

## **KnowRisk - Ethics Report**

Implementing ethics in practice: ethics for supply chain risk identification, commercial property insurance, and the advanced technology that underpins it

# June 2021



## Table of contents

<b>Section I: Executive summary</b>	<b>3</b>
<b>Section II: Introduction to KnowRisk Project</b>	<b>4</b>
<b>Section III: Introduction to the current challenges within technology ethics, Digital Catapult's practical approach to ethics</b>	<b>5</b>
<b>Section IV: Core activities undertaken &amp; areas of ethical concern for the project</b>	<b>8</b>
<b>Section V: The Ethics Roadmap</b>	<b>12</b>
<b>Section VI: interpretation of results and discussion</b>	<b>16</b>
<b>Section VII: Feedback and how to improve ethics as a service</b>	<b>22</b>
<b>Closing remarks</b>	<b>23</b>

## Section I: Executive summary

The need for supply chains to be redesigned is critical, and this has been felt acutely during the pandemic. Supply chains often leave a 'black hole' for commercial property insurers, and the companies within supply chains are not able to have a full awareness of their supply chain risks. This is what the KnowRisk project aims to achieve: provide insurers and companies with an enhanced snapshot and understanding of risk. The opportunity for enhancing supply chains using a number of technologies - such as the combination of machine learning, distributed ledger technologies and geospatial data - is vast, and would be a great benefit to both business and society. As with all great opportunities and benefits, there are also great risks. The complexity in the challenges is balancing the number of tradeoffs in design, development and deployment.

In the case of the KnowRisk consortium, which uses all the above technologies, critical ethical questions have included:

- How can data *responsibly* be used to create and evaluate supply chain risk and commercial property insurance models?
- What may be some of the unintended consequences of using self-sovereign identity within distributed ledger technologies and blockchain?
- What are some strong alternatives to traditional business models that can ensure ethics is embedded throughout a product's lifecycle?
- Will transparency in supply chains negatively impact SME businesses, and how might this be mitigated?
- How can consortia be governed to ensure participants best engage and consider ethical risks?

This paper and the KnowRisk consortium does not purport to have all the answers to these complex questions. Rather, this aspect of the project is an exploration of these issues in real life; experimented with and applied to the pilot that is KnowRisk. Responsible innovation and applied technology ethics is not just about having the 'right' answers, but instead about how to embed repeatable and auditable processes to ensure risks have been identified, given adequate thought, and plans to mitigate against them. This procedural regularity<sup>1</sup> is what success has looked like for the KnowRisk consortium.

This report will cover: 1) identification of ethical challenges for the KnowRisk consortium, 2) the engagements and activities undertaken to be able to address them, 3) the practical ethical problems dissected for this piece of work and resulting choices for the consortium, 4) feedback on the process of ethics as a service and how to improve applied ethics in future applications and deployments of these technologies in what is a critical area for the UK and global economies.

As a result of the documentation of this work, we hope that other companies, and groups of companies working together with shared goals across a variety of supply chains, can learn and use this as a reference for their own successes.

---

<sup>1</sup> See page 9, <https://arxiv.org/pdf/2102.09364.pdf> for more in depth information on the concept of procedural regularity.

## Section II: Introduction to KnowRisk Project

In today's globalised economy supply chains are highly complex, and a problem in one part of the network can impact a multitude of businesses resulting in risks that are highly fluid and dynamic. We currently have little visibility into these risks. A business will have one view of its risk, auditors another, lawyers yet another, and insurers will have a detailed view for 5% of commercial sites, but are left with statistical models for the other 95%. Currently, these fragments are not brought together to create a more holistic view of risk, and nobody has a view of the flow of goods and services through the end-to-end supply chain. Many supply chain businesses are unable to access insurance, and are suffering \$500bn of uninsured losses per year, losses which often result in their unnecessary closure.

The KnowRisk Consortium was created to solve this problem. Using the latest technologies it aims to bring together a business' own internal data alongside Accounting, Insurance and Legal (AIL) data, which is then augmented with geo-spatial data, IoT data, and over 300 third party data sources to create a 360 degree view of risk. Today businesses reactively deal with a problem after it's occurred. With KnowRisk's real-time data they will be able to collaboratively manage risks and proactively reduce their frequency and impact. By creating visibility into an individual business' risk, as well as risk right across its supply chain, KnowRisk can help business avoid problems, and ensure they have the right insurance that they need when things do go wrong.

The KnowRisk project has combined AI, distributed ledger technologies (DLT), and geospatial intelligence (GEOINT) to serve this critical pain point for insurers across the supply chain, and move towards adaptive and robust supply chains. The partner companies involved in the KnowRisk project are SweetBridge, Engine B, Cystellar, Digital Catapult, Industria Tech and Intelligent AI, with Sweetbridge being the leading partner.

As with any nascent technology there are always associated ethical challenges. Given this pilot integrates a number of technologies together, involves a number of players, and has the potential to impact numerous companies financially, it is critical that the ethical issues be properly identified and tackled head on. This is the first experimental attempt to apply practical ethics to a consortium and this work explores the multiple players and considerations in a digital supply chain on both sides.

## Section III: Introduction to the current challenges within technology ethics, the KnowRisk consortium's practical approach to ethics

This section outlines the current challenges within the field of technology, ethics and responsible adoption, highlighting the importance and need for the approach undertaken by the KnowRisk consortium. This section also details Digital Catapult's previous work within

applied ethics, and how a modified approach has been undertaken to drive impact for the KnowRisk consortium.

