21

ALTRUISTIC COLLECTIVE INTELLIGENCE FOR THE BETTERMENT OF ARTIFICIAL INTELLIGENCE

Thomas Maillart, Lucia Gomez, Mohanty Sharada, Dipam Chakraborty and Sneha Nanavati

1 Introduction

Despite its transformative economic and social outcomes, artificial intelligence (AI) is faced with several operational, legal, and ethical challenges (Ntoutsi et al., 2020), mainly associated with algorithm robustness (Dietterich, 2019), explainability (Confalonieri et al., 2021), and biases (Osoba et al., 2017). However, as reported by Percia David et al. (2023), AI seems to be still largely in its infancy and may be decades away from becoming a mature technology. Despite concerns and room for improvement, the opportunity for a bright AI future exists. Navigating toward the betterment of AI implies engaging in a critical reflection that builds on the lessons learned and experience of computer revolutions from their start in the late 1960s (Levy, 2010). While computer systems development was mainly driven by closed source and intellectual property (IP)-protected software at the time, an open-source community developed early on. This open-source movement was initially seen as fringe and utopist, as it introduced a new form of IP, which would primarily require sharing software code and crediting code authors[1] in a non-exclusive non-rival economic regime, instead of extracting profit from exclusive, rival goods (Tirole & Lerner, 2002). The open-source movement led to the disruption of large parts of the software industry, from operating systems (Moody, 2009), to the Internet (Zittrain, 2009), to cryptography (Landau, 2022), to the World Wide Web (Benkler, 2011), and eventually impregnating the world of commercial software (Fitzgerald, 2006). The dominance of open-source software has culminated since Microsoft acquired GitHub, the prime open-source online development platform, in 2018. Reflecting on how past technology development has been well managed and how it has failed in its social outcomes is informative to establish a sustainable development path for AI, under economic, social, and ethical constraints.

Inspired by the open-source software (OSS) movement, we posit that two interacting mechanisms are key to the sustainable betterment of AI. First, collective intelligence (CI), embodied by peer production – composed of task self-selection, peer review, and transparency (Benkler, 2002) – plays the role of democratized control over the development of AI systems, while also boosting innovation, through decentralized, modular, and creatively destructive dynamics (Maillart et al., 2008). Second, the open-source movement has allowed a strong culture of altruism, born

DOI: 10.4324/9781003468615-25

from the imperative of collective action to overcome significant challenges in complex environments (Ostrom, 1990). It also grew largely with the recognition of software programming as a form of art (Bonaccorsi & Rossi, 2004), whose practice stems largely from intrinsic motivation (Ryan & Deci, 2000; von Krogh et al., 2012).

In this chapter, we critically reflect on the development of AI by drawing from the concrete example of AIcrowd,[2] an AI startup that has heavily bet on peer production, intrinsic motivation of a sense of community, with emphasis on fun and a strong drive for social good. Namely, AIcrowd proposes challenges offered by its paying customers or for free by not-for-profit organizations for its 68k+ community members to tackle. During challenges, contributors can submit and test for performance as many AI models as they wish. The winners receive a monetary or in-kind prize at the end of the competition period. Participants may also be offered to be co-authors of research papers to be posted to, e.g., *arXiv.org* or submitted to a computer science conference, such as NeurIPS, the famous AI scientific conference (competition track). If accepted, the lead authors are invited to attend the conference with all fees covered. AIcrowd is therefore walking a thin line combining extrinsic motivation and competition (i.e., competition for a prize and potential royalty proceedings from paying use of AI models down the road), intrinsic motivation (i.e., sense of achievement in a fun and social online environment), cooperation, and a sense of purpose. The latter sense of purpose is powerful with challenges organized free of charge for not-for-profit organizations (e.g., International Telecommunication Union – AI for Good or the United Nations). By investigating the AIcrowd community as a heterogeneous set of engaged individuals fueled by various incentives, we investigate how the selfless philanthropic side of people acting collectively, through their cooperation and competition at once, brings more ethical value to AI, enhancing cooperation, utility, robustness, and transparency.

2 Background

To envision how altruistic collective intelligence could benefit from the development of AI, we consider how the open-source movement has durably shaped the development of software, which is the most important precursor technology of AI together with abundant data and the development of specialized hardware, such as graphical processing units (GPUs) chips.

2.1 How altruistic collective intelligence has shaped the digital revolution

The open-source "*hacker*" movement (Levy, 2010) and the peer production approaches (Benkler, 2002), characterized by the collaborative efforts of individuals working voluntarily toward a common goal, have significantly contributed through history to alleviating operational, organizational, and ethical challenges in software development. Operationally, peer production facilitated the rapid identification and resolution of bugs and the enhancement of software features through the diverse expertise of contributors, leading to more robust, reliable, and adaptive software (Raymond, 1999; Maillart et al., 2017). Organizationally, it democratized key development processes, breaking down barriers to entry and allowing for a more inclusive group of participants, which in turn fostered innovation and accelerated development cycles (Bosu et al., 2017). One overarching example of the influence of the open-source movement on the development of software development practices is GIT (Loeliger & McCullough, 2012), a distributed control version system invented by Linus Torvald in 2005, that has become the technical backend of widely used social coding platforms, like GitHub and GitLab.

Ethically, the open-source approach has encouraged transparency and accountability, as the open review process ensured that code was scrutinized by a broad community. This openness helped in identifying and mitigating biases or unethical uses of software, aligning development practices more closely with societal norms and values (Coleman, 2012; von Krogh et al., 2012).

The open-source movement has indelibly shaped the software and hardware development landscape, embedding principles of transparency, collaboration, and accessibility deeply within the fabric of technological innovation. It marks a pivotal shift from the exclusivity of proprietary systems toward a more democratized approach to technology creation and dissemination. The genesis of the open-source movement can be linked to the collaborative ethos among early computer scientists and hobbyists who believed in sharing software openly as a means to foster innovation and solve problems more efficiently. In his book *Hackers: Heroes of the Computer Revolution* (2010) on the history of the hacker culture, Steven Levy highlighted the commitments of this burgeoning community to openness, information, and sharing. This period laid the groundwork for developing major open-source projects that would later revolutionize the technology sector.

The influence of open-source principles became more pronounced with the development of foundational operating systems, notably the Linux kernel, spearheaded by Linus Torvalds in 1991. Moody's exploration (2009) of open-source operating systems illustrates how this model disrupted the traditional software development paradigm, enabling a global community of developers to contribute to and improve existing codebases. This collaborative approach not only accelerated innovation but also ensured that software could be modular and adaptive (Maillart et al., 2008), as such, more secure, reliable, and adaptable to the needs of diverse users.

