The Innovation Game

John Fletcher, Ying Chan, Philip David

April 2024. Version: 2.2

Abstract

The Innovation Game (TIG) is a novel market-based framework designed to accelerate development of computational methods crucial to data-driven sciences. TIG conceptualises a "synthetic market" based on proof of work. In this market, "Innovators" contribute methods for solving "asymmetric" computational science problems and are rewarded based on adoption by "Benchmarkers" who, in turn, are objectively rewarded based on their performance in generating proof of work by solving instances of those problems. Accordingly, this market incentivises the open development of an important class of computational methods and provides price discovery for both commercial and pre-commercial research in this area, allowing the research to attract private investment. TIG introduces a hybrid licensing model that combines open data licenses with traditional commercial licenses. This model aims to manage intellectual property rights effectively, ensuring a route for open collaboration while also enabling value capture. Additionally, mechanisms to ensure market competitiveness and prevent monopolistic dominance are introduced.

The aim of The Innovation Game is to accelerate development of computational methods for solving "asymmetric" problems. These problems require significant computational effort to solve, yet their solutions are easily validated once proposed. Numerous problems belong to this category. For example, NP-complete problems, which are simple to check and generally considered unsolvable within polynomial time. Such problems are fundamental in science and engineering. Prominent examples include the Hamiltonian Cycle Problem, involving the identification of a Hamiltonian cycle in a given graph, and the Boolean Satisfiability Problem (SAT), which entails deciding if a certain Boolean formula can be satisfied.

Another well-known example of an asymmetric problem is creating a model based on given data, which in the field of artificial intelligence, equates to training a machine learning model. This approach has a wide range of applications, from generation of language models to the creation of images. In scientific research, these techniques can allow for the processing of extensive datasets to formulate and test more complex hypotheses, an essential capability in the age of data-centric science.

1 Accelerating Science: Openness and Efficiency

Scientific research can be broadly divided into two categories: fundamental and applied. Fundamental science research, which aims to expand our understanding of the natural world, often leads to discoveries that form the foundation for new technologies. For instance, the development of the internet, now a cornerstone of global communication and commerce, originated from fundamental research in computer science and electrical engineering. Similarly, advances in medicine, such as MRI machines, are direct outcomes of fundamental research in physics. Applied science focuses on practical applications of scientific knowledge to create products and solutions that enhance the quality of life. These advancements not only improve living standards but also drive economic growth. The revenue generated from these technologies can be, and often is, reinvested into fundamental research, with a view to creating a cycle of innovation and discovery. However, despite the significant returns driven by this virtuous relationship between fundamental and applied research, there remain two major unresolved issues in the economics of science, which present obstacles to further progress:[1]

(a) For commercial (applied) scientific research, private investment is typically available, but provided in return for exclusive rights to exploit the research output (including data). In other words, private investment normally requires that the research output is **proprietary**, which tends to discourage open collaboration.

(b) For pre-commercial scientific research (such as fundamental research[2]) outputs are more often non-proprietary. However, private investment is typically unavailable due to a well-known market failure which prevents the market from returning downstream value generated by fundamental scientific research to the researchers themselves. This is believed to result in an **inefficient allocation** of resources for fundamental research.

[1] See for example [1].

[2] We acknowledge that some applied research is also pre-commercial (for example, research into nuclear fusion power, at least until recently). Nevertheless, we will use the terms "fundamental" and "pre-commercial" somewhat interchangeably.

We seek a framework for open collaborative development of computational methods that will accelerate scientific research. This framework should operate on a commercial basis, without requiring subsidy. As such, Open Source development is a natural path to consider. However, as Eric Raymond explains in his seminal essay, *The Magic Cauldron*, Open Source development is not well suited[3] to the type of code-based projects that are common in computational science [3]. Raymond further explains that open collaborative development of such projects *can* be viable, provided that there is:

- A **value capture** mechanism (a method of capturing the value generated by the project), and
- A **price discovery** mechanism (a mechanism for pricing contributions to the project, enabling a portion of the captured value to be returned to contributors in accordance with the value of their contribution).[4]

In the next section, we present a framework called The Innovation Game, which provides a value capture and price discovery mechanisms for computational methods for solving asymmetric problems. Following Raymond, we aim to achieve this by addressing issues (a) and (b) for the open development of computational methods required to solve these problems. See Figure 1.