## Making applied ethics impactful

Within the advanced digital technology space, technologies such as machine learning and blockchain are moving at a rapid pace, meanwhile the legislations and regulations around how the technology should be used, are not. Indeed, given how fast technology progresses, it would be a near impossible feat to have regulation constantly updating itself and covering all grey areas of technology development and application. In the absence of 'hard' governance mechanisms,<sup>2</sup> such as explicitly clear legal rules, there has been a surge in 'soft' governance mechanisms, such as codes, guides and principles to encourage those working within the technology space to consider the parameters of application. Typically, these guides and frameworks share values in common, and often refer to concepts such as 'justice', 'beneficence' or 'autonomy'.<sup>3</sup> It is, however, difficult to define what these concepts mean in practical terms, and there may be differences in their perception or application depending on cultural or geographical variance, with priorities shifting over time. This may leave practitioners with questions on how to implement these values and institutionalise ethics in practice, into their products and services. In addition to this embedded challenge, many AI ethics or technology ethics tools often are used or interpreted as one off events, or worse, used as 'ethics washing', i.e. to exaggerate or window-dress a company's interest in making ethical decisions. This thereby limits the amount of impact these ethical tools can have on mitigating against the technologies' potential harms and risks.

Consequently, the approach undertaken by the KnowRisk consortium in response to applied AI ethics is one of 'procedural regularity'.<sup>4</sup> It is not about necessarily looking to provide answers to very complex questions, i.e. what is 'justice', but rather procedural regularity looks to create repeatable and auditable processes, which engrain responsibility, deliberateness, and conscientiousness into product, culture and business model. With a growing backdrop of distrust towards technology companies,<sup>5</sup> companies may sometimes use the defence that they were unaware of the impact the choices would have on the technology being developed.<sup>6</sup> Whilst it is impossible to be omniscient, we hope that practical ethics will enable practitioners to truly consider, from the outset, the implications of their technology and product development, to mitigate against risk. And, more importantly, we hope to create positive case studies around the value and commercial benefits that applied ethics can bring to technology.

---

<sup>2</sup> Consider policies leaving a lot of 'grey areas' for technology usage, for example how GDPR is insufficient to protect individuals' privacy in light of inferences that machine learning models can make. See:

<https://www.law.ox.ac.uk/business-law-blog/blog/2018/10/right-reasonable-inferences-re-thinking-data-protection-law-age-big>

<sup>3</sup> <https://www.digicatapult.org.uk/for-startups/other-programmes/applied-ai-ethics-typology>

<sup>4</sup> <https://arxiv.org/pdf/2102.09364.pdf>

<sup>5</sup> <https://www.computerweekly.com/news/252483007/British-public-largely-distrustful-of-technology-companies-says-report>

<sup>6</sup> See CNN transcript around Facebook and Cambridge Analytica scandal:

<http://transcripts.cnn.com/TRANSCRIPTS/1803/30/qmb.01.html>

## Digital Catapult's previous work in practical ethics: ethics as a service

In previous work, Digital Catapult experimented and tested a practical methodology to apply ethics to AI. In particular, the practical work focused on implementing and embedding ethics into early stage machine learning startups. Digital Catapult developed a framework with their Ethics Committee, (chaired by Prof. Luciano Floridi at the University of Oxford) an independent group of experts in this field. The Framework translates high level principles into practical questions that illuminate how they are relevant to business, people and technology decisions.<sup>7</sup> The framework comprises seven core principles; each of its core principles has an associated set of questions to facilitate a reflective, consultative and deeply practical approach. The AI Ethics Framework is used to support conscientious decision making and promote responsible, questioning, and thoughtful startup cultures.

## Applying this work to a consortium project: KnowRisk

Given this previous experimentation of applying ethics to early stage machine learning startups, the KnowRisk consortium were keen to apply this same approach to the development of their platform. This exercise in application of ethics to the KnowRisk platform, was far more challenging and complex than with just a single startup. Given KnowRisk is a multi-technology, multi-stakeholder environment, with a number of companies working on different technologies for the same mutual goal, it was necessary to amend this methodology and process to meet the needs of the consortium.

The two ethics advisors, Professor Burkhard Schafer and Dr Laura James, drove the ethics journey for the consortium. They designed and delivered a number of ethics workshops with the consortium as a whole, as well as individual company members, to discuss the risks on a macroscale and other company specific concerns. These risks will be outlined in greater detail in sections 4 and 5 of the report.

Within the KnowRisk project, ethics serves a central utilitarian and commercial function, as well as a practical conduit for harm and risk mitigation. Ethics has been embedded into the development and coordination of this project with the objective of sustained, long term success of the project. This process encourages all different parties to make thoughtful decisions, consider unintended consequences and justify their approaches in development.

The KnowRisk project at its heart has a constructive ideal of responsibility, that is, doing 'good' where possible, (opposed to 'constrained' ideal, which is strictly about following legal rules) and therefore are committed to embedding these ideals through design. This approach, we hope, is a tangible method of mitigating short and long term risk; avoidance of myopic choices.

---

<sup>7</sup>[https://assets.ctfassets.net/nubxhjiwc091/xTEqMcYudwQ7GHZWNoBfM/c2a2d55a0ee1694e77634e240eafdfdf/20200430\\_DC\\_143\\_EthicsPaper\\_\\_1\\_.pdf](https://assets.ctfassets.net/nubxhjiwc091/xTEqMcYudwQ7GHZWNoBfM/c2a2d55a0ee1694e77634e240eafdfdf/20200430_DC_143_EthicsPaper__1_.pdf)

## Section IV: Core activities undertaken & areas of ethical concern for the project

This section outlines:

- 1) The sequence of activities undertaken for 'ethics as a service'
- 2) The independent ethical advisors and their professional backgrounds
- 3) Development of a bespoke ethics framework
- 4) Key areas of ethical concern for the project

### Sequence of activities undertaken and led by independent ethics experts and the KnowRisk consortium

1) External data and AI ethics experts onboarded to provide impartial consultative advice →	2) Ethics experts briefed and given independent time to scope out the specific risks →	3) Ethics experts to build an ethics roadmap for consortium to use →	4) Individual workshop with each consortium member to discuss individual risks →	5) Quarterly consortium ethics workshop as to support with ongoing developments →	6) Company led ethical risk identification activity →	7) Brown bag lunches on ethical risks most pertinent to them →	8) 'Flow risk' and 'node risk' office hours held with advisors	9) Final workshop with entire consortium to discuss final issues as the pilot is developed →	10) A roundtable to review the ethics process and advise on what could be improved
--	---	---	---	--	--	---	--	---	--

Throughout the project, the KnowRisk consortium worked in an agile manner and iteratively. As a consequence, activities 6 and 8 listed above were not originally within the plan for delivery, but the consortium members input these activities as a fruitful way to increase and continue engagement. As will be discussed in more depth in the paper, the success of the project is hugely attributed to the enthusiastic participation of the individual KnowRisk consortium members and their attentiveness to the needs of the ethics work and product, with such responsive tweaks exemplifying how they were able to achieve that.