The expansion of the Internet and the World Wide Web further exemplified the power of open-source methodologies. Jonathan Zittrain (2009) and Barbara van Schewick (2012) have discussed how Internet's open architecture facilitated an unprecedented level of innovation and creativity, allowing individuals and small teams to create impactful technologies without the need for substantial resources. The invention of the World Wide Web by Tim Berners-Lee at CERN, who made this technology available on a royalty-free basis, epitomizes the ethos of the open-source movement, democratizing access to information and enabling the explosive growth of online content and services.

Cryptography, as explored by Susan Landau (2022), is another area where open-source principles have been instrumental. The move toward open cryptographic standards and the public sharing of encryption algorithms have significantly enhanced security and privacy in the digital age, underscoring the movement's role in building trust and safeguarding freedoms online.

The culmination of the open-source movement's integration into the commercial sector was symbolized by Microsoft's acquisition of GitHub in 2018. Once seen as antithetical to the open-source *ethos*, major corporations have now largely embraced open development practices, recognizing the value of community-driven innovation (yet without relinquishing proprietary, closed source code). The GitHub acquisition by Microsoft, as discussed by Brian Fitzgerald (2006), marks a significant acknowledgment of the impact of the open-source model on commercial software development, highlighting a shift toward more open, collaborative, and transparent practices associated with increased competitiveness for firms (Nagle, 2018) and even for nations (Nagle, 2019).

2.2 Altruistic collective intelligence is a form of philanthropy

At the root of the open-source movement is collective intelligence (CI). CI refers to the shared knowledge, expertise, and problem-solving capabilities of a group or community of individuals (Malone, 2019). It is the idea that the collective wisdom of a group and their collective decision-making can bring additional performance to that of any single member within the group: CI emerges when people collaborate, share information, and pool their insights and abilities to tackle complex problems, make decisions, or create innovative solutions. Deeper, CI is best predicted by social interactions (Kim et al., 2012) and the capacity to sense the emotional states of others through non-verbal social cues (Woolley et al., 2010). The CI concept has become particularly relevant in the digital age, where technology and connectivity enable large and diverse groups to collaborate and generate insights and solutions collectively (Benkler, 2011).

When predominantly relying on intrinsic motivation (Ryan & Deci, 2000), CI can be seen as a form of philanthropy, albeit in a non-traditional sense. While philanthropy typically involves the individual donation of financial resources or volunteering time to support charitable causes directly through the provision of individual skills, CI offers a different kind of contribution, which is collective, integrative, and, most importantly, involves the mobilization of empathy (Woolley et al., 2010). Hence, the philanthropic contribution lies primarily in how it proceeds and delivers value collectively, by leveraging people's inner social capabilities (Dunbar, 1998). *In fine*, it enables collective action to overcome significant challenges and strive in adverse environments (Ostrom, 1990).

CI involves harnessing the wisdom, knowledge, and expertise of a diverse group of individuals to address complex problems, make informed decisions, or create innovative solutions (Hong & Page, 2004). In essence, CI is a form of giving back to society through the productive collision and integration of intellectual capital (Engel & Malone, 2018). By collaborating and pooling their collective knowledge, people can collectively benefit others by solving challenges, advancing research, or improving decision-making in areas such as science, technology, and governance (Fink, 2018). CI is even thought to provide answers to how humankind should consider tackling global catastrophic risks (Yang & Sandberg, 2023). In this sense, the act of contributing individual insights and expertise to collective efforts can be seen as a valuable and altruistic form of philanthropy, one that goes beyond financial donations and embodies the spirit of communal support for the greater good. As for contributing, collaborative giving elicits similar intrinsic rewards, and the "whole is more than the sum" financial contributions (Proulx et al., 2023), while also possibly producing the "whole is more than the sum" (Sornette et al., 2014) through peer production (Benkler, 2002).

2.3 Outstanding challenges in AI and how altruistic collective intelligence can help with a dose of competition

As we stand on the brink of a fundamental reshaping of many facets of human life by AI, the lessons from the open-source movement are more pertinent than ever. The principles of transparency, collaboration, and ethical responsibility that have driven the open-source movement can, and already largely, serve as guiding lights for the development of AI technologies. By fostering an open AI ecosystem, we can encourage a broad and diverse community of developers, ethicists, and users to contribute their perspectives and expertise, thereby ensuring that AI technologies are not only advanced but are developed in a manner that is socially responsible, inclusive, and aligned with human values.

Moreover, the OSS model can address some of the most pressing concerns in AI development, including biases, transparency, and accountability, in the same way OSS has helped alleviate similar problems for previous information technology developments. By making AI algorithms and datasets publicly available, the community can facilitate scrutiny, peer review, and iterative improvement, ensuring that AI systems are fair, reliable, and understandable. This collaborative approach to AI development has the potential to democratize AI innovation, making it accessible to a wider range of stakeholders and enabling solutions that are tailored to a variety of social, economic, and environmental challenges. However, in AI development, benchmarks serve as computational Olympic arenas, where algorithms and pipelines compete for improvement toward a progressively optimal solution or fork through radical innovations. These standardized testing grounds for AI evaluation constitute unique tools for combining cooperation and competition, as all developers can observe and learn from each other's solutions to create the next best one. Potentially serving as one of the major catalyzers of CI within AI, benchmarks are fundamental pieces for a future open and democratic AI landscape.

The journey from the early advocacy for open computing environments to the present-day ubiquity of open-source methodologies underscores a fundamental change in how digital technologies have been developed, distributed, and perceived. The transformative impact of open-source principles across various technological milestones offers a compelling narrative that not only charts the evolution of technology but also sets a precedent for the development of artificial intelligence (AI) in a sustainable, ethical, and inclusive manner.

3 Theoretical framework: altruistic collective intelligence contributing to better AI

Altruistic CI is pivotal in advancing and improving AI, as witnessed by the fast advancement of worldwide open-source communities and hubs such as *HuggingFace*,[3] providing new AI tools and evaluation benchmarks daily (such as their *open_llm_leaderboard*). It brings together individuals from diverse backgrounds, approaches, and focus areas, each of which contributes unique tools. This diversity of thought helps AI developers and researchers consider a wide range of perspectives and opens the doors for a fruitful peer-supported exchange, leading to more well-rounded and ethically sound AI systems.