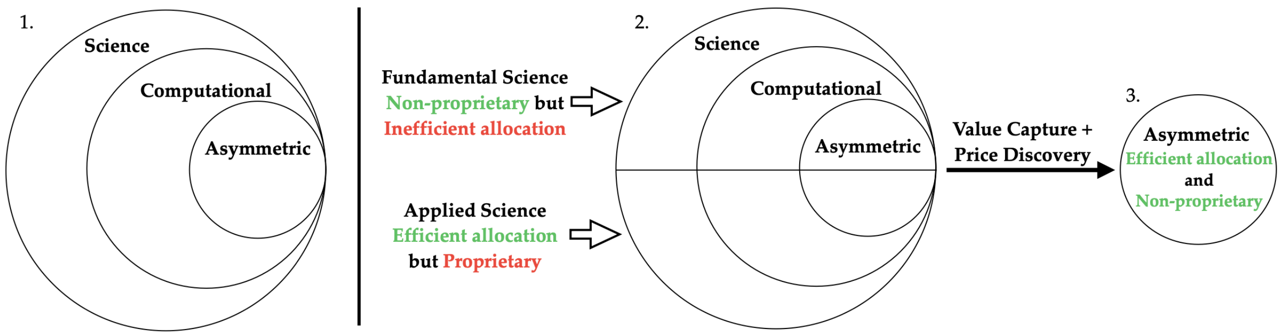

Figure 1: 1. Computational science is a subset of science; a subset of problems in computational science are asymmetric (meaning that solutions are difficult to find but easy to verify—see above); 2. Scientific research can be broadly divided into fundamental (pre-commercial) and applied (commercial); 3. Value capture combined with price discovery can address issues (a) and (b), above, for research in asymmetric computational problems.

Open collaborative development in science also faces significant challenges relating to data, particularly when machine learning is employed. Data sets, often viewed as factual lists, fall outside traditional property rights in major jurisdictions, including the U.S., making a copyleft[5] effect unattainable. Consequently, open projects sharing data risk free riding by closed projects who may use the data without reciprocal sharing [6]. It follows that, in certain domains (notably machine learning) the lack of availability of relevant data diminishes the intrinsic value of computational methods. To address this issue, The Innovation Game aims to provide incentives for sharing relevant data through its licensing strategy.

2 Our Solution: The Innovation Game

The Innovation Game seeks to accelerate scientific progress by addressing issues of resource allocation and proprietary outputs in pre-commercial and commercial scientific research respectively (issues (a) and (b), above) for asymmetric problems in computational science. The Innovation Game achieves this by enabling price discovery and a value capture mechanism. **Price discovery** is facilitated by creating a "synthetic" market in implementations of algorithms (which we call "Methods") for solving important problems in computational science (referred to here as "Challenges"). We call this market synthetic because it arrises from an artificial source of demand created by agents called Benchmarkers, who are introduced in order to facilitate price discovery for Methods that they use to solve instances of Challenges. Benchmarkers receive rewards for submitting these solutions. The effect is to create demand for optimised Methods in an analogous way to how Bitcoin miners create demand for optimised mining hardware [7]. Agents called Innovators satisfy Benchmarkers' demand for improved Methods. Innovators receive rewards based on their Methods' performance, measured by the degree of its adoption by the Benchmarkers. The Innovation Game achieves **value capture** by securing the intellectual property embodied in the Methods and creating incentives for certain third parties to license them under the

[3]Raymond recognised the issue of pricing code contributions as an instance of F.A. Hayek's renowned 'calculation problem' that can be resolved via a market pricing mechanism [2]. See Appendix A for further discussion.

