Introduction

It was a Monday morning in early January 2020, and most of the 25-or-so US-based staff members of data science and artificial intelligence (AI) company, Rho AI, were online for the company’s daily stand-up. Issues that needed deeper thought were flagged for smaller discussions at a later stage. Gilman Callsen, the company’s Chief Technology Officer, raised the topic of a brand new project to develop an open-access and open-source, environmental, social and governance (ESG) investment tool. ‘In many ways this project puts us into new territory, so we need to discuss the ethics involved in this project and the potential business case, data quality and impact’, he said. Callsen and David McColl, Rho AI’s Head of Commercialisation, agreed to speak later that day to consider the issue.

Background on Rho AI

With the slogan, ‘Let’s uncover the value of your data’, Rho AI had evolved from Pit Rho, a company that Callsen and four others had founded in 2012, when they developed an AI product that provided statistical insights into races held by the National Association for Stock Car Auto Racing (NASCAR) (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December; Weaver, 2018). Having
made progress with Pit Rho, the founders then set out to build a bigger company that used data science and machine learning to develop products and services that would enable companies to obtain, organise and utilise data to their benefit. ‘We definitely do try to steer the ship towards impact projects whenever possible – it’s deeply rooted in our culture – we are always looking for meaningful projects that have something tangible as an end product. We want to be doing data science that impacts the world. We don’t want to be doing something that analyses tweets, or anything like that’, said Callsen (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December).

Using the Gartner Hype Cycle for Artificial Intelligence, 2019, McColl described the nature of the industry in which Rho AI operated. ‘In the beginning there were a few behemoths that were producing black box AI models that were supposed to be generalized enough to work in any application’, he said.

Things like IBM Watson were supposed to replace everything humans did (Peak of Inflated Expectations). But that approach didn’t prove to be as fruitful as anticipated because AI hadn’t reached a state where it was truly general (Trough of Disillusionment). Then people started to understand the need to develop models in tandem with industry knowledge and make the process more transparent, so many companies started forming around specific verticals and areas of expertise (Slope of Enlightenment). And now there are many companies carving out niches, either making AI platforms for one specific process/category or deeply integrating with the expertise of the partners they work with (Plateau of Productivity). To thrive in this space you need to have a good understanding of both the mathematics of the models and how they apply to the specific use cases. (Fischer and Beswick, 2019, Online interview with David McColl, 21 November; Gartner, 2019)

A desire to foster ESG investment

The ESG investment tool that Rho AI was developing was designed to make it easier for small-scale investors in particular to understand and integrate information on companies’ climate impact into their decisions about where to invest. The appetite for ESG investment was at the start of 2020 increasing globally. Sustainable investing assets in the major developed economies amounted to approximately US$30.7 trillion in 2018, a 34% increase on 2016 figures and 68% increase on 2004 figures (Global Sustainable Development Alliance, 2019; McKinsey, 2019). In the United States, ESG assets grew from US$8.7 trillion at the start of 2016 to US$12.0 trillion at the start of 2018, an increase of 38%. ESG assets constituted 26% of all investment assets under professional management in that country (Global Sustainable Development Alliance, 2019).

At the beginning of 2020, following a period of 20 years, climate change had become an ever more important component of ESG investment (Common Fund Institute, 2013), with a need for such investment coming into even sharper focus as the effects of global warming were becoming more apparent.4 With the signing of the United Nations (UN) Paris Agreement in 2016, the countries which were signatories to this agreement had undertaken to put in place measures to keep the global rise in temperature to well below 2°C above pre-industrial levels and to attempt to keep the increase to 1.5°C by the end of the century (UN, 2016). The Intergovernmental Panel on Climate Change (IPCC) estimated that to achieve this goal, there would have to be an annual average investment in the energy system of around US$2.4 trillion between 2016 and 2035, representing about 2.5% of the world GDP (IPCC, 2018).

Air pollution, that is, carbon dioxide and methane, released, for example, by burning fossil fuels, contributes to and exacerbates climate change. In addition to this, once the coronavirus pandemic developed in spring 2020, initial findings examining the link between COVID-19 mortality rates and air pollution showed a correlation, though not necessarily a causation, with a relative higher mortality rate in areas with higher air pollution and higher population density compared on a regional basis5 (Gross and Hook, 2020).