With the increasing complexity of AI systems, ethical considerations have become paramount. Altruistic CI fosters discussions and debates around AI ethics, helping to establish guidelines and best practices that ensure AI technologies are developed with fairness, transparency, and accountability in mind. An exemplar case is the safety and ethics AI leaderboard hosted at *HuggingFace*, setting standards developed and continuously improved by an open community for evaluating AI models (*llm-trustworthy-leaderboard*). CI initiatives can offer valuable feedback on AI systems, helping developers identify weaknesses, vulnerabilities, and areas for improvement. This iterative process is crucial for enhancing the robustness and security of AI technologies.

Besides model development and testing, AI systems heavily rely on data. By engaging in altruistic data sharing, data annotation efforts, and collaborative-competitive development of AI algorithms, CI can help improve the quality and diversity of training data to mitigate biases, but also improve the balance in representing minorities and rare data sources, for example. This collective data sharing and pruning efforts, in turn, lead ultimately to AI systems that are more accurate, unbiased, and reliable. In sum, altruistic CI is a source of diverse and rich data, AI models, and ethics testing.

Further, as the existence of AIcrowd shows, CI can also contribute to the accelerated development and the betterment of machine learning algorithms and neural networks. Nevertheless, the origins of CI performance in AI algorithm development have remained unclear.

We posit that performance is largely due to collective "*trial-and-error*" by teams. Our hypothesis is that ***team size, number of submissions***, and incidentally ***waiting time between submissions*** are important, but the ***diversity of submissions*** by the same team matters most. The number and diversity of submissions by a team reflect the collective capacity to take risks in accepting that some submissions will not overperform the current benchmark of (i) own team submissions as well as

(ii) submissions by other teams. We contend that this capacity stems from using foraging-like *explore* and *exploit* human search algorithms (Wilson et al., 2014).

As we dive into the data of *the AIcrowd Food Recognition Challenge* (see Section 4), we focus on the structure of ***cooperation*** and ***competition*** in communities and how the teams achieve performance during the challenges.

4 AIcrowd: an exemplar case study on altruistic CI for AI

4.1 AIcrowd: democratizing access to AI challenges and harnessing collective intelligence

AIcrowd is a pioneering platform at the intersection of data sharing, AI, and CI. As a business, AIcrowd operates as a hub that connects problem-solvers, data scientists, and machine learning enthusiasts with organizations seeking innovative solutions to complex AI challenges. Since it started, AIcrowd has hosted over 500 challenges and competitions in a wide range of AI sub-disciplines. Participants from around the world team up, cooperate, and submit their machine learning models to solve real-world problems.

AIcrowd challenges cover diverse domains, from computer vision and natural language processing to robotics and healthcare. The business model thrives on creating a collaborative ecosystem that allows organizations to tap into the CI of a global community of AI enthusiasts and researchers. Through these challenges, AIcrowd facilitates knowledge sharing, fosters innovation, and accelerates AI advancements while offering organizations access to cutting-edge solutions and talents. This unique business model positions AIcrowd as a catalyst for collective problem-solving in the machine learning domain.

AIcrowd faces competition primarily from platforms like Kaggle[4] – known for its broad range of data science and machine learning competitions, Topcoder[5] – which connects businesses with global developers and data scientists, DrivenData[6] – specializing in data science challenges with social impact, CrowdAI – focusing on computer vision competitions,[7] or Zindi – catering to African-specific AI challenges.[8] These platforms vary in their focus, community, and types of challenges they offer, creating a dynamic landscape for AI competitions and CI in the field.

AIcrowd sets itself apart from competitors by emphasizing fostering a collaborative AI community and diverse problem domains. Unlike competitors, AIcrowd encourages open knowledge sharing and collaborative learning, creating a strong sense of CI among participants. Its challenges cover a wide array of AI fields, addressing real-world problems faced by organizations and making it practical and industry-oriented. AIcrowd also actively collaborates with academic institutions and research organizations, promoting and leading cutting-edge AI research that emerges from the community breakthrough offered solutions. These aspects distinguish AIcrowd as a platform not only for innovation but also for nurturing a culture of open learning and sharing, and collective problem-solving in the AI community, making it unique in the field of AI challenge business.

4.2 The AIcrowd Food Recognition Challenge

In the landscape of modern AI-driven challenges, the *AIcrowd Food Recognition Challenge* stands out as a compelling endeavor[9] with implications spanning different sustainability matters such as SDGs 3 (good health and well-being) or 2 (zero hunger). This challenge revolves around the intricate task of recognizing food items from images, a capability that could empower individuals to effortlessly monitor their dietary habits by merely photographing their meals, providing a powerful tool for personal health management and nutrition tracking (Mohanty et al., 2022). Beyond personal applications, the challenge's focus on food recognition holds substantial medical relevance, addressing a long-standing methodological need in research studies. Traditionally, such studies relied on low-scale and imprecise food frequency questionnaires. This challenge harnesses the power of CI and AI to provide improved and scalable detection solutions.

A core element of CI in this challenge is the collaborative annotation process. The dataset used in this competition is a collection of images and a rich repository meticulously annotated with segmentation, class-belonging, and weight/volume estimation of individual food items.[10] This annotation effort harnesses the collective wisdom of contributors, ensuring the quality of the dataset and enhancing its utility (Mohanty et al., 2022). The Food Recognition Challenge is by no means unique in that regard. Most recent advances in AI have been possible through highly annotated datasets and evaluation benchmarks, using some form of crowdsourcing for collaborative building and benefiting from a standardized and open ground for model evaluation and optimization. Cases such as the transformative ImageNet challenge illustrate the power of CI for AI, an iconic benchmark competition that has significantly advanced image recognition datasets, research, and AI solutions widely used today, such as the AlexNet, VGG, and ResNet (Deng et al., 2009).

Beyond annotation, the Food Recognition Challenge encourages the development of new machine learning models. Participants are invited to submit their innovative algorithms and approaches, creating a forum for collective problem-solving. By pooling together diverse talents and perspectives in both the annotation and AI development parts, the Food Recognition Challenge exemplifies the principles of altruistic CI. The design of the challenge is an example of best practices in the pursuit of solutions for good and for the resolution of intricate problems.

The inner mechanisms of how performance emerges from *coopetition*, i.e., a subtle equilibrium between collaboration and competition, have remained unclear. While coopetition enjoys some popularity in the context of industrial organization (Bouncken et al., 2015), it is also applicable to collaborative communities (Gulley & Lakhani, 2010). Here, we investigate the fine-grained origins of performance in developing AI algorithms in the context of the Food Recognition Challenge.