[4]More generally, price discovery is the process by which the market determines the price of an asset through the interactions of buyers and sellers, allowing for efficient allocation of resources. See for example [4].

[5]Copyleft is a licensing approach that grants the right to freely distribute and modify creative work, with the condition that any derivative work must be distributed under the same or similar license terms [5]. See also Property 1, Appendix A.

terms of a paid commercial licence. These features allow The Innovation Game to provide sustainable incentives for development of improved Methods for solving asymmetric problems in computational science.

There are three types of Players in The Innovation Game:

• **Benchmarkers** (Token miners; supporters of The Innovation Game; anyone with a computer). Benchmarkers solve random instances of the Challenges featured in The Innovation Game using Methods that have been submitted by Innovators (see below), and earn token rewards in return. Benchmarkers can gain a competitive advantage by solving Challenge instances more efficiently. This results in a clear incentive to use the most efficient Method in order to reduce their operating costs and increase profitability. In The Innovation Game, the Benchmarkers' role is to provide a measure of the relative performance of the Methods. See Figure 2.

• **Commercial Enterprises** (Entities that sell products based on the IP secured by The TIG Foundation). Commercial enterprises must pay a fee to use the IP secured by The Innovation Game for commercial purposes other than Benchmarking, *unless they are willing to make their data and/or code available.* The fee permits use of all IP secured by The TIG Foundation, is non-discriminatory and rate card based, and is payable in TIG tokens (see Section 2.2 for details). This generates demand for the token.

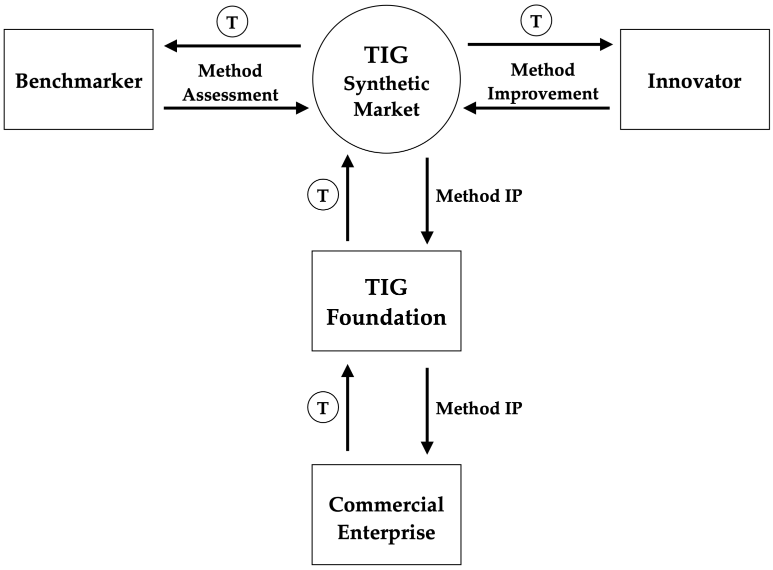

Figure 2: Economic model of The Innovation Game. Arrows represent asset flows (goods, services, tokens). Circled 'T' denotes TIG tokens.

• **Innovators** (Coders; academic scientists; entities conducting commercial research). Innovators develop and submit more efficient Methods, in the form of code, to The Innovation Game, aiming to be the most efficient at solving instances of the featured Challenges. As long as the Method reaches a threshold of adoption by Benchmarkers, the Innovator receives a reward of tokens. The role of the Innovator in The Innovation Game is to contribute the intellectual property embodied in the Methods.[6] Note that the role of an Innovator may be automated [8].

Our objective is for the rewards for Innovators to correlate with the expected value of their contributions in terms of the downstream value generated by the contribution. Our strategy for achieving this is outlined in the following sections. On the face of it, the creation of a synthetic market in computational methods might appear simple. However, without careful design, such a market would face two significant problems:

(1) The ability to commercially exploit the Method could remain exclusively with the Innovator, meaning there is no value capture in the token, meaning that the system is unsustainable.

(2) The market would tend to centralise towards monopoly, reducing competition and the incentive to innovate.