While at the time of writing this case study, the investment performance of ESG-centred portfolios through the COVID-19 crisis was still inconclusive, and there was growing evidence that companies that focussed on ESG issues within their organisation performed better over the long term (DB Climate Change Advisors, 2012; McKinsey, 2019). Thus, investment in such companies would also maximise monetary returns for their investors. Despite this, adoption of ESG metrics in investment decision-making had not reached the levels that the IPCC had indicated were necessary to keep global warming at the desired levels. The major reasons for this related to information. According to the Department for International Development in the United Kingdom, two of the three people in the United Kingdom agreed with the statement, ‘I have a responsibility to make the world better, and I want my investment choices to make a difference’. However, 54% of people cited a lack of information as a barrier to sustainable investment (HM Government, 2019).

Research by the Stanford Global Projects Center identified three barriers to integration of ESG data into investment decision-making. These were the following:

- Unlike traditional financial data that were structured and quantitative, ESG data tended to be ‘unstructured, qualitative, scattered, and incomplete’ (In et al., 2019, p.5);
- Evaluation of ESG data required more robust tools to ensure that they were used appropriately;
• Investors were not using new tools of data collection and analytics effectively (In et al., 2019).

Brad Foster and David Tabit, respectively, global head of enterprise data content and global head of equity data at Bloomberg, remarked that lack of access to consistent and reliable ESG data was forcing investors to ‘expend an excessive amount of already limited resources trying to standardize and interpret unstructured data, slowing down experienced investors and inhibiting new entrants from joining the field’ (Foster and Tabit, 2019).

It was this lack of information that Rho AI intended to address. The company planned to use automated web scraping technology, natural language processing (NLP) and machine learning to collect and to analyse climate-related data on US-listed companies and to build an open-access software platform that provided transparency on how companies were contributing to, or were tackling, environmental problems. Access to this information would be provided on an intuitive, user-friendly interface. Initially, the project would focus on the areas of pollution monitoring, air quality and efficient use of energy, including companies’ transportation systems and supply chains. The areas covered could expand to include other ESG factors in the future (McColl, 2019, Email correspondence with Isabel Fischer, 17 September).

Implementation of the project

Rho AI’s conventional development methodology was to put in place a road map for each project with dedicated developers assigned to the project and responsible for hitting their milestones. The development process started with building a full technical assessment – ‘everything from how it is going to be used, what our end data will look like, to what format it needs to be in so that our users can use it’ explained McColl.

Then we work backwards: what original data do we need to find, where is it located, and how do we bring it in? Then we start to discuss all the pipelines that handle that data: which databases they are stored in, how to pull data in and out of that database, what types of models we want to use to run this data through, what machine learning techniques we want to use.’ (Fischer and Beswick, 2019, Online interview with David McColl, 21 November)

When it came to information sources, this project was unusual for Rho AI. Ordinarily, the company worked with a client’s own dataset. In this instance, there was no client, and Rho AI had to obtain information from a variety of public sources. The information would come from structured and unstructured sources, including company reports and social media pages, as well as free online information pages containing ESG metrics, such as the World Benchmarking Alliance and WikiRate (McColl, 2019, Email correspondence with Isabel Fischer, 17 September). ‘That is why we decided to go for public equities first – because that data is readily available’, explained McColl (Fischer and Beswick, 2019, Online interview with David McColl, 21 November).

As was its custom, Rho AI had assembled an advisory group for the project that included domain experts (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December). Among other responsibilities, this team’s role was to establish what constituted current, relevant and ethical data collection and reporting for this platform. ‘Part of the process is judging the quality of the sources. Normally we partner with domain experts who understand what is quality data in this vertical’, explained McColl. ‘At Rho AI, we concern ourselves with whether the data is clean and consistent. Then we turn to the domain expertise to assess whether it is valuable – is it the right data to be used?’ (Fischer and Beswick, 2019, Online interview with David McColl, 21 November).