4.3 Mapping collective intelligence dynamics of the AIcrowd Food Recognition Challenge

4.3.1 Data and method

To understand how team dynamics evolved throughout the Food Recognition challenge (2021–2022), we studied anonymized data on timestamped[11] solution submissions annotated by:

1. Round type (challenge versus benchmark)[12];
2. Round identification (ID) number (1–4, 1–2)[13];
3. Anonymized participant ID;
4. Anonymized team ID;
5. Solution precision score (proportion of true positive predictions among all positive predictions); and
6. Submission latency (in days, zero being the start of each challenge round).

To comprehensively visualize and analyze team activity across the entire challenge (challenge rounds and benchmark rounds), we structured this data in a bipartite network[14] (see Figure 21.1) capturing the relationships between teams (white dots) and submissions (dark dots), weighted by

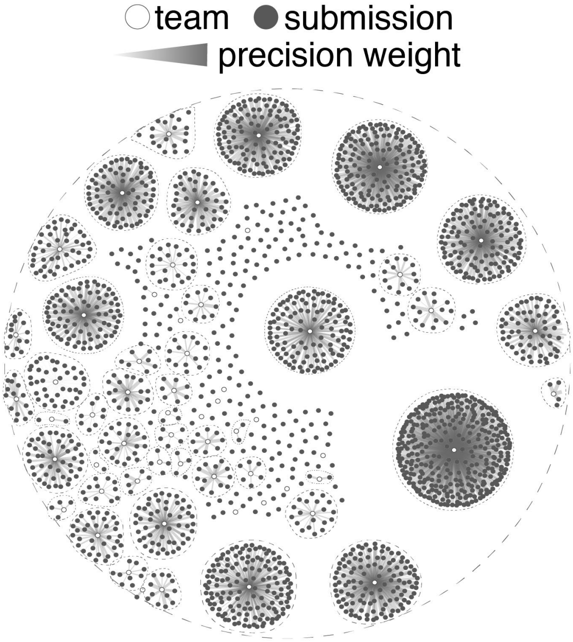


Figure 21.1 Bipartite network of relationships between teams and submissions as weighted by precision scoring.

the precision of the solution (gray lines), as systematically evaluated by AIcrowd. The temporal evolution of the team submission was obtained by slicing the core bipartite network per round (one to four challenge rounds, one to two benchmark rounds; see Figure 21.2).

Upon visual inspection (see Figure 21.1), we find evidence of activity heterogeneity: some teams displayed strong engagement through their submission activity (i.e., a high number of submissions – many dark dots in dense circular clusters). Conversely, other teams sparsely contributed, with a low number of submissions (sparse blue dots, scattered). Additionally, some teams display an additional layer of heterogeneity in their activity patterns: although all circular cluster teams submitted enough to create a dense cluster, some teams submitted significantly more solutions and benchmarked their performance more frequently, while others submitted less solutions (fewer dark dots; smaller clusters).

We propose that team submission density in the network could be a behavioral indicator for potential hubs of CI emergence through *coopetition* in the benchmarking of submissions open to the entire community. In other words, densely clustered teams cooperate between members and compete against themselves and other team submission. The entire network thus interacts (hub), engaging in a CI-generating behavior altogether. Real-time standardized evaluation (benchmarking) could be fundamental for indicating the status of the collective, accessing knowledge and pipelines developed by others, and, therefore, fostering CI coopetition. As all community members are informed in real time about the performance and algorithms of others, cooperation and competition fuel the emergence of innovative and ever-improving solutions.

Next, we considered the teams' submission dynamics during the challenge: Figure 21.2A shows how performance arises for each team and across the four rounds of the challenge.

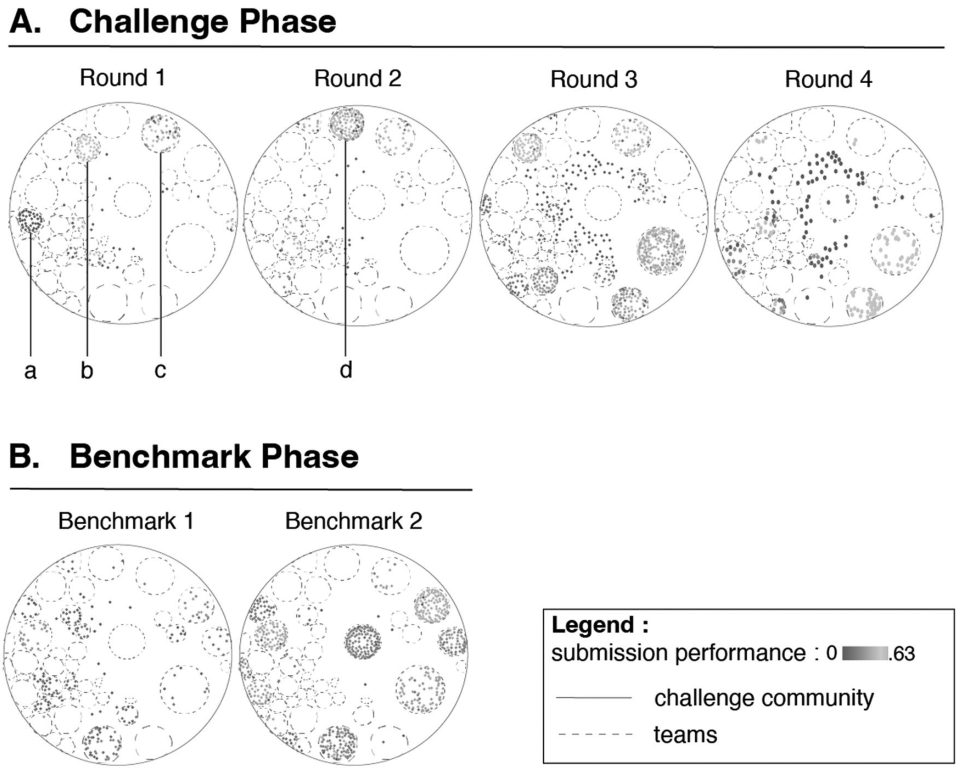
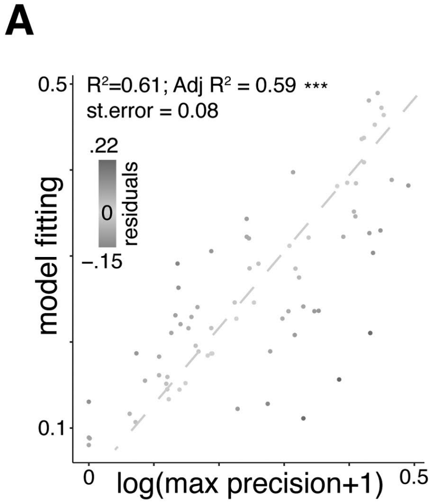


Figure 21.2 Network slice by submission rounds, in challenge and benchmark phases.