To address (1), The Innovation Game requires that any Method used by a Benchmarker must be made available to all Benchmarkers. In addition, the Innovator must grant rights to exploit the new Method to the TIG Foundation, which will make the Method available for license under both a license consistent with open collaboration (called the TIG Open Data license[7]) and a paid license called the TIG Commercial license. Regarding (2), the nature of The Innovation Game as an open, global competition, means that national antitrust regulations cannot be relied upon to prevent Benchmarkers from tending to monopoly. We therefore introduce a novel mechanism (the "antitrust mechanism") designed to prevent any one Benchmarker from unduly dominating.

[6] The Innovation Game makes a distinction between algorithmic and code optimisations, and this is reflected in the Innovator rewards. The details of the rewards mechanism, including a way to establish the type of optimisation (algorithm or code), and the division of rewards between optimisation types, will be presented in our forthcoming technical paper.

[7] The TIG Open Data license requires that input data is made generally available if the output data, or the covered code used to generate it, is distributed. In this way, the TIG Open Data license provides a unique incentive to share data: The right to use all IP captured by The Innovation Game for free. See Section 2.1 for details.

2.1 Price Discovery in The Innovation Game

The mechanism that we use to enable price discovery in The Innovation Game is proof of work. Proof of work (POW) was first introduced to provide a pricing function for sending email in order to deter email spam [9]. Fundamental to digital currencies such as Bitcoin, POW is a mechanism that enforces a cost on the execution of a task while being straightforward for others to verify the completion of that task.

Despite their significance in science and technology, problems with scope for optimisation have been excluded from application in proof of work mining, because of the potential for significant speedup through algorithmic optimisation, which could unduly advantage a miner by massively improving efficiency. This would lead to centralisation concerns, potentially jeopardising the security and fairness of the network [10].

To address this issue, The Innovation Game introduces *optimisable* proof of work (OPOW). OPOW encompasses a set of n distinct POW Challenges, each potentially open to algorithmic enhancement. We require that these Challenges satisfy two conditions: (i) Challenges must be asymmetric (as defined above), (ii) Challenges should be sufficiently independent so that an optimisation for one Challenge should not lead to similar improvements across most other Challenges. The Innovation Game is designed such that private optimisation advantages remain tightly constrained. This makes the system resistant to monopolisation and retains its decentralised nature. In this way, OPOW provides an "antitrust mechanism" within the synthetic market of The Innovation Game, helping to maintain competition and security essential to a functioning market for innovation.

To compute the rewards earned by Benchmarkers performing OPOW, denote by f_x^i the number of qualifying solutions[8] for the x^{th} Challenge submitted by Benchmarker i. The normalised f_x^i is defined $\hat{f}_x^i \equiv f_x^i / \sum_j f_x^j$ where the sum ranges over $1, 2, ..., N_m$, and N_m is the number of Benchmarkers. For n Challenges, we call the condition $\hat{f}_1^i = ... = \hat{f}_n^i$ "parity." A Benchmarker i's rewards R^i are increasing in the mean of the $\hat{f}_{x=1,..,n}^i$, and decreasing in the spread of the $\hat{f}_{x=1,..,n}^i$, as follows:

$$R^i \propto \langle \hat{f} \rangle \exp\left(-\frac{k\,(C_V^i)^2}{n-1} \right), \quad \langle \hat{f} \rangle^i \equiv \frac{1}{n} \sum_y \hat{f}_y^i, \quad C_V^i \equiv \frac{\sigma^i}{\langle \hat{f} \rangle^i}, \quad \sigma^i \equiv \left\langle (\langle \hat{f} \rangle^i - \hat{f}^i)^2 \right\rangle^{1/2}$$

Above, C_V^i is the *coefficient of variation*, and k parameterises the penalty for deviation from parity.