Using a variety of software tools, including those which it had developed itself, Rho AI would develop processes that automatically and regularly downloaded and stored the required data as unstructured information. The goal of this step was simply to identify text that may be relevant to evaluating a company’s climate impact. After this, NLP models based upon methods that included deep learning recurrent neural networks would be used to identify paragraphs, passages, sentences, phrases and/or words that were associated with climate impact (McColl, 2019, Email correspondence with Isabel Fischer, 17 September).

Having done this, Rho AI then planned to use a set of machine learning models to convert this NLP text into structured information (numerical values or rankings). Thus, the tool would make a numeric estimate of the climate impact of the companies concerned, based upon the language used in their text, and impute values where no public comment was made or identified (McColl, 2019, Email correspondence with Isabel Fischer, 17 September).

McColl explained the process further:

A lot of the data that we are pulling in is unstructured. So we have to run natural language processing to understand what it is actually trying to say. A lot of it is documents that are more qualitative than quantitative. So we have to be able to take sentences and run a sentiment analysis to gauge how what they are saying will translate into a sustainability score on a certain metric. (Fischer and Beswick, 2019, Online interview with David McColl, 21 November)

Using data that linked companies’ language use with their historical climate impact, Rho AI could build a model that would predict the linkage between language use and actual climate impact currently and in the future. Then, it could use these data to train the AI tool to predict how companies’ current language in the area of climate change would translate into action. ‘You have to have something to train it on’, said McColl. ‘Then the tool can look at the new information that is coming in and make its best predictions’
We build into our models our confidence in our predictions – so ‘this language translates to this metric and we’re about 50% confident in it’. We don’t want to be binary with these things. Some things you can be more confident in – so things like SEC data you can have much more confidence in than a LinkedIn post – but we try to develop methodologies for saying how confident we are in a prediction, based on historical data. This confidence is integrated into the score that the company gets in the end. (Fischer and Beswick, 2019, Online interview with David McColl, 21 November)

‘It’s not a perfect science to do this – there is definitely human input to determining what should be weighted and where’, McColl added. To establish the methodology for this project, Rho AI planned to establish a group of its own data scientists along with the advisory group. The role of the domain experts was to look beyond the numbers and to evaluate how the data should be factored (Fischer and Beswick, 2019, Online interview with David McColl, 21 November).

These factors were important in mitigating against personal bias. In addition, Rho AI used a training set to highlight instances of bias caused by technical issues. ‘At the end of the day, if you’re doing something in data science, the model is only as good as the data you build on’, said McColl. ‘The reduction of bias has to come at the very initial stages, when you put together the data set. Because afterwards, this is what the model will listen to’ (Fischer and Beswick, 2019, Online interview with David McColl, 21 November).

Since the company intended to develop this as an open-source project, Rho AI wanted to ensure transparency and integrity within the platform. The company wanted the platform’s inner workings to be auditable and for users to be able to discuss and understand how the data were collected, factored, organised and reported, as well as its implications. Similarly, Rho AI intended to work with industry partners to confirm the outputs of their model. Integrating domain expertise would allow the company to assess the accuracy and verify the applicability of the modelling methods used.

AI and ethics under the microscope

In the late 1990s, computing power finally caught up with the ambitions that humans had held for AI for decades, and IBM was able to develop Deep Blue, the chess programme that defeated the reigning world chess champion Garry Kasparov. Since then, the use of AI has become so ubiquitous that the term ‘Fourth Industrial Revolution’ was coined to describe it impact (Anyoha, 2017; Madiega, 2019; Schwab, 2020). AI tools are involved in an increasingly broad spectrum of activities, including credit scoring, detection of tax evaders, algorithmic share trading, prediction of car robberies, parole decisions, online dating and drone warfare (Rahwan et al., 2019). Writing in Nature, Rahwan et al. (2019: 477) observed, ‘Machines powered by artificial intelligence increasingly mediate our social, cultural, economic and political interactions’.

Among the concerns that Rahwan et al. (2019) expressed relating to AI were the following:

- Although the code could be simple, the results could be complex, resulting in ‘black boxes’;
- The code could be difficult to interpret;
- Imperfections in data could impact on the results produced;
- Much of the source code and training data were proprietary.