Further visual inspection suggests that teams with high and low numbers of submissions at each round can achieve both good and bad performances (high heterogeneity – visualized by color gradient). Additionally, the number of submissions per round does not directly indicate overall performance. For example, *team c* that won the Food Recognition Challenge submitted their winning solution in round 2, while they submitted most solutions during rounds 1 and 3 (61 submissions in round 1, 46 in round 2, 60 in round 3, and 2 in round 4).

Contrary to other teams, such as *a*, *b,* and *d*, participating in single rounds, team *c* showed a protracted engagement across all rounds. For teams participating in several main rounds (i.e., challenge rounds), an overall increase in both solution precision and number of submissions can be observed in consecutive phases (rounds 3–4 display more submissions compared to rounds 1–2). Additionally, during challenge rounds, the performance of submitted solutions within teams shows high variability: while some teams, such as *a* and *b,* submit solutions with a similar level of precision across trials, submissions by other teams, such as *c* and *d,* show a high degree of variability across trials.

Among these, team *c* provided the best solution to the challenge, with a precision of 0.62 (and recall of 0.88),[15] a high score considering the stringent food identification criteria set by AIcrowd for this challenge (c.f., footnote 7 for reading all details at the AIcrowd website). Team *d*, participating in rounds 1 and 2 (2 submissions in round 1, 117 in round 2) and displaying a high-performance variability, was the second-best team, with their top solution having a precision of 0.59 (and recall of 0.82). Conversely, the overall submission precision in the *benchmark* phase was visibly lower than during the *challenge* phase.



B

lm(formula = log(max precision + 1) ~
n submitters + sd precision +
sd diff latency + n submitters)

	est	st.err	t	Pr(>ItI)	VIF
(intercept)	.06	.03	2	.04 *	
n submissions	.0	.0	2.4	.01 *	1.02
sd precision	1.7	.17	9.7	.00 ***	1.05
sd diff latency	.0	.0	.8	.38	1.03
n submitters	.01	.01	.7	.44	1.01

Figure 21.3 Determinants of collective intelligence in AIcrowd Food Recognition Challenge.

This was due to the post-challenge re-definition of both the dataset and task requirements (among which, amount of food types to be identified) toward continuous development of the AI solution after an initial successful competition (challenge rounds 1–4). Despite the decrease in performance (caused by performance metric redefinitions from challenge to benchmark phases), an improvement in submission precision is still observable between phases 1 and 2 of the benchmark. This suggests that award expectation, which was present during the challenge phase but absent during benchmarking, is not the sole driver of continued coopetition.

To systematically understand the recipe of successful CI dynamics, we modeled ***maximum precision*** (dependent variable) by ***team*** and ***round*** across the competition, using a standard linear regression model (see Figure 21.3A–B). As regressors (independent variables), we considered:

1 The ***number of submissions*** by team and round (proxy for how active a team is in testing solutions);
2 The ***number of submitters per team*** (indicating how many different participants are active submitters testing solutions in parallel);
3 The ***standard deviation of the precision*** across submissions (proxy for their trial-and-error progress variation and thus appetite for risky trial and fail);
4 And the ***standard deviation of the waiting time between two submissions*** (metric of how much time it took to develop a given submission).

These independent variables were individually examined for their contributions to the maximum precision. Prior to modeling, we ensured that these predictors did not display collinearity (*VIF* score, Figure 21.3B).

4.3.2 Results

We find that the independent variables of interest here explain a considerable amount of the maximum performance per team and round, with an *adjusted-*R^2 of *0.59* (*non-adjusted-*R^2 of

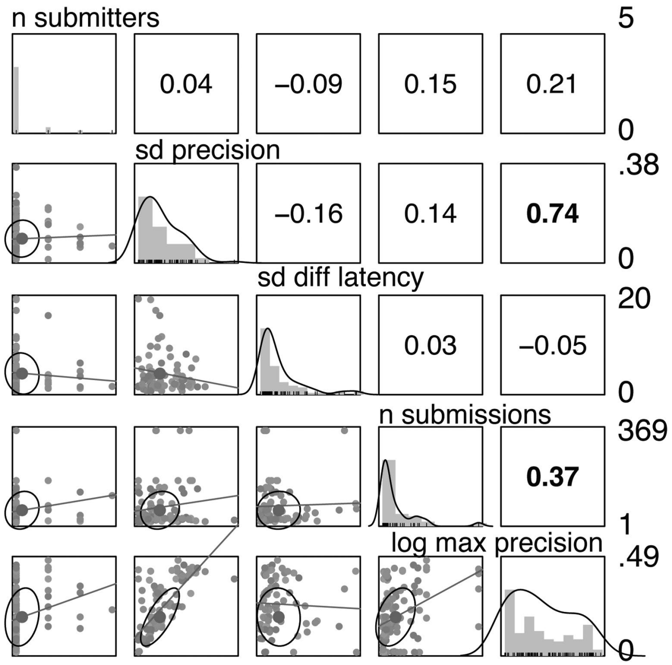


Figure 21.4 Detail on the relation between the variables used for modeling.

0.61), a residual standard error of *0.08*, and statistical significance on modeling ($p < 0.05$) (see Figure 21.3A).[16] Among predictor variables, the number and the heterogeneity of submissions were significant predictors, in particular heterogeneity (standard deviation of solution precision; $p < 0.001$) (see Figure 21.3B). Importantly, as found for the best and second-best performing teams *c* and *d,* risky *"trial-and-error*" exploration of solutions with variable accuracies is an important contributor to successful CI (see Appendix 21.A2). Although this model is not able to explain all observations (high residuals associated with outlier submission: exceptionally performant and under-performant ones), it does capture the overlying trend behind CI success, fueled by a mighty predictive power of the standard deviation of submission precision (see Figures 21.3B and 21.4).

Our results indicate that the more a team tries heterogeneous solutions associated with more performance outcome risks, the more likely it is to obtain a higher maximum precision across trials. Related, however not correlated, the number of submissions per team significantly indicated overall success. Considering the volatility of submission latency, we find that the variability in the time taken for developing solutions can result in an overall successful or unsuccessful competition. Overall, indicators of *coopetition* support CI success.

