news/briefing/21721634-how-it-shaping-up-data-giving-rise-new-economy>.
- Thomas, S. 2014. *2014: The year in digital agency acquisitions*. Ventureburn. Retrieved from <http://ventureburn.com/2014/12/2014-year-digital-agency-acquisition/>.
- Thomas, S. 2015. *Why Automattic's WooThemes acquisition is so important to South Africa's startup space*. Ventureburn. Retrieved from <http://ventureburn.com/2015/05/automattic-woothemes-acquisition/>.
- Timm, 2017. *Chance meeting helped SA startup GetSmarter secure \$100m sale*. Ventureburn. Retrieved from <http://ventureburn.com/2017/05/chance-meeting-helped-sa-startup-getsarter-secure-100m-sale/>.
- US Geological Survey (USGS). 2012. *2012 Minerals Yearbook*. Retrieved from <https://minerals.usgs.gov/minerals/pubs/country/2012/myb3-sum-2012-africa.pdf>.
- Ventureburn. 2017. *Data mining firm BrandsEye crowned winner of 2017 FNB innovation awards*. Retrieved from <https://ventureburn.com/2017/06/data-mining-firm-brandseye-crowned-winner-2017-fnb-innovation-awards/>.
- Wikipedia. 2017a. *Sentiment analysis*. Retrieved from [https://en.wikipedia.org/wiki/Sentiment\\_analysis](https://en.wikipedia.org/wiki/Sentiment_analysis).
- Wikipedia. 2017b. *The Terminator*. Retrieved from [https://en.wikipedia.org/wiki/The\\_Terminator](https://en.wikipedia.org/wiki/The_Terminator).