Some of the algorithm-specific questions Rahwan et al. (2019) raised as being relevant in the domain of machine behaviour were as follows:

- Does the algorithm create filter bubbles? Does the algorithm disproportionately censor content? Do algorithms manipulate markets? Does the behaviour of the algorithm increase systemic risk of market crash? Do algorithms of competitors collude to fix prices? Does the algorithm exhibit price discrimination? (p. 479)

European Union (EU) research has shown that while AI technologies can be very beneficial from a social and economic point of view, they raise several ethical, legal and economic concerns with regard to the risks facing fundamental freedom and human rights. Thus, the EU said, AI poses risks to privacy and to personal data protection, as well as discrimination when algorithms focus on profiling people (Madiega, 2019).

In 2019, the European Commission introduced a set of Ethics Guidelines for Trustworthy Artificial Intelligence, which it is set to confirm in 2020. The guidelines set out seven requirements for AI systems to be regarded as trustworthy. The guidelines cover the areas of human agency and oversight, technical robustness and safety, privacy and data governance, transparency, diversity, non-discrimination and fairness, societal and environmental well-being and accountability (European Commission, 2019).12

Regarding the broader ethics associated with Rho AI’s products, Callsen said the following:

*It is a constant topic on the one hand, and a little in the background on another. Some of our products involve no ethical considerations at all – for example Pit Rho: if a team uses our information and we have a biased training set that gives them the wrong answer, they are not going to win the race and they will stop using our product. We are not endangering anybody, there is no real ethical problem there. (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December)*
Exhibit 1. Simplified overview of techniques used as part of this project.
Source: Developed by Isabel Fischer and David McColl for the purpose of this case.

Exhibit 2. Areas covered by the EU AI guidelines.

1. ‘Human agency and oversight: AI systems should empower human beings, allowing them to make informed decisions and fostering their fundamental rights. At the same time, proper oversight mechanisms need to be ensured, which can be achieved through human-in-the-loop (HiTL), human-on-the-loop (HoTL), and human-in-command (HiC) approaches’;
2. ‘Technical Robustness and safety: AI systems need to be resilient and secure. They need to be safe, ensuring a fall back plan in case something goes wrong, as well as being accurate, reliable and reproducible. That is the only way to ensure that also unintentional harm can be minimized and prevented’;
3. ‘Privacy and data governance: besides ensuring full respect for privacy and data protection, adequate data governance mechanisms must also be ensured, taking into account the quality and integrity of the data, and ensuring legitimised access to data’;
4. ‘Transparency: the data, system and AI business models should be transparent. Traceability mechanisms can help achieving this. Moreover, AI systems and their decisions should be explained in a manner adapted to the stakeholder concerned. Humans need to be aware that they are interacting with an AI system, and must be informed of the system’s capabilities and limitations’;
5. ‘Societal and environmental well-being: AI systems should benefit all human beings, including future generations. It must hence be ensured that they are sustainable and environmentally friendly. Moreover, they should take into account the environment, including other living beings, and their social and societal impact should be carefully considered’;
6. ‘Accountability: Mechanisms should be put in place to ensure responsibility and accountability for AI systems and their outcomes. Auditability, which enables the assessment of algorithms, data and design processes plays a key role therein, especially in critical applications. Moreover, adequate an accessible redress should be ensured (European Commission, 2019).’

EU: European Union; AI: artificial intelligence.

Exhibit 3. Indicative business models.
The aim of this open-access and open-source tool is to be socially useful by contributing to the democratisation of information and by limiting the impact of climate change. Rho AI’s initial focus might therefore be to apply for grants from either public or private funding bodies. Alternatively, there are several business models Rho AI might want to consider for monetising their tool. These include the following:

- A subscription model linked to either one-off access fees or transactional fees;
- Digital marketing offered to other parties, for example, selling display advertising;
- A ‘freemium’ pricing strategy where the standard product or service is provided free of charge, either in combination with advertisement served to users as part of the ‘free of charge’ option, and/or where additional features are chargeable, possibly linked to a subscription model;\(^a\)
- Payments by third parties within a multisided platform where the product or service is available free of charge in one layer, with value being monetised at a different layer. Financial institutions, which act as custodians of securities and other assets, are an example of a third party (‘side’) that might find value in the data linked to the continual learning algorithms that this tool offers, and that therefore might be willing to pay for the service.\(^b\)