5 Conclusion

The AIcrowd Food Recognition Challenge analyses reveal non-trivial dynamics of collective intelligence (CI) in action. The varying degrees of team engagement and the performance differences illustrate the intricate balance between competition and collaboration in driving AI advancements. Notably, the significant predictors of success – *variability in submission performance* and *active* *engagement in solution testing* – underscore the efficacy of a "*try-and-error*" sub-mechanism in the broader CI universe. These findings not only highlight the importance of seeking diversity, and hence collective intelligence, to achieve higher precision but also reflect the broader mission of AIcrowd: to democratize AI challenges and harness CI for the collective good.

The mission of AIcrowd serves as a Petri dish for the potential transformation of the AI landscape fueled by the philanthropic contribution of AI developers cooperating and competing at once. By fostering an environment where individuals from varied backgrounds contribute toward common goals, AIcrowd exemplifies how CI can lead to innovative solutions that might emerge neither from isolated nor homogeneous efforts. This approach, rooted in peer production and transparency, leverages the intrinsic motivation of participants, blending competition with cooperation for societal benefit. This, in turn, enables the strength of collective problem-solving through which AI practitioners can tap into a global pool of expertise for solving the intricate types of challenges only this technology has been able to address. The collaborative approach of CI accelerates AI advancements, helps overcome challenges more effectively and democratically, educates the public about AI and its implications, and fosters discussions of how it should be the AI for all.

The implications of such a model extend beyond AIcrowd to inform broader discussions on AI and philanthropy. The altruistic underpinnings of CI, as demonstrated through challenges organized by AIcrowd, offer a blueprint for advancing AI that prioritizes ethical considerations, inclusivity, and fairness. This collaborative innovation model, where success is derived from collective effort and diversity of thought, not only accelerates technological advancements but also ensures that these advancements are aligned with societal needs and ethical standards. As AI continues to evolve, integrating principles of CI and philanthropy into its development can help mitigate biases, enhance transparency, and ensure that AI serves as a force for good, reflecting a shared commitment to improving the human condition.

Notes

1 A wide spectrum of open-source licenses have existed and co-evolved over the years. See Carver (2005) for a full review.
2 https://www.aicrowd.com
3 https://huggingface.co/
4 https://www.kaggle.com/
5 https://www.topcoder.com/
6 https://www.drivendata.org/
7 https://www.crowdai.com/
8 https://zindi.africa/
9 https://www.aicrowd.com/challenges/food-recognition-challenge and https://www.aicrowd.com/challenges/food-recognition-benchmark-2022
10 Segmentation = image annotation delimiting the perimeter of regions of interest (ROI) (food items for training); Class-belonging = Region of Interest (ROI) annotation stating the true food category; weight/volume estimation is self-explanatory.
11 Each submitted solution is time-stamped so that the challenge can be analyzed with temporal resolution.
12 Challenge submissions are registered for the competition phase toward a prized solution whereas benchmark submissions are post-challenge submissions for the permanent improvement of the algorithm outside the competition.
13 IDs 1–4 belong to phases 1–4 of challenge rounds; IDs 1 and 2 belong to corresponding phases of benchmark rounds.
14 Data structure representing two different types of samples (nodes of the network), in this case teams (white) and submissions (dark gray), connected by the performance score (precision) of the submitted solution (gray lines whose thickness, weight, is proportional to its precision).

15 Precision : proportion of true positive predictions among all positive predictions; Recall : Proportion of true positive predictions among all positive and negative predictions.

16 *adjusted-R^2*: Measure of model goodness of fit indicating the proportion of the variance in the independent variable explained by the model when using selected dependent variables for regression. Ranging from 0 to 1, it is calculated as the ratio of the explained sum of squares to the total sum of squares. Adjusted means that this R^2 considers the number of independent variables used and penalizes the addition of unnecessary variables by adjusting for degrees of freedom. *non-adjusted-R^2* does not penalize the addition of unnecessary variables by adjusting for degrees of freedom. The residual standard error is used to measure how well a regression model fits a dataset. In simple terms, it measures the standard deviation of the residuals in a regression model. The *p value* helps determine that the null hypothesis cannot be rejected, i.e., that the model used to represent the data cannot be ruled out statistically.

References

Benkler, Y. (2002). Coase's Penguin, or Linux and "The nature of the firm." *The Yale Law Journal*, *112*(3), 369+. https://doi.org/10.2307/1562247

Benkler, Y. (2011). *The Penguin and the Leviathan: How Cooperation Triumphs over Self-Interest* (1st ed.). Crown Business. http://www.worldcat.org/isbn/0385525761

Bonaccorsi, A., & Rossi, C. (2004). Altruistic individuals, selfish firms? The structure of motivation in Open Source software. *First Monday*, *9*(1). https://doi.org/10.5210/fm.v0i0.1476

Bosu, A., Carver, J. C., Bird, C., Orbeck, J., & Chockley, C. (2017). Process aspects and social dynamics of contemporary code review: Insights from open source development and industrial practice at Microsoft. *IEEE Transactions on Software Engineering*, *43*(1), 56–75. https://doi.org/10.1109/TSE.2016.2576451

Bouncken, R. B., Gast, J., Kraus, S., & Bogers, M. (2015). Coopetition: A systematic review, synthesis, and future research directions. *Review of Managerial Science*, *9*(3), 577–601. https://doi.org/10.1007/s11846-015-0168-6

Carver, B. W. (2005). Share and share alike: Understanding and enforcing open source and free software licenses part I: Law and technology: Subpart IV: Business law: Section A: Notes. *Berkeley Technology Law Journal*, *20*(1: Annual Review), 443–484.

Coleman, E. G. (2012). Coding freedom: The ethics and aesthetics of hacking. In *Coding Freedom*. Princeton University Press. https://doi.org/10.1515/9781400845293

Confalonieri, R., Coba, L., Wagner, B., & Besold, T. R. (2021). A historical perspective of explainable Artificial Intelligence. *WIREs Data Mining and Knowledge Discovery*, *11*(1), e1391. https://doi.org/10.1002/widm.1391

Deng, J., Dong, W., Socher, R., Li, L.-J., Li, K., & Fei-Fei, L. (2009). ImageNet: A large-scale hierarchical image database. *2009 IEEE Conference on Computer Vision and Pattern Recognition*, 248–255. https://doi.org/10.1109/CVPR.2009.5206848

Dietterich, T. G. (2019). Robust artificial intelligence and robust human organizations. *Frontiers of Computer Science*, *13*(1), 1–3. https://doi.org/10.1007/s11704-018-8900-4