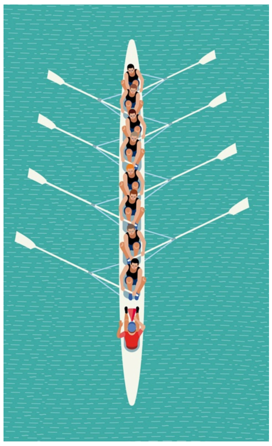

Figure 3: In TIG, a Benchmarker is analogous to a rowing boat, in that they benefit from synchronisation.

The effect of distributing Benchmarker rewards this way can be illustrated with an analogy (Figure 3): Consider a rowing boat (analogous to a Benchmarker) with n rowers (analogous to one of the n Challenges). In this boat, one rower is a hundred times stronger and faster than any other rower (we assume that they are also a similar weight). Although this obviously confers an advantage, because the oars must remain reasonably synchronised, the "super rower" can provide only a limited boost to their boat. Similarly, an Innovator/Benchmarker in The Innovation Game can only achieve limited benefit from being relatively "stronger" in a minority of Challenges, resulting in a diminished incentive to keep an improved Method for private use.

OPOW means that a far wider class of Challenges can be used as a proof of work than previously possible—specifically—the class of asymmetric Challenges which includes a vast array of problems of considerable scientific interest and/or with valuable real-world applications. This is leveraged in The Innovation Game to create a synthetic market where an Innovator earns more rewards the broader and more sustained the adoption of their Method. These factors are used as a measure of the Method's performance, which we expect to be correlated with the additional value the Method will enable The Innovation Game to capture (see Section 4 for details). This is how the synthetic market enables price discovery for Methods corresponding to the featured Challenges.

Selection of Challenges in The Innovation Game

For a Challenge to be added to The Innovation Game, there should be a high degree of consensus among experts as to it's significance in science or technology. To facilitate the process of adding Challenges, a committee of experts will nominate Challenges for consideration. Following the committee's nomination, token holders vote on whether to add the Challenge to The Innovation Game.[9] The vote serves as a check on the committee's power to decide which Challenges to add to The Innovation Game and when. Token holders have a vested

[8] A qualifying solution is a solution to a proof of work Challenge which is of sufficient difficulty and is submitted to The Innovation Game within a certain time period. This notion will be precisely defined in our technical paper.

[9] Token holders may delegate their voting power to another user if they do not feel equipped to make an informed decision.

interest in the timing and frequency of Challenge additions to The Innovation Game because of the impact that this has on captured IP value and, consequently, TIG token value.[10]

The incentive to add Challenges is to increase the potential capture of licensable IP by The TIG Foundation. However, as all Challenges in The Innovation Game share a single reward pool, adding too many Challenges would dilute the reward available for each Challenge and diminish the incentive for Innovators to contribute. Achieving a balance between the total number of Challenges and individual Challenge rewards is vital for maximising valuable innovation and token value. Consequently, token holders are incentivised to vote to maintain the optimal balance. Challenges can also be retired from The Innovation Game through a similar mechanism: a proposal by the committee, followed by a ratifying vote from the token holders. We anticipate that the set of Challenges will stabilise over the long term.

2.2 Value Capture in The Innovation Game

For the TIG token value, and therefore the incentives that underpin The Innovation Game, to be sustainable, we recognise the requirement for the token to represent an asset from which revenue can be derived. Submissions to The Innovation Game will be protected by copyright and may embody patentable inventions. Value capture in The Innovation Game requires[11] the TIG Foundation to secure or exclusively control at least some IP rights for innovations for solving instances of the Challenges. This IP represents the primary asset from which revenue may be derived.

Benchmarkers offer a market signal as to which Methods represent improvements over previous bests. This signal is used to make decisions regarding patent applications for any underlying inventions. Regardless of whether patent applications can be successfully prosecuted to grant, we believe that there will be significant value in the copyright in the submitted optimised code for the Methods which are difficult to reimplement without either infringing copyright, or loss of the advantage intrinsic to that optimised code.