\(^a\)See, for example, https://hbr.org/2014/05/making-freemium-work (accessed 15 April 2020).
On the other hand, some of Rho AI’s tools have a real ethical dimension. ‘These can have a truly positive impact’, Callsen said,

but there is a risk there that what our tools predict might have a negative impact on what would otherwise have been good decisions. So we are really focussed on making sure that the tools yield appropriate results for the investment community, so that we can actually be putting capital and resources towards positive, impactful technologies, specifically around carbon mitigation. (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December)

Up until now, however, Rho AI had not codified its approach to the ethical development of AI. ‘We are small enough that our methods are far more informal than would be required for a larger organisation’, said Callsen. But as the company has grown, it is becoming more necessary for Rho AI to codify its approach. ‘When you’re a team of 10 people, you just don’t need to’, he explained.

You know who you’re working with, everybody gets it, you talk about it with everybody, all the time. That core founding group of maybe 10 people: you’re extremely tight knit. Everyone is walking in lockstep. You don’t need to write anything down. Then you get into the kind of 10-20 people range and you’re starting to feel the growing pains. But in that range, I’d consider any employee to be part of the founding team – with the same ethos of ownership. In the 20-30 range, you will still get people who have that same sense of ownership, but you are going to get people who don’t have that – for whom it is just a job. When you get to that stage, it is important to have things codified. (Fischer and Beswick, 2019, Online interview with Gilman Callsen, 4 December)

Concluding thoughts

‘So, do you think we have everything covered?’ asked Callsen, as he and McColl spoke later that Monday. McColl was concerned about the business model, in particular, the potential conflict between making the product open-source and freely accessible to any user, with transparent training sets, and the fact that it would cost in the region of US$300,000 to develop the beta version. There would be additional costs to roll it out, to maintain it and, potentially, to expand into other areas and regions. Rho AI had not been funded by venture capital and did not have the wherewithal to bankroll this project indefinitely. Perhaps the company could apply for grants, or consider a different business model (see Exhibit 3 for indicative suggestions) to help fund the development and to ensure that the product was of a high-enough specification that others could find value in the data. He and Callsen had to consider whether it was feasible to allow investors to benefit from the insights ‘for free’. Also, would transparency enable competitors to copy the model and then market it on a proprietary basis?

McColl was also reflecting on the data quality that could be achieved by web scraping and interpreting ESG-related unstructured data, and the impact the data and the tool overall might have on companies and society. Had the ethical considerations been considered fully and the potential for bias been reduced as far as possible? Would the tool contribute to the democratisation of information? And would the tool, via the investments of the users, pressure corporations to intensify their focus on limiting their climate impact in a meaningful, effective and sustainable way?

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The author(s) received no financial support for the research, authorship and/or publication of this article.

ORCID iD

Isabel Fischer  https://orcid.org/0000-0001-7185-7579

Notes

1. So-called because participants initially remained standing during these short meetings. For Rho AI, most of staff worked remotely and the daily stand-up was conducted online via dial-in.
2. This product was still in Rho AI’s portfolio in 2020 and was being used by five National Association for Stock Car Auto Racing (NASCAR) racing teams to enable them to make better racing strategy decisions in real time.
3. Various terms are used when referring to investment for social, environmental and governance purposes: socially responsible investment (SRI), ESG investment and impact investment.
4. For example, when comparing Milan to other parts of the Lombardy region.
7. See Exhibit 1 for a simplified illustration of the techniques used as part of this project.
10. Brief clarification on the role of the domain experts: The contributions of the experts are to check general approaches...
and data quality, the experts do not continuously check each
decision, nor do they provide (training) data.

11. See https://www.investopedia.com/terms/d/deep-learning.asp
(accessed 10 January 2020); also see www.investopedia.com/
terms/n/deep-learning.asp (accessed 10 January 2020).

12. See Exhibit 2 for more details on these guidelines.

13. See, for example, https://cranetool.org (accessed 29 March
2020).

14. For example, to increase the adoption rate of the tool, that is,
for marketing to make users aware of the salience of the topic
and/or the user-friendliness of the tool.

