THE UNIVERSITY OF SYDNEY

INFO5992 Understanding IT Innovations

Tutorial 05 Summary

Crowdsourcing

Q1. Briefly describe the crowdsourcing tasks and goals of the provided business examples in your own words.

Amazon Mechanical Turk: Amazon Mechanical Turk (MTurk) is a crowdsourcing platform that connects businesses and researchers with a diverse group of individuals, known as "workers," who perform small online tasks in exchange for payment.

- **Tasks**: Virtual tasks that require human intelligence, such as identifying objects in a picture or video, transcribing audio recordings, and researching data specifics.
- **Goal**: To provide businesses with a varied, on-demand, scalable workforce through the Mechanical Turk service, and to allow workers to choose from thousands of tasks to complete at their convenience.

Kaggle: Kaggle is a crowdsourcing platform that specializes in data science and machine learning tasks. The platform connects businesses and organizations with a community of data scientists and machine learning experts who compete to provide the best solutions to specific problems or challenges.

- **Tasks**: Data science, machine learning, and predictive analytics problems.
- **Goal**: To give developers and data scientists a platform to interact and compete in solving real-life problems, and to host machine learning contests with monetary prizes. Kaggle also provides a platform for users to upload their data and for developers to make a solution and share it.

iStockPhoto: iStockPhoto is a crowdsourcing platform that specializes in stock photography and other forms of digital media. The platform connects businesses and individuals with a global community of photographers, videographers, and illustrators who create and sell digital media assets.

- **Tasks**: Providing an online image and video library with ready-to-use photos, vector graphics, and video clips.
- **Goal**: To collect a diverse set of user-generated stock content that can be used for a variety of purposes, and to enable users to upload their own images and license them to other users who require them for projects.

Comma AI: Comma AI is a crowdsourcing platform that specializes in developing open-source autonomous driving technology. The platform connects businesses and individuals with a community of developers, data scientists, and engineers who collaborate to create and improve autonomous driving software and hardware.

- **Tasks**: Maintain and provide the open source for the semi-automated driving system. It crowdsourced the development of the system, including curating the AI-developing dataset and enhancing the source code.
- **Goal**: Provide a semi-automated driving system to assist drivers and allow the users can modify their existing cars.

Q2. Explain with reasons and determine which type of crowdsourcing each example is.

Amazon Mechanical Turk: Distributed Human Intelligence Tasking, where workers are asked to complete specific tasks that require human intelligence, such as data labeling, content moderation, and transcription. It breaks down complex tasks into smaller ones and distributes them to many people to complete. The goal is to connect businesses with a scalable workforce.

Kaggle: Broadcast Search. It provides a platform for companies to offer prizes to data scientists who compete to solve real-life problems. The goal is to foster collaboration and competition in the data science community.

iStockPhoto: Peer-Vetted Creative Production. It involves a community of photographers, videographers, and illustrators who create and submit digital media assets that can be licensed and sold to businesses and individuals. It

provides a platform for users to upload their images, which are then organized into categories with searching capabilities and where the photos need to meet with user's taste. The goal is to collect a diverse set of user-generated stock content with high quality and follow the market needs.

Comma AI: Distributed Human Intelligence Tasking \ Broadcast Search. It provides a platform for contributors to label images to improve the accuracy of Comma ai's open-source semi-autonomous driving system. The goal is to assist drivers in improving their visual perception and electromechanical actuator control. The development of this technology requires both software development skills and domain expertise in autonomous driving, making it an example of both peer-vetted creative production and distributed human intelligence tasking. Also, it is partially having a type of broadcast search where the open source system is gathering the solutions of empirical problems (semi-automatic driving) into its source code.

Q3. For each example, discuss how they solve the challenges associated with managing crowd contributors.

Amazon Mechanical Turk:

- MTurk has several features designed to help manage crowd contributors, such as worker qualification tests, performance monitoring tools, and a rating system for workers. These features allow requesters to filter out low-quality workers and identify the best performers, which can help ensure high-quality results and minimize the risk of fraud or unethical behaviour.
- They test and filter the crowd contributors to ensure that they are qualified and have the required skills for the tasks. They also anonymize or obscure the data before sending it to the crowd to protect confidentiality. They carefully outline the process before initiating crowdsourcing and consider possible interdependencies to avoid future coordination problems. They also need to evaluate answers submitted by individual users.

Kaggle:

- Kaggle uses a leaderboard system to manage crowd contributors, where participants can see how their solutions compare to others and track their progress over time. This encourages healthy competition and motivates participants to produce high-quality solutions.
- Kaggle provides clear instructions, offers incentives, monitors progress, provides feedback, and uses quality control measures to manage crowd contributors. The competition prepares the data and provides a detailed description of the problem at hand. They also offer an additional consulting service that can help prepare data and describe the problem in the best possible format.

iStockPhoto:

- iStockPhoto has a community of contributors who can provide feedback and support to each other, which helps to build a sense of community and encourage high-quality work. Additionally, iStockPhoto has a team of reviewers who evaluate submitted assets and provide feedback to contributors, which can help improve the quality of work over time.
- They manage the rights associated with images on behalf of contributors, ensuring that the content is properly licensed and legally available for customer use. They handle all copyright issues and monitor the marketplace to prevent unauthorized use of their contributors' work.

Comma AI:

- Comma AI manages crowd contributors through a combination of collaboration tools and technical standards. The platform has a set of guidelines and technical standards that contributors must follow when developing autonomous driving technology, which helps to ensure consistency and quality across contributions. Additionally, Comma AI has a team of moderators who review and approve contributions, which can help maintain the quality of work over time.
- They use training, quality control, communication, incentives, and community building to manage crowd contributors. They have a growing community that maintains various source code forks of OpenPilot, which consists of experimental features such as stop-light detection.