### Independent data and AI ethics experts consulted

For the partners of the KnowRisk consortium, it was important to engage experts who were impartial and independent of the product development itself. It is uncontroversial to say that raising ethical concerns within a company can typically be met with resistance, sometimes even persecution of the ethics whistle-blowing individual.<sup>8</sup> This approach looks to create a system and dynamic whereby the ethics consultants are working as trusted external and independent ethics advisors, with the sole interests of the collective consortium in mind, and

<sup>8</sup> <https://www.ft.com/content/4f8ebc3a-d714-4d79-9805-613303a29420>

it cannot be left to the discretion of any one company to engage or disengage from their advice or recommendations. It is stressed that they are purely advisory in nature, and businesses are not coerced to take their advice - conversely, in our experiences businesses do wish to take a lot of their advice as typically these recommendations will also create a positive impact on their product quality. The structure of the engagements, with quarterly check-ins from the advisors, is intended to create a long standing culture of ethics within the consortium, and individual companies.



**Professor Burkhard Schafer**

*Professor of Computational Legal Theory at the University of Edinburgh*

Burkhard Schafer studied Theory of Science, Logic, Theoretical Linguistics, Philosophy and Law at the Universities of Mainz, Munich, Florence and Lancaster. His main field of interest is the interaction between law, science and computer technology, especially computer linguistics. How can law, understood as a system, communicate with systems external to it – be it the law of other countries (comparative law and its methodology) or science (evidence, proof and trial process)? As a co-founder and co-director of the Joseph Bell Centre for Legal Reasoning and Forensic Statistics, he helps to develop new approaches to assist lawyers in evaluating scientific evidence and develop computer models which embody these techniques. A special interest here is the development of computer systems that help law enforcement agencies to co-operate more efficiently across jurisdictions, assisting them in the interpretation of the legal environment within which evidence in other jurisdictions is collected. This research is linked to his wider interest in comparative law and its methodology, the idea of a “Chomsky turn in comparative law”, and the project of a computational legal theory.



**Dr Laura James**

*Entrepreneur in Residence at the University of Cambridge*



Holding a PhD in Engineering from the University of Cambridge, Laura James works with emerging technologies in new and growing organisations across sectors, and has been active in the tech responsibility space since 2016, with a focus on practical ways to improve industry practice. Working with businesses and learning about their technologies, challenges and opportunities has always been fascinating to her, and she enjoys supporting early stage and growing organisations. Laura is very experienced in enabling startups and scaleups to act responsibly with regards to their users, society more broadly, and other stakeholders, as well as exploring the tradeoffs and choices they face.

### **Developing a bespoke Ethics Framework**

As mentioned in the previous section, Digital Catapult Machine Intelligence Garage has an ethics framework dedicated specifically to early stage machine learning startups. Using this (experimented and piloted) framework as a foundation, Professor Schafer and Dr James worked collaboratively to update the framework to reflect the needs of an advanced technology consortium. The updated framework also moves away from primarily considering the impact of machine learning, to include additional technologies such as decentralised and distributed ledger approach to the storage of potentially commercially sensitive information.

The thought process behind revising the framework was to attempt to maintain the generality and applicability of the framework as far as possible, whilst respecting that some aspects may be quite specific due to the nature of the KnowRisk ethics work. Unlike ethics consultations for machine learning individual startups, where harm is more likely to be limited to individuals or groups, the KnowRisk consortium's harms could impact entire market economies or countries. The full amended Ethics Framework for the consortium is included in the Appendix.

### **Key areas of ethical concern for the project**

This section will outline key areas of ethical concern for the KnowRisk project. These were the areas identified at the *start of the project*, to aid the consortium members in thinking about the most pertinent challenges. This includes: data and machine learning ethics; supply chain ethics and legislation; power dynamics around transparency; issues of unrepresented stakeholders. By the end of the project, and as will be discussed in sections 5 and 6, these issues are explored by the consortium in depth.

#### **Data and machine learning ethics**

Typically, the law emphasises the areas of focus and concern for companies. The past few years, the 'ethics' conversation has been dominated by questions around personal data and GDPR compliance. This is especially the case due to the severe financial penalties from noncompliance. However, there are still a number of harms that can be derived from using data combined with machine learning: even if this is outside of the parameters within which GDPR operates. Environmental harm is a key aspect of this for example; given the KnowRisk prototype will be used in an insurance context, if environmentally harmful practices are rewarded with a low economic risk score, it may incentivize, or at least facilitate, myopic environmental practices. We have seen insurance as an industry previously incentivise all kinds of behaviours, some potentially positive and negative - for example, insurance in the

past has 'nudged' people to implementing more safety features in their homes due to the rewards in lower insurance premiums. It is important to understand the power of incentivisation that insurance companies therefore have on different people.

The Ethics Advisors on this project highlight the importance of understanding data: data does not always reflect reality: its selection expresses particular concerns, interests and world views. The process of measuring, or data collection in itself can distort what it is trying to measure.

This project is particularly privacy conscious: it uses federated learning, where analysis software is run on site where data is stored, without having to explicitly share data among parties, enabling parties to benefit from an ML model that has been trained on a variety of datasets, but still preserving each of the parties' privacy. Whilst there are clear benefits in privacy preservation, its utilisation may not be without its own challenges: how can the efficacy and utility of federated learning be demonstrated and evidenced? Furthermore, how can a consortium using federated learning maintain transparency when the underlying data is only partially visible (each party can only see their local dataset)? It is imperative to ensure that effectiveness or predictive accuracy of the model are evaluated on an ongoing basis. Providing practical mitigations to these shortcomings of Federated Learning in a consortium context through the use of carefully selected applied AI Ethics tools is the focus of the accompanying Ethics Tools Report for KnowRisk.