15. For example, for cloud storage and customer service.

16. See, for example, Raisch and Krakowski, 2020.

References

Anyoha R (2017) The history of artificial intelligence. Science in the
News: Blog Special Edition on Artificial Intelligence, 28 August,
Harvard University Graduate School of Arts and Sciences.
Available at: http://sitn.hms.harvard.edu/flash/2017/history-

Common Fund Institute (2013) From SRI to ESG: The Changing
World of Responsible Investing, September. Available at: https://

DB Climate Change Advisors (2012) Sustainable Investing:
Establishing Long-Term Value and Performance, June.
Deutsche Bank Group. Available at: www.db.com/cr/en/docs/

AI, April. Available at: https://ec.europa.eu/digital-single-
January 2020).

Foster B and Tabit D (2019) As demand for ESG investing grows,
so too does demand for high-quality data. Pensions and
Investments, 22 April. Available at: www.bloomberg.com/
professional/blog/demand-ESG-investing-grows-need-high-

Gartner (2019) Smarter with Gartner. Top Trends on the Gartner
Hype Cycle for Artificial Intelligence, 12 September.
Available at: https://www.gartner.com/smarterwithgartner/
top-trends-on-the-gartner-hype-cycle-for-artificial-intelligence-
(accessed 3 September 2020).

Global Sustainable Development Alliance (2019) 2018 Global

Gross A and Hook L (2020) Academics probe links between corona-
virus and toxic air pollution. Financial Times, 30 April. Available at:
https://www.ft.com/content/21f4ff6f-57ac-43cc-aeb5-602a127553de


In SY, Rook D and Monk A (2019) Integrating alternative data
(also known as ESG data) in investment decision making.

Intergovernmental Panel on Climate Change (IPCC) (2018)
Intergovernmental panel on climate change. Global Warming
of 1.5°C. Intergovernmental Panel on Climate Change.
Available at: https://report.ipcc.ch/sr15/pdf/sr15_spm_final.

McKinsey (2019) Five ways that ESG creates value. Report,
November. Available at: www.mckinsey.com/business-func-
tions/strategy-and-corporate-finance/our-insights/five-ways-
that-esg-creates-value?cid=other-eml-tnn-mippckhlkid=
f053a85eba64e2092ba2117fe231610&hkey=11599349&h
dpid=9d5ed7cf6db4ec6a9f49-de48c258ed0 (accessed 2
January 2020).

Madiega T (2019) EU Guidelines on Ethics in Artificial
Intelligence: Context and Implementation. European
Parliamentary Research Service, September. Available at:


Raisch S and Krakowski S (2020) Artificial intelligence and man-
agement: The automation augmentation paradox. Academy
of Management Review. Epub ahead of print 11 February.
DOI: 10.5465/2018.0072.

Schwab K (2020) Schwab’s profile page as founder and execu-
tive chair of the World Economic Forum with a link to
Schwab’s book. The Fourth Industrial Revolution Seen
as the Origin to the Aforementioned Term. Available at:
https://www.weforum.org/about/klaus-schwab (accessed 4
January 2020).

climate change. The Paris Agreement. Available at: https://
unfccc.int/process-and-meetings/the-paris-agreement/the-
paris-agreement (accessed 2 January 2020).

Weaver M (2018) Andrew Maness leading a NASCAR statistics
revolution. Autoweek, 2 January. Available at: https://
autoweek.com/article/nascar/andrew-maness-leading-nasc-

Author biographies

Isabel Fischer is an Associate Professor of Information Systems
and Management at Warwick Business School, UK. As part of her
work for the UN PRME UK network she focuses on how impact
investing can stimulate innovative technologies that address cli-
mate change. Prior to joining academia Isabel worked in
Technologies and the Financial Services Sector.

Claire Beswick is the Case Centre Manager at Wits Business
School, RSA, where she is responsible for writing and editing
teaching cases on topics that cover the full range of management
disciplines. She has written more than 30 case studies.

Sue Newell is a Professor of Information Systems at Warwick
Business School, UK. Her research focuses on innovation pro-
cesses, studying this from a knowledge and organizational net-
working perspective. She has published over 100 peer-reviewed
articles in both IS and Management journals.