Intellectual Property and Licensing in The Innovation Game

The Innovation Game aims to capture valuable intellectual property embodied in the Methods which it will make available under several licenses. There are five distinct licences in The Innovation Game:[12]

- **The TIG Innovator Outbound Game licence**: A licence which will gateway access to Methods previously submitted to The Innovation Game for use by Innovators who wish to use those Methods to participate in The Innovation Game.
- **The TIG Benchmarker Outbound Game licence**: A licence which will gateway access to Methods previously submitted to The Innovation Game for use by Benchmarkers who wish to use those Methods to participate in The Innovation Game.
- **The TIG Inbound Game licence**: A licence which will gateway submissions of new Methods to The Innovation Game and which, in conjunction with the TIG Game Rules, will enable the TIG Foundation to secure Intellectual Property rights in the innovation resulting from participation in The Innovation Game.
- **The TIG Open Data licence**: A licence which aspires to meet objective standards of openness for code and data, and which includes a **share alike obligation**. The *TIG Open Data License* will require that if data generated by execution of the subject Method ("**Output Data**") on certain input data is made available to third parties as data or in the form of a product derived from the data, then the input data processed by execution of the Method to generate the Output Data together with any additional information to generate a product derived therefrom must also be made **generally available** to the extent necessary to allow a reproduction of the Output Data or product, as the case may be (we call this input data and additional information "**Relevant Data**"). The share alike obligation will require that; (i) where the subject Method alone is distributed, the source code for the Method must be made available under the TIG Open Data License to the recipient of the Method; and (ii) where Output Data is distributed, the source code for the Method and the Relevant Data must be made generally available for distribution under the TIG Open Data License.

[10]For example, if the committee were to nominate a Challenge that is clearly not sufficiently important to justify inclusion in The Innovation Game, then its inclusion can be vetoed by the token holders—the incentive for this is clear: Challenges of lower significance imply less valuable IP in The Innovation Game, less spillover value, and lower demand for the TIG token.

[11]We may later introduce additional modes of value capture such as charging a fee for a Challenge to be proposed.

[12]See the *TIG IP Policy Rationale* for details [11]. We emphasise that, despite superficial similarities, the TIG licensing model is fundamentally different from traditional dual licensing. For example, dual licensing is synonymous with *single vendor licensing*, which requires assignment of copyright by contributors. TIG licensing does not require such an assignment.

- **The TIG Commercial licence**: In contrast with the TIG Open Data License, The TIG Commercial License provides greater freedom with respect to downstream licensing, allowing the use of any and all TIG IP while exempting the licensee from the obligations (under the TIG Open Data licence) to make Relevant Data and source code available, in return for the payment of a fee (non-discriminatory, and rate card based). The fee is used to underpin value in The Innovation Game's incentives.

The intellectual property embodied in Methods submitted by Innovators and secured by the TIG Foundation will be managed and licensed by the TIG Foundation. The TIG Foundation provides undertakings to Innovators to make contributed IP available under the TIG licences, as described above. For algorithmic optimisations, Benchmarkers provide a performance measure that acts as a market signal, informing the decision to seek patent protection for any patentable inventions within the Methods.[13] Long term, the demand for tokens should correlate with the demand for TIG Commercial Licences, which require a fee payable in tokens. This, in turn, will be correlated with the value of the intellectual property held by the TIG Foundation.

This intellectual property will also gain value from several key features integral to The Innovation Game:

Spillover Effects. A significant increase in value is anticipated from spillover benefits arising from the inclusion of a wide array of challenges within The Innovation Game, covering diverse domains of fundamental and applied research. Such spillovers allow The Innovation Game to mirror the strategy employed by governments to earn a return from investments in fundamental research, where broader industry benefits from commercialisation of subsequent spinoff innovations.

Open Collaboration. The open sharing of code and data is crucial for enabling Innovators to augment and refine each other's work. This principle, which lies at the heart of the Open Science movement, has also been a crucial factor in the remarkable successes of Open Source. The TIG Open Data license is designed to facilitate this kind of collaborative innovation. The availability of open data indirectly increments the value of The Innovation Game's IP assets. The wider The Innovation Game's role in bringing about the sharing of open code and data resources, the greater its perceived contribution to the open movement will be, enhancing the community's regard for the initiative.