### **Supply chain ethics and legislation**

The KnowRisk project happened and is happening in tandem with two major events: leaving the European single market and Covid-19, which have highlighted how supply chains are essential and yet extremely fragile in the face of shocks. Currently, social-economic factors are such that supply chains are optimised for efficiency, to maximise flow of goods, yet, there is less focus on how to make them resilient. Whilst this may be seen as following the objectives and profit structures of revenue driven businesses, this starts to become an ethical problem when essential services, such as food or medicines fail to reach countries or cities in the quantities needed as a result of failures in the supply chains. Richer areas or countries might be less impacted than less affluent areas. It is therefore important to create systems, infrastructure and incentives which take into account the needs of societies as well as company profit margins.

It has been also recognised that there are a number of human rights violations or harmful environmental practices within supply chains, often at entry points. In 2017, the UK Parliament Committee called for prosecution of parent [companies linked to supply chain abuse](#), where UK companies have been found to not respect human rights in their operations overseas. For KnowRisk, this is an important contextual background. Given the duties of due diligence, regulators might become interested in data made available through the prototype of KnowRisk, and where appropriate, use it against the companies that generated this data. Consequently, it is another question around alignment of incentives: it is imperative to ensure that companies are encouraged to uncover (and address) human rights violations within their supply chain without being placed at a competitive disadvantage for generating this data. The questions of how to engage with regulators, and how to communicate this with other stakeholders is key, as well as being wary of new incentives for companies to falsify records, which may impact on the degree in which tools are openly shared. If KnowRisk's technology becomes used for certification, or third parties use the technology to

certify transactions, this could increase the demand on accuracy and correctness of results, as well as new transparency duties. It could also make it de-facto impossible for participating companies to switch to another platform, creating technology lock-in, which might impact companies with less financial resources more strongly.

### **Power dynamics around transparency**

Transparency is often referred to as an unalloyed good. However, access to information does not equally benefit all stakeholders. In the case of supply chains, larger parties might sometimes use transparency to extract profit from smaller ones. This might for example happen if players want to undercut other companies in the supply chain or try to exclude them entirely. There are also risks around economic warfare, particularly in the current geopolitical landscape. As a result of this tension, the KnowRisk consortium would like to make clear that the project does not aim to increase supply chain transparency, rather the aim is to increase accountability without requiring transparency (and the associated problems). It is hoped that this is what will make the project so valuable in light of the current issues with supply chains.

## Section V: The Ethics Roadmap

The ethics roadmap is a critical part of this ethics workstream and engagement. Many technology ethics frameworks exist - there are over 160 frameworks around ethical use of AI as of 2020.<sup>9</sup> Many of these guides are aspirational, or encourage us to consider what we may believe to be right or wrong when it comes to technology and ethics. Whilst this body of literature has laid the groundwork to bring us closer to what “ethical” technology might look like, what these guides often lack is how to operationalise these values to create business outcomes. The intention for this bespoke roadmap, is a critical tool to bridge this gap between aspiration and reality.

This ethics roadmap was developed by our independent ethics consultants. It was produced after the first group consortium session, as well as each of the individual company sessions. The advisors picked up on prevalent themes throughout each of the discussions and developed them into actionable recommendations. The consortium was asked for feedback on the roadmap, on any omissions or amendments - making this map a collaborative document on what could be achieved realistically within the timeframe.

### In the first month

#### Across the consortium

- Continue work to help all consortium members understand and feel involved in the Knowrisk vision and overall aims.
  - Consider a ‘storytelling’ style informal remote meeting, where key project leaders can talk through the vision and ambition (with examples and ideas, rather than slides and diagrams), for the whole project team.
- Identify a forum or process by which individual or team concerns as well as positive suggestions about ethics in the project can be raised at consortium level, and also internally if there are not yet clear procedures in place. This includes ways to assure more exposed members (on short term contracts, with less social capital) that raising concerns is a positive contribution to the collective effort, and makes it a task for everyone to identify opportunities for the growth of ethical practices.
  - Proposed activities:
    - That team ethics concerns are raised with Burkhard Schafer (BS) and Laura James (LJ) if it is felt that there is no other route to raise them within the consortium
    - That there should be informal talks, perhaps scheduled over coffee or lunch in a ‘brown bag’ style, from different team members or invited guests, run remotely, with discussion time, once a month or so during the project duration, to build team cohesion
    - That some thought should be given to informal moments where team members can get together and chat; for instance, a weekly remote coffee slot where you can drop in, or not, at will
- Relevant team members across the group should meet to discuss the threat model which impacts data privacy in Knowrisk, and write up the results of this.

---

<sup>9</sup>See:

[https://assets.ctfassets.net/nubxhjiwc091/xTEqMcYudwQ7GHZWNoBfM/c2a2d55a0ee1694e77634e240eafdfdf/20200430\\_DC\\_143\\_EthicsPaper\\_\\_1\\_.pdf](https://assets.ctfassets.net/nubxhjiwc091/xTEqMcYudwQ7GHZWNoBfM/c2a2d55a0ee1694e77634e240eafdfdf/20200430_DC_143_EthicsPaper__1_.pdf) and <https://inventory.algorithmwatch.org/>

## Cystellar

- Devote some dedicated time (perhaps a couple of hours) as a team working on Knowrisk to reviewing the Ethics Framework and considering how each section impacts the work of the organisation on Knowrisk

## Digital Catapult

- Devote some dedicated time (perhaps a couple of hours) as a team working on AI for Knowrisk to reviewing the Ethics Framework and considering how each section impacts the work of the organisation on Knowrisk

## EngineB

- Devote some dedicated time (perhaps a couple of hours) as a team working on Knowrisk to reviewing the Ethics Framework and considering how each section impacts the work of the organisation on Knowrisk

## Intelligent AI

- Devote some dedicated time (perhaps a couple of hours) as a team working on AI for Knowrisk to reviewing the Ethics Framework and considering how each section impacts the work of the organisation on Knowrisk

## Sweetbridge

- Devote some dedicated time (perhaps a couple of hours) as a team working on Knowrisk to reviewing the Ethics Framework and considering how each section impacts the work of the organisation on Knowrisk

## For November 2020

### Across the consortium

- Ethics check-in with LJ & BS
- Relevant team members to connect and discuss in some way questions of how the ever-evolving nature of machine learning, used in several parts of Knowrisk, might be presented and worked with to users of Knowrisk in the future.
  - Consider creating some materials to explain this to future stakeholders (an explainer, comic, short video, or other form)
- Set up a Knowrisk webpage including a statement on its approach to ethics and contact details/procedures that would allow organisations or individuals who fear they have been unjustly affected to ask for remedies.
- Have had at least one stakeholder workshop or conversation, with for example a risk assessor or commercial insurance buyer. This could either be a workshop with a range of such people, or smaller conversations with relevant project team members and one or two stakeholders.