Dunbar, R. I. M. (1998). The social brain hypothesis. *Evolutionary Anthropology: Issues, News, and Reviews*, *6*(5), 178–190. https://doi.org/10.1002/(SICI)1520-6505(1998)6:5<178::AID-EVAN5>3.0.CO;2-8

Engel, D., & Malone, T. W. (2018). Integrated information as a metric for group interaction. *PLoS One*, *13*(10), e0205335. https://doi.org/10.1371/journal.pone.0205335

Fink, A. (2018). Bigger data, less wisdom: The need for more inclusive collective intelligence in social service provision. *AI & SOCIETY*, *33*(1), 61–70. https://doi.org/10.1007/s00146-017-0719-2

Fitzgerald, B. (2006). The transformation of open source software. *MIS Quarterly*, *30*(3), 587–598. https://doi.org/10.2307/25148740

Gulley, N., & Lakhani, K. R. (2010). *The Determinants of Individual Performance and Collective Value in Private-Collective Software Innovation* (SSRN Scholarly Paper 1550352). https://doi.org/10.2139/ssrn.1550352

Hong, L., & Page, S. E. (2004). Groups of diverse problem solvers can outperform groups of high-ability problem solvers. *Proceedings of the National Academy of Sciences*, *101*(46), 16385–16389. https://doi.org/10.1073/pnas.0403723101

Kim, T., McFee, E., Olguin, D. O., Waber, B., & Pentland, A. "Sandy." (2012). Sociometric badges: Using sensor technology to capture new forms of collaboration. *Journal of Organizational Behavior*, *33*(3), 412–427. https://doi.org/10.1002/job.1776

Landau, S. (2022). The development of a crypto policy community: Diffie–Hellman's impact on public policy. In *Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman* (1st ed., Vol. 42, pp. 213–256). Association for Computing Machinery. https://doi.org/10.1145/3549993.3550002

Levy, S. (2010). *Hackers: Heroes of the Computer Revolution* (O'Reilly, Ed.; 3rd ed.). O'Reilly. http://www.worldcat.org/isbn/0141000511

Loeliger, J., & McCullough, M. (2012). *Version Control with Git: Powerful Tools and Techniques for Collaborative Software Development*. O'Reilly Media, Inc.

Maillart, T., Sornette, D., Spaeth, S., & von Krogh, G. (2008). Empirical tests of Zipf's law mechanism in open source Linux distribution. *Physical Review Letters*, *101*(21), 218701. https://doi.org/10.1103/PhysRevLett.101.218701

Maillart, T., Zhao, M., Grossklags, J., & Chuang, J. (2017). Given enough eyeballs, all bugs are shallow? Revisiting Eric Raymond with bug bounty programs. *Journal of Cybersecurity*, *3*(2), 81–90. https://doi.org/10.1093/cybsec/tyx008

Malone, T. W. (2019). *Superminds: How Hyperconnectivity Is Changing the Way We Solve Problems*. Oneworld Publications.

Mohanty, S. P., Singhal, G., Scuccimarra, E. A., Kebaili, D., Héritier, H., Boulanger, V., & Salathé, M. (2022). The food recognition benchmark: Using deep learning to recognize food in images. *Frontiers in Nutrition*, *9*. https://www.frontiersin.org/articles/10.3389/fnut.2022.875143

Moody, G. (2009). *Rebel Code: Linux and the Open Source Revolution*. Hachette UK.

Nagle, F. (2018). Learning by contributing: Gaining competitive advantage through contribution to crowdsourced public goods. *Organization Science*, *29*(4), 569–587. https://doi.org/10.1287/orsc.2018.1202

Nagle, F. (2019). *Government Technology Policy, Social Value, and National Competitiveness* (SSRN Scholarly Paper 3355486). https://doi.org/10.2139/ssrn.3355486

Ntoutsi, E., Fafalios, P., Gadiraju, U., Iosifidis, V., Nejdl, W., Vidal, M.-E., Ruggieri, S., Turini, F., Papadopoulos, S., Krasanakis, E., Kompatsiaris, I., Kinder-Kurlanda, K., Wagner, C., Karimi, F., Fernandez, M., Alani, H., Berendt, B., Kruegel, T., Heinze, C., & Staab, S. (2020). Bias in data-driven artificial intelligence systems – An introductory survey. *WIREs Data Mining and Knowledge Discovery*, *10*(3), e1356. https://doi.org/10.1002/widm.1356

Osoba, O. A., IV, W. W., & Welser, W. (2017). *An Intelligence in Our Image: The Risks of Bias and Errors in Artificial Intelligence*. Rand Corporation.

Ostrom, E. (1990). *Governing the Commons: The Evolution of Institutions for Collective Action (Political Economy of Institutions and Decisions)*. Cambridge University Press. http://www.worldcat.org/isbn/0521405998

Percia David, D., Maréchal, L., Lacube, W., Gillard, S., Tsesmelis, M., Maillart, T., & Mermoud, A. (2023). Measuring security development in information technologies: A scientometric framework using arXiv e-prints. *Technological Forecasting and Social Change*, *188*, 122316. https://doi.org/10.1016/j.techfore.2023.122316

Proulx, J. D. E., Aknin, L. B., & Barasch, A. (2023). Let's give together: Can collaborative giving boost generosity? *Nonprofit and Voluntary Sector Quarterly*, *52*(1), 50–74. https://doi.org/10.1177/08997640221074699

Raymond, E. (1999). The cathedral and the bazaar. *Knowledge, Technology & Policy*, *12*(3), 23–49. https://doi.org/10.1007/s12130-999-1026-0

Ryan, R. M., & Deci, E. L. (2000). Intrinsic and extrinsic motivations: Classic definitions and new directions. *Contemporary Educational Psychology*, *25*(1), 54–67. https://doi.org/10.1006/ceps.1999.1020

Sornette, D., Maillart, T., & Ghezzi, G. (2014). How much is the whole really more than the sum of its parts? 1⊞ 1= 2.5: Superlinear productivity in collective group actions. *PLoS One*, *9*(8), e103023.

