

Appendix A

Overview of cases

Data crawlers

- *Mercury* is a data platform that aims to build privacy-safe insights from 2 billion human-generated daily interactions. The company has linked up with social networks by parsing its public data, such as public posts and tweets and growing advertising across social networks, where the primary focus is on brands, businesses, financial markets, news, and public opinion. Value is created by data collection, data set aggregation, anonymisation, and classification and reporting for a private need. Value is delivered using a platform, and it captures value through partnerships with commercial B2B customers.
- *Venus* goes beyond social listening and analyses billions of social media and marketing websites by filtering the noise with human-powered artificial intelligence. As a result, Venus creates value by providing real-time insights about leads, brand reputation, dissatisfied customers, and trending sentiment (positive, negative, or neutral). It enables early warning alerts to help firms prevent a crisis from impacting a brand and improves customer satisfaction. The value is delivered using a platform and is captured through contracts.
- *Earth*, a data platform for pet owners, enables users to upload and share decorated pictures of pets for fun and enjoyment. The value is created by providing games and arranging competitions for cutest pet determined by voting, such that pet owners can receive prizes such as gift cards for their accomplishments. In addition, animal owners are invited to participate in surveys about their pets so that these data can be used to target ads. Value is delivered through an app and captured by standard contracts with

animal nutrition companies.

Data marketplaces

- *Mars* is a data platform that aims to disrupt the conventional meat marketplaces by enabling data-driven productivity improvements across the whole value chain of livestock production (e.g., higher-quality products with less waste, higher returns for farmers) and by providing transparency for cattle selling. Mars provides every farmer in the UK with a free app to record private individual animal data (e.g., measurements, vet records). Value is created by enabling data-driven improvements for farming (e.g., optimal date of sale for peak profit) and providing a digital profile for each animal that can be matched with a contract and traded online. Value is captured by taking a small commission of 1%–3% depending on the trade when the cattle is marketed. Value is delivered through an app and captured by the transaction fees of the marketplace.
- *Jupiter* is a data platform to organise and share private health-related records from wearables, lab reports, and physician and healthcare organization visits. Value is created by better managing private medical and wellness records and targeted sharing with physicians or personal trainers so that the data contributors are more aware of their health, including changes and options to improve it. In addition, the data contributors can receive rewards by participating in research and clinical trials, campaigns, and challenges to help advance healthcare. Value is delivered through an app and captured by standard contracts with healthcare organizations.
- *Saturn* is a data platform that helps individuals generate and sell their private data after DNA tests and uses these data to develop personalised products (e.g., drugs). The value is created by providing a marketplace where scientific community members, pharmaceutical firms, and fast-moving consumer good companies (data buyers) can find

sellers. Saturn provides a free genetic test and is funded by owners using blockchain. Value is delivered using a decentralised blockchain platform and is captured by cryptocurrency payment.

Data aggregators

- *Neptune* is a data platform aimed to accelerate drug discovery and development by providing access to the full set of ‘omics’ (e.g., genomics, proteomics, metabolomics, metagenomics, phenomics) data generated by the scientific community so far, which includes millions of data samples in multiple data types, and replacing the manual work of bioinformaticians to standardise them. Value is created by clinical trials' success rate prediction and is captured by standard contracts with pharmaceutical companies.
- *Uranus* is a data platform to enable direct communication between smart cities and their citizens using road infrastructure, including bicycles and scooters. To collect data, Uranus sells IoT-enabled lights that generate data on the surface and cyclists' behaviour (e.g., braking data), which can help data scientists predict accidents and collision deaths, inform road upgrading, and improve cycling safety (value creation). Therefore, the community of cyclists and community of data scientists co-influence transport planning decision makers to improve cycling infrastructure and conditions for cycling.
- *Pluto* is a data platform to raise global IT security by tracking virus activity. It embraces a community of engineers who share data on viral activity on the internet. First, Pluto creates value by aggregating, visualising, and disseminating data, thereby increasing global awareness of viral outbreaks. Second, it provides communication and ranking tools for community members, thus capturing the social capital of experienced engineers.

Data disseminators

- *Phoebus* is a data platform to lower the barrier to the practical use of big data in agricultural research and disseminate data-driven outcomes for smaller farms. The value is created via sharing, visualising, and collaboratively analysing agricultural data provided by community members. Phoebus enables secondary services to secure the transfer and storage of computer program codes and data files via its data platform, to hold the trust and deliver data-related global public goods. Its purpose is to find new ways to reduce hunger and poverty, contributing to the United Nations' Sustainable Development Goals. Value is delivered through the platform and captured by its contribution to sustainability.
- *Lune* is a data platform to prevent human trafficking by facilitating data sharing. It partners with financial institutions, nongovernmental organizations (NGOs), law enforcement, and government agencies to collect data to the shared 'data lake' and prepare it for natural language recognition tools. Lune provides value by serving as a single source of truth on where trafficked victims come from, how they get where they are, and in which regions and industries they are most likely to end up working. Lune enables companies, NGOs, and individuals to identify trafficking patterns, networks, and hotspots across all industries and sectors. Financial organisations, for example, can use this information to identify where people are financially benefiting from this activity and then design processes that make it riskier for them to do so. The value is captured by attention to the firm that provided the grant.
- *Sun* is a data platform that aims to reduce the volume of antimicrobials in producing pork, chicken, and eggs by identifying how much antibiotic has been used, what specific drugs are used, for how long, and what animals consume. It owns the largest UK animal

health database of 92 million transactions and activities, independent of any major food producer or manufacturer. Value is created by measuring antimicrobial consumption per animal production at farm, batch, and flock levels, delivered to veterinary pharmaceutical companies and large multinational food producers. Value is delivered using a platform and captured by standard contracts with these large firms.