## EngineB

- Deliver an informal remote seminar to the overall consortium project team about how EngineB has been set up to balance purpose and profit, protect the mission against bad outcomes and actors, and so on. (This will help showcase a different way of shaping an organisation or consortium, which may support thinking for the Knowrisk work after 2020-21).

## Intelligent AI

- Explore issues of debiasing, explainability/transparency, and other core ML ethics questions for IntelligentAI's work on Knowrisk, and make a short informal written briefing or presentation for the project

## Sweetbridge

- Publish an accessible essay, video or other output describing Knowrisk and the Sweetbridge system and model
- Share some of Sweetbridge's thinking about ethics for the future of the project and for Sweetbridge's platform, with the whole project and ideally a wider audience online (eg informal webinar talk with Q&A)

## For January 2020

### Across the consortium

- Ethics check-in with LJ & BS

### Cystellar

- Evaluate Cystellar work on Knowrisk and consider how explainable it is to a general audience (perhaps an FT reader); consider writing a short blog post or similar informal article that sets out the work done, and how it is appropriate, fair and so on. (You may find the ethical framework a useful tool to inform thinking about what such an article might cover.)

### Digital Catapult

- Evaluate Digital Catapult work on Knowrisk and consider how explainable it is to a general audience (perhaps an FT reader); consider writing a short blog post or similar informal article that sets out the work done, and how it is appropriate, fair and so on. (You may find the ethical framework a useful tool to inform thinking about what such an article might cover.)
- Explore issues of explainability/transparency, privacy, robustness and other core ML ethics questions for the Catapult's work on Knowrisk, and make a short informal written briefing or presentation for the project

### EngineB

- Evaluate EngineB work on Knowrisk and consider how explainable it is to a general audience (perhaps an FT reader); consider writing a short blog post or similar

informal article that sets out the work done, and how it is appropriate, fair and so on. (You may find the ethical framework a useful tool to inform thinking about what such an article might cover.)

### Intelligent AI

- Evaluate Intelligent AI work on Knowrisk and consider how explainable it is to a general audience (perhaps an FT reader); consider writing a short blog post or similar informal article that sets out the work done, and how it is appropriate, fair and so on. (You may find the ethical framework a useful tool to inform thinking about what such an article might cover.)

### Sweetbridge

- Evaluate Sweetbridge work on Knowrisk and consider how explainable it is to a general audience (perhaps an FT reader); consider writing a short blog post or similar informal article that sets out the work done, and how it is appropriate, fair and so on. (You may find the ethical framework a useful tool to inform thinking about what such an article might cover.)

## By project end (tbc in 2021)

### Across the consortium

- Consider how to clearly articulate the project vision and components in the final reports and other assets (such as any project website or archive) for accessibility to non-expert stakeholders
- Evaluate the project progress in light of the ideas uncovered in the first ethics workshop around “what we could be proud of” and “what would be the worst future headline”, and consider progress towards/away from these ambitions and fears

### Digital Catapult

- LJ & BS input to final ethics report, with any final conversations/checks within the project complete [not really a roadmap item but including here for completion]

## For future work after this proof of concept project

### Across the consortium

- Review and consider the final ethics report from the 2020-21 project
- Review and consider the accompanying ethics tools report from the 2020-21 project
- Review and consider any parts of the Knowrisk ethics roadmap (this document) which were not completed during the 2020-21 project
- Prioritise ethics and governance questions in the design of any future projects, including allocating responsibilities. Establish regular feedback processes to ensure stakeholders are appropriately engaged and on board with the work

## Section VI: interpretation of results and discussion

This section will outline the outcomes of the KnowRisk ethics project. For clarity, outcomes will be centred around the principles in the ethics framework that was developed for the consortium. Under each principle in bold, the consortium's efforts and the impact of the interventions will be discussed. In addition to this, there may be ethical challenges which arose out of deeper consultation with the ethics experts in the later stages of the project, of which direct actions are still being undertaken. The thoughts and discussions are documented here as they pose important questions for both ethical issues in technology and supply chains. Innovating responsibly is a journey, and it is impossible to make changes overnight.

This work has looked at ethics from a consortium point of view. This has meant that each partner has had to align with one another on the approach and implementation of ethics. This is vastly different to implementing and embedding ethics into sole startups, as companies are typically composed of 2-10 people (most usually, 2-4), and therefore have all decision makers in a room together, and the ability to make quick decisions. The dynamic within startups is also mainly having full context and knowledge of the stages of development and potential issues. Embedding ethics consortium wide is a different endeavour, as typically meetings may have a couple of representatives from each organisation (not its entirety), and whilst there will be strong contextual knowledge it may be more fragmented due to the distributed nature of product development in a consortium. This therefore requires an additional level of engagement, bonding and governance throughout the consortium as a whole to be successful.

**"Ethics isn't a tick box exercise, it is the life and blood of the organisation. It is a requirement to doing business"**

Anthony Peake - CEO, Intelligent.AI

### **Be clear about the benefits of the product or service**

KnowRisk aims to benefit supply chains and mitigate against supply chain disruptions. It provides a first step towards modern, adaptive and robust supply chains for manufacturers, retailers, suppliers as well as insurer, legal and accounting firms.