Tirole, J., & Lerner, J. (2002). Some simple economics of open source. *Journal of Industrial Economics*, *50*(2), 197–234.

van Schewick, B. (2012). *Internet Architecture and Innovation*. The MIT Press. http://www.worldcat.org/isbn/026251804X

von Krogh, G., Haefliger, S., Spaeth, S., & Wallin, M. W. (2012). Carrots and rainbows: motivation and social practice in open source software development. *MIS Quarterly*, *36*(2), 649–676. https://doi.org/10.2307/41703471

Wilson, R. C., Geana, A., White, J. M., Ludvig, E. A., & Cohen, J. D. (2014). Humans use directed and random exploration to solve the explore–exploit dilemma. *Journal of Experimental Psychology: General*, *143*, 2074–2081. https://doi.org/10.1037/a0038199

Woolley, A. W., Chabris, C. F., Pentland, A., Hashmi, N., & Malone, T. W. (2010). Evidence for a collective intelligence factor in the performance of human groups. *Science*, *330*(6004), 686–688. https://doi.org/10.1126/science.1193147

Yang, V. C., & Sandberg, A. (2023). *Collective Intelligence as Infrastructure for Reducing Broad Global Catastrophic Risks*. https://doi.org/10.25740/mf606ht6373

Zittrain, J. (2009). *The Future of the Internet–And How to Stop It*. Yale University Press. http://www.worldcat.org/isbn/0300151241

Appendix 21.1

A1. Food Recognition Challenge Case Study

Purpose and Intention: The Food Recognition Challenge was established to address a common yet complex problem in nutritional informatics – accurately identifying food from images. It intends to harness deep learning capabilities to develop tools that assist users, ranging from individuals tracking their dietary intake to medical professionals conducting nutritional studies.

Problem Statement: Despite advancements in computer vision, food recognition from images remains an intricate task due to the diversity of food appearance, presentation, and context. The challenge's problem statement revolves around creating models that can robustly identify and analyze food items in varying conditions, pushing the limits of current technology.

Motivation: The initiative is driven by the need for precise food-tracking mechanisms and the desire to foster community engagement in solving this problem. The challenge focuses on technological innovation and aims to improve public health outcomes and support medical research.

Impact of the Challenge: Over successive iterations, the challenge has made significant impacts by (i) providing an open, evolving benchmark for food recognition, encouraging ongoing participation and development, (ii) releasing annotated datasets to the public, facilitating research and application in real-world scenarios, and (iii) creating a platform for developers and researchers to collaborate and compete, spurring advancements in the field.

Solution and Contribution: Participants contribute solutions that utilize a novel dataset from the MyFoodRepo app, which is continuously updated with images and annotations for segmentation, classification, and weight/volume estimation. These contributions have led to improved models capable of detecting individual food items and understanding dietary patterns.

Challenge Cycle: The challenge operates in multiple rounds, each with specific tasks and datasets that reflect the growing dataset and technological advancements. This cyclical nature ensures the challenge remains relevant and continues to adapt to community feedback and the latest research findings.

Common Goal: The shared goal of the Food Recognition Challenge is to provide a high-quality dataset that serves as a foundation for developing effective food recognition algorithms. Unlike the "beautiful" but unrepresentative stock photos found online, these algorithms are expected to work with real-world images. By doing so, the challenge aims to create AI tools that can be widely adopted for personal and medical use.

Methods: The Food Recognition Challenge focuses on developing AI models to identify food items in images. These models should be capable of detecting and annotating individual food items with accurate segmentation, classification, and weight/volume estimation. The challenge uses a novel dataset collected through the MyFoodRepo app, contributed by volunteer Swiss users documenting their daily food intake. This dataset has been annotated to map the individual food items onto an ontology of Swiss Food items.

Dataset: The dataset provided by the AIcrowd Food Recognition Challenge is an evolving collection of food images with annotations in MS-COCO format. It includes:

- A Training Set with 24,120 RGB images and 39,328 annotations.
- A suggested Validation Set with 1,269 RGB images and 2,053 annotations.

This is a debug Test Set for Round 3, offering the same images as the validation set. The dataset is designed to overcome the limitations of existing food databases, which often feature unrepresentative stock photography without proper annotations. The challenge dataset aims to provide real-world images with proper segmentation, classification, and volume/weight estimates.

Results: The Food Recognition Challenge offers substantial prizes to incentivize participants. For Round 4, significant cash prizes were awarded for scores above specific thresholds, with the top prize being 10,000 CHF for a score greater than 0.70. Additionally, the top four winners received an Oculus Quest 2, and a travel grant to AMLD 2021 was also provided. These incentives aim to encourage high-quality submissions and advancements in food recognition.

Submissions: Participants were required to set up a proper repository structure and create a private Git repository at GitLab with the contents of their submission. Submissions were identified using an aicrowd.json file containing specific fields, including the challenge ID and whether the submission required a GPU for evaluation. If needed, a NVIDIA-K80 GPU was made available for the submission evaluation.

Here are some key takeaways from the Food Recognition Challenges:

Community Engagement Is Crucial: The challenges have consistently emphasized community involvement, leveraging crowd-sourced data from the MyFoodRepo app and encouraging developers globally to contribute to the evolving dataset.

Real-World Application Focus: The practical use case of the challenge – to help track dietary intake for personal and medical purposes – highlights the importance of AI applications that can be integrated into everyday life.

Evolving Datasets Enhance Relevance: The datasets have grown over time, ensuring that the challenge remains relevant and that the algorithms developed are tested against a diverse and up-to-date range of food images.

Difficulty of Image-Based Recognition: Despite advances in deep learning, food recognition from images is still a difficult problem due to the variability in food presentation, which these challenges aim to address.

Quality of Data Over Quantity: The emphasis on high-quality, well-annotated datasets underscores the challenge's commitment to creating reliable and accurate AI models, moving away from unrepresentative and misleading Internet images.

Continuous Improvement Through Iterative Rounds: The challenge's multi-round structure fosters ongoing improvement and innovation, allowing participants to build upon previous work and adapt to new data.

Incentives Drive Innovation: Substantial prizes and recognition, such as co-authorships in papers and cash rewards, are significant incentives for participation and pushing the boundaries of current AI capabilities.

Openness and Collaboration: By establishing an open benchmark and providing resources like starter kits and discussion forums, the challenge encourages transparency and collaboration within the AI community.

Accessibility and Ease of Entry: The challenges have lowered the barrier to entry, allowing a broad range of participants, from those with access to powerful AI models to individuals who may just be starting.

Results Demonstrate Feasibility and Progress: The results and solutions generated from these challenges demonstrate the feasibility of using AI for food recognition and tracking, showcasing progress and paving the way for further advancements in the field.

Appendix 21.2

A2. Relation between variables used for modeling

Exploring the relation of target variables and model prediction further indicates the overall linear nature of the phenomenon under study and indicates that the precision of exceptionally performant solutions fails to be predicted accurately given their scarcity in the data. Figure 21.4 shows the pairs' scatter plots (left quadrant), histograms (diagonal), and correlations (right quadrant) for variables used for modeling. Color-coding of scatter plots represents the different teams.