As the KnowRisk platform has the potential to impact a large number of global supply chains, economies and countries, who the platform is benefiting needs to be made explicit. The consortium has considered at length for example, if there will be a disparate effect on different countries, and if it would contribute to the gap between poorer and richer countries. Given this application is also being used in a global context, it is likely that what might be defined a 'benefit' to one country, may not necessarily apply universally. For example, while transparency is often seen as an intrinsic benefit, it is actually much more instrumentally beneficial to a select number of parties: by having complete transparency, retailers could drive down prices of smaller suppliers across the network to uncomfortable or unprofitable



levels. The consortium are also acutely aware of not trying to impose their world view into other countries that they operate in.

### **Know and manage your risks**

Identifying and mitigating against risk is critical to the success of technology products and platforms. In Digital Catapult's previous work, it was identified that having a good grasp of risks can help with securing investment and customers more effectively.<sup>10</sup> Whilst the KnowRisk platform as a single entity is not yet at the stage to be talking to investors, this has played an integral part in being prepared for the future. Equally, Intelligent AI has successfully raised investor funding during involvement in KnowRisk, and they believe their involvement in ethics was useful in this process to raise this funding.

"Intelligent AI has successfully raised [investment] during KnowRisk... [the] ethics work was discussed with investors, and I believe that made a difference in raising the funding"  
**Anthony Peake, CEO of Intelligent AI**

From an outsider's perspective, it was interesting to deep dive into how very technical individuals consider ethics. In engaging with different companies' technical teams during the initial stages of the consortium, there was sometimes a tendency to defer solely to technical features as providing a solution to ethical problems. For example, if asked, "what happens if data is leaked for unintended purposes?", a response may have been to talk about specific features which would make data leakages very difficult to achieve technically. Throughout the project, it was notable how these responses greatly evolved and transformed. By the end of the project teams were much more empowered to discuss hypothetical (but possible) ethically challenging scenarios and how to best manage and deal with them. In particular, this included the kinds of processes that might be implemented to ensure users can effectively complain about problems using the platform and how to rectify their results. By way of example, the Intelligent AI team have implemented bi-weekly show and tell sessions with the development team, whereby they challenge ethical implications and discuss any ethical concerns candidly and openly.

### **Use data responsibly**

It is important to understand responsible use of data in a machine learning context as not only being an ethics question, but also fundamental to a high quality product. The consortium recognised early on that responsible use of data would also lend itself to strong predictive accuracy of algorithms, whilst balancing precision. This is critical; if users find the platform to not work in the way it was promised to, KnowRisk would struggle to retain customers and find a high churn rate. Consequently, we see ethics also aligning with strong product outcomes.

#### *Using data responsibly in machine learning aspects of KnowRisk*

---

<sup>10</sup><https://www.digicatapult.org.uk/news-and-insights/publication/unveiling-the-commercial-value-of-the-responsible-use-of-ai>

The machine learning elements of the consortium have a very rigorous approach to biases in data. Understanding that data bias has no one quick fix, as all data that is collected, has been collected by someone, for a specific purpose, looking at a defined number of variables, is important for robust and repeatable outputs. Data will have a number of historical ills to it - for example, sloppy data labelling may indicate that the area directly outside of a police station suffers acutely from higher crime rates than the surrounding areas, but the explanation will be that the data has been labelled incorrectly by the person collecting the data. Equally, there may be biases in higher crime rates outside and near police stations, simply because people are more likely to report crime if it is easier to do so, which again creates problems of bias in the dataset.

As a result, the machine learning team, Intelligent AI have built specific toggles in their data visualisations, to be able to discern whether a result changes as a result of the input of additional variables. For example, when looking at the impact of value 'X' on crime statistics in an area, users of the platform could remove the variable 'X' and visualise what other (potentially more pertinent) factors contribute to crime in the same area. This might highlight wrong attributions of causality to specific problems.

Sweetbridge and the wider consortium also have built ethics into data analytics and data visualisation processes, i.e. the ability to see raw analytics and ethics corrected analytics, so clients can see the impact of different features and decide accordingly.

Digital Catapult have also demonstrated the benefits of using a Federated Learning system in order to train machine learning models while protecting sensitive data and maintaining a level of transparency using tools like Model Score Cards for Model Reporting and Record On Negative Impact (RONI).

### *Using data responsibly within blockchain and distributed ledger aspects of KnowRisk*

The consortium discussed at length with the independent ethics advisors some of the issues in using self sovereign identity on blockchain with respect to being responsible with data. There are clear benefits to having autonomy over identity, i.e. not sharing unnecessary excess data, but there are also some ethical issues.

First, as the notion of self sovereign identity (SSI) is individualistic, it may emphasise the individual too much and therefore offer disincentives to collective action or communal collaboration. There are technical approaches that can mitigate against this individualisation. For example, it can be demonstrated that a person has contributed to a group discussion or engagement, without disclosing the nature of the engagement, and then have all individuals who contributed to form a group digital signature on a blockchain. This thereby creates a collective vision and potential for group responsibility, however, these will only go so far as company culture enables it to.

There are also concerns around biases resulting from who decides to not disclose data; research indicates that those who are more privacy conscious tend to be more affluent and more educated,<sup>11</sup> and usually only these groups will exercise their rights within data privacy. This begs the question as to who will benefit most from self sovereign identities, and if it is

---

<sup>11</sup> <https://hbr.org/2020/01/do-you-care-about-privacy-as-much-as-your-customers-do>

perpetuating a system of continuing to protect and provide more security to only the most advantaged groups.

In addition, using self sovereign identity (SSIs) in a corporate context may have some other added benefits and concerns. Zero Knowledge Proofs, used by Sweetbridge enables businesses to ask questions of datasets and get answers without ever revealing the underlying data, protecting privacy. These answers have proofs, demonstrating to businesses that the answers they receive are correct without being able to see the underlying data. This is a crucial aspect of the system, and fully focused around the issue of privacy within supply chains, as well as the protection of commercially sensitive data (protecting the commercial interests of businesses).

As blockchain and distributed ledger technology more broadly operates on the premise of decentralisation, there is a bigger question about what happens on a large scale level when individuals choose not to share data. For example, in the case of discrimination and unfair algorithms, how much knowledge is needed to understand that groups are being marginalised as a result? If data is kept secret, we may not have the bigger picture vision of what groups are being discriminated against and how.

The Sweetbridge team are developing “Proof-Oriented Programming”: instead of sharing data itself, this enables parties to exchange proof that the data they have meets and has certain properties. This enables private aggregation, whereby users are able to see aggregated results, but what specific values contributed to the end aggregation are private. In this context, the ethics advisors spent time with the Sweetbridge team to discuss: how can developers of this system know if this is working correctly? Within a legal context, when it comes to litigation, someone would have to demonstrate that the platform hasn’t worked; consequently, what would be needed to prove that there is a mistake or failure of the system? This emphasised the distinct notions of ‘evidence’ and ‘proof’; and that ‘evidence’, should be understood as ‘evidence for a certain thing’. These sessions evolved the consortium’s approach to the validity of forensic methodology required for the system’s data to be admissible in court.

“It is likely our product will be (technologically) ahead of regulation - we needed to ensure we built the correct processes to go above and beyond, and protect all users”

Scott Nelson - CEO, Sweetbridge

## Be worthy of trust

There is a paradigm when it comes to trust, and trust experts have defined a combination of behaviours and drivers, such as reliability and competence in behaviours, and integrity and empathy as drivers, that make individuals and companies ‘worthy of trust’.<sup>12</sup> Trust then, is something earned over time through consistent behaviours, and the willingness to take action even if it is deemed as potentially undesirable to business in the short term (in the long term the argument would be that it would pay off in the business sense, through trust resulting in retention of customers and winning new clients, investment etc.).

<sup>12</sup> <https://medium.com/@rachelbotsman/being-more-trustworthy-the-basics-6354e504917f>

A way to consolidate trustworthiness in the consortium was through a number of activities, for example each company within the consortium undertook Ethics training and workshops for their entire teams. In addition, Intelligent AI are developing personas for user groups, and for non-user groups, so that they can test the platform against the impacts of each persona. They are also building machine learning algorithms that are auditable and not black boxes. They are also working to ensure that the platform cannot be used to profile and prevent smaller organisations from gaining favorable insurance rates. Other members of the consortium, for example CyStellar, have implemented ethics into data infrastructure and architectures by providing automated DevOps checkpoints, automated tests and validation points, accepted and non-accepted datasets, supplier classifications, and specific guidelines to project managers.

As part of building trust in the platform externally, the KnowRisk project team have engaged with real users throughout the project through the 2Build consortium. This is a group primarily constituted of SMEs, and they have been engaged to ensure that the companies' whose risks (through the KnowRisk platform) are being assessed are understood. This thereby limits the risk that key stakeholders are not represented as part of the product development process.

As the KnowRisk consortium works towards commercialization of the platform by the end of 2021, they are expected to find a third party to collaborate with, such as the social tech trust to enable audits of the agreed ethics standards.

### **Open and understandable in communications**

A key takeaway for individual members of the consortium and more broadly as a group, was learning the importance of communicating the approach to the product externally. The perception of public communications were initially perceived as more of a public relations exercise, but now, the consortium views external communications a critical part of ensuring that the KnowRisk consortium has a measure by which success can be judged internally and externally. The consortium now has draft documents of their ethics statement and stakeholder engagement processes, and intend to publish this to the website in the upcoming months.

### **Promote diversity equality and inclusion in the project partners and teams**

The Ethics framework and workshop enabled the KnowRisk consortium to better identify and appreciate the risks that could result from the KnowRisk platform. This was important, as for a number of members in the consortium, 'doing good' is/was a key motivator for them as individuals and companies. Whilst this is undoubtedly a desirable trait to have, companies and individuals may be led to believe that good will and good intentions are sufficient to avoid harm. The Ethics exercises that the consortium did together led members to think more deeply about the risks that could derail their objective to be fair and balanced, and to ensure that there were processes and governance structures that would be robust enough to mitigate against these risks. It underscored the need to establish a proactive and inclusive governance approach that seeks and welcomes input about ethical biases and offers a fair and transparent path to resolution. This is particularly needed in an endeavor such as

KnowRisk, where the technology is likely to be ahead of the regulatory bodies' awareness and ability to protect its citizens from possible adverse effects.

One particular mechanism that reflectively worked well was the appointment of a project manager for the consortium who took ethics in their stride. In this case, the lead company was Sweetbridge, and they established a role specifically for managing the KnowRisk consortium - whilst this was outside of the recommendations in the roadmap, and was formally included in the consortium building agreement, this is a notable way to ensure ethics is being considered effectively. This role ensured that all different stakeholders within the consortium were being listened to, included, and encouraged throughout the ethics work. Furthermore, this consortium project manager was able to adeptly identify when ethics needed more attention and how to mitigate against the risk of individual companies feeling left behind. Interestingly, throughout the literature of consortia governance, there is no established best practice on how to delegate and assign 'leads' of projects. Depending on the objectives of the consortium, different structures might work best. For example, consortia whose main priority is to be as democratic as possible may want to have rotating heads, whereas consortia who hope to make decisions quickly and more effectively may look to have a smaller steering group or individual managing decisions.<sup>13</sup> Conversely, through the experience of the KnowRisk project, having an established point of contact for the ethics workstream was useful for effective communication, governance and inclusion of parties.

The importance of this overarching principle is not to be underestimated; strong relationship building and bonding from the outset of the consortium were critical to the success of the consortium. This sense of interdependency and shared vision across consortium members was fostered through "brown bag lunches" whereby members could convene every couple of months over (virtual) lunches to discuss the project and updates in an informal way. The process of having consistent ethics workshops, with independent advisors, was also described as being a fruitful mechanism for bonding, interaction and ensuring members were developing a uniform culture across the consortium. This was also important for Industria Tech - the blockchain company joined the consortium later than the other partners, and they managed to meaningfully engage with the ethics workstream as a result of these different consortium engagements. Their enthusiasm and willingness to learn from the ethics experts was instrumental to them getting up to speed quickly. Consequently, these mechanisms can help when bringing in new partners and ensuring ethics continues to be instilled across partners.

### **Promote diversity, equality, and inclusion in consortium governance**

New technologies can change existing power hierarchies, and it is important to ensure that the governance within the KnowRisk platform fairly manages power between stakeholders and prevents powerful groups from co-opting the platform to further their own interests. As a small starting point, the consortium is developing a statement that will be made publicly available on how to engage any stakeholder, to ensure that any one party is not using the data and platform to further their business interests unfairly.

---

<sup>13</sup> [https://gh.bmj.com/content/4/Suppl\\_8/e001450](https://gh.bmj.com/content/4/Suppl_8/e001450)

In addition, the consortium has considered the possibility of legally creating a single entity/joint venture agreement, to help align organisations in a way in which there can be a collective article of association. Having said this, the ethics advisors stressed that “culture eats strategy for breakfast”, and potentially the most important aspects to keep ethics alive are mechanisms to ensure it continues to be embedded in the culture. For example, a few suggestions were 1) around creating hard criteria for new board members that would ensure they met some standard of technology conscientiousness (etc.), 2) run workshops and role plays when setting up future consortia - this could be translated into governance and economic guidance.

### **Consider the business model**

By the end of the project, members felt that ethics was a top priority for the business, and continue to want to ensure that ethics is not pushed down the pecking order. Members expressed the desire to keep ethics alive, to be embedded into the product, and to not simply become ‘ethics washing’ or another corporate social responsibility exercise, whereby it is merely for window dressing, and doesn’t do enough to protect against potential harms to users or other groups. To this end, the consortium have considered legacy planning of the KnowRisk project to be an important one: how can it be certain that there is ‘institutional memory’ embedded into the consortium, no matter if project partners are eventually replaced by other ones? There are also broader questions around what happens to data in the event of corporate death, and the legal frameworks available to prevent subsequent takeovers.

Some practical steps taken have been in identifying who the business model impacts: Intelligent.AI’s platform is focused on Commercial Property Insurance - B2B model - and therefore has less ethical impact on citizens and individuals, but instead could present ethical biases against SME businesses, and favour large businesses. As a result, they are building analytics capabilities into the platform to monitor this.

On a broader level, areas of future research and exploration may include delving into emerging venture capitalists that are structured in nontraditional methods and do not encourage companies towards an exit. It is useful to frame innovation building around the incentives imposed due to financial support in VCs and investors, as this has historically created a moral duty on founders to sell their businesses and prioritise extremely high valuations. Second generation business models that may focus on other variables such as revenue share opposed to company valuations may promote the creation of slower but more sustainable and conscientious businesses.

## Section VII: Feedback and how to improve ethics as a service

“The quality of the input from Dr Laura James and Professor Burkhard Schafer was extremely high, and the framework and roadmap are highly thoughtful and detailed, and have become essential reference documents for KnowRisk.”

**Jason Cresswell - KnowRisk Consortium Manager**

Reflecting on the process, it was felt that the quality of input from both ethics advisors was extremely high, highly thoughtful and their documents (roadmap and ethics framework) have become reference documents for the KnowRisk project. Across the board, consortium members felt that the ethics workstream did everything it promised from the outset.

In terms of constructive feedback, from a user experience perspective, it was felt that this could be improved for 1) future consortia, or 2) for commercialisations of this as a product/service. In particular, the layout of the ethics service, which was predominantly documented in long form in documents, made the barrier to entry quite high, as it was quite difficult to ensure everyone would read, engage and evaluate the documents. This format was perhaps more academic in nature, and gave the feeling of being ‘taught’ ethics. As the KnowRisk consortium had many members who are highly experienced professionals, this approach was potentially not the most fruitful way of engaging them. In addition, the group activities were, initially, all large online workshops. This was done to create the feeling of a single entity as a consortium - however, it meant that this larger group setting meant not everyone wanted to contribute openly.

This problem was identified during the course of the project. The KnowRisk consortium manager worked closely with the Digital Catapult team to increase engagement in the ethics activities. The KnowRisk consortium manager and Digital Catapult team developed a risk identification activity,<sup>14</sup> to be held during a ‘brown bag lunch’ within individual consortium members. This meant that internal teams were holistically engaging with the most relevant questions for their specific technologies in a critical way. These conversations enabled teams to truly understand the ethical issues covered in the framework and beyond. To further this understanding, the Digital Catapult team organised an additional “Office Hours” session with the two independent ethics experts to discuss the issues that came to light during their internal ethics meetings. The Office Hours were split into two groups: node risk and flow risk, and this forum allowed the teams to dig even deeper into ethical considerations. These sessions took people outside of their comfort zone and enabled them to understand the complexity and competing tradeoffs involved in lots of different decisions.

Ultimately, while the content offered and produced was top calibre, the way it is presented has potential to be optimised to ensure it reaches the highest levels of engagement. It is likely that the most adept way of achieving this also depends on consortium to consortium; potentially, academic consortia might have seen a preference for the initial framing. In any

---

<sup>14</sup> See the appendix for more information.

case, the design and user experience of 'ethics as a service' has a way to go, and every trial and iteration of its experimentation will continue to reap even stronger results.

For future industry led consortia, the approach of having individual teams run through an ethics identification activity, and then having a small number of participants in "office hours" with the ethics experts to discuss the issues identified and to dig deeper, is recommended as a complement to the larger consortia ethics sessions. It seems both are essential to making ethics have an impact at the local level (i.e. within teams) as well as at the consortium level.

## Closing remarks

It is evident that cultivating ethics within the KnowRisk consortium has been immensely beneficial, both in culture across teams, quality in product design and development choices, and in its approach to the platform's potential impact on wider society. Responsible innovation doesn't happen overnight, but it is a worthwhile endeavour. It is especially important to have these conversations in the early stages of projects, as this can set the precedent for how entire projects are governed and technologies developed.



## *Appendix*

[Revised ethics framework for KnowRisk](#)

[Identification of ethical risks worksheet](#)