Community Building. Scientists' behaviour often adheres to specific norms, with motives for participation that include a desire for openness and a drive to build a reputation within their community. The competitive dimension to The Innovation Game is encouraged, with leaderboards identifying the top performers in algorithmic and code optimisations for each Challenge. We intend for The Innovation Game to establish its own community of Innovators, motivated by factors similar to those which motivate academic scientists and contributors in the Open Source community. Strong community norms will also help reduce free riding on The Innovation Game's intellectual property.

3 Summary of The Innovation Game

The Innovation Game is designed to accelerate the development of computational methods for tackling a broad and critical category of problems known as asymmetric Challenges. These Challenges are pivotal in fields such as machine learning, combinatorial optimisation, and various mathematical problems, including prime number factorisation. They are also common in scientific inverse problems, which are essential for advancing data-driven science. The Innovation Game carefully selects Challenges to maximise the potential for knowledge spillovers, which are beneficial exchanges of information that can lead to further advancements.

To fulfil its goal, The Innovation Game confronts two primary obstacles: the inefficient allocation of resources in fundamental research and the limitations brought about by the proprietary nature of applied research. To overcome these, The Innovation Game employs two key mechanisms: price discovery and value capture. Price discovery within The Innovation Game is facilitated by a synthetic market. This market not only determines how resources should be distributed but also actively directs those resources towards the Innovators who develop and submit new computational methods, or "Methods," to The Innovation Game. The market's design, which relies on the proof of work from participants called Benchmarkers, is robust against manipulation and naturally incorporates the economic considerations for producing specialised hardware needed to run these Methods.

Value capture is achieved through a strategic multi-licensing approach. The Innovation Game incentivises the purchase of a TIG Commercial license by allowing entities to keep their source code and data proprietary, bypassing the share-alike requirements of the TIG Open Data license. Conversely, The Innovation Game also encourages the open sharing of source code and data by offering free access to its captured intellectual property under the TIG Open Data license. These dual incentives are crucial for The Innovation Game to meet its aims. An equilibrium is expected to form between entities drawn to the benefits of sharing and those willing to pay for the privilege of exclusivity. The fees collected are then redistributed to enhance the rewards for Innovators.

[13] See Section A.1 for discussion of the utility of patents in the context of The Innovation Game.

References

[1] Dasgupta Partha and Paul A David. Toward a new economics of science. In *Science bought and sold: Essays in the economics of science*, pages 219–248. University of Chicago Press, 2002.

[2] Friedrich Hayek. The use of knowledge in society. *The American economic review*, 35(4):519–530, 1945.

[3] Eric Raymond. The magic cauldron. In *The cathedral and the bazaar*, pages 113–166. O'Reilly, 1999.

[4] Paul Milgrom. *Discovering Prices.* Columbia University Press, 2017.

[5] Free Software Foundation. What is Copyleft? https://www.gnu.org/licenses/copyleft.en.html, (2024).

[6] L. Villa. Copyleft, Attribution, Data. lu.is/blog/2016/09/21/copyleft-attribution-and-data-other-considerations, 2017.

[7] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. 2008.

[8] Bernardino Romera-Paredes et al. Mathematical discoveries from program search with large language models. *Nature*, pages 1–3, 2023.

[9] Cynthia Dwork and Moni Naor. Pricing via processing or combatting junk mail. In *Annual international cryptology conference*, pages 139–147. Springer, 1992.

[10] Andrew Poelstra. ASICs and decentralization FAQ, 2015.

[11] Philip David, John Fletcher. TIG IP Policy Rationale. 2023.

[12] Josh Lerner and Jean Tirole. Some simple economics of open source. *The journal of industrial economics*, 50(2):197–234, 2002.

[13] James Bessen. Open source software: Free provision of complex public goods. In *The economics of open source software development*, pages 57–81. Elsevier, 2006.

[14] Carliss Y Baldwin and Kim B Clark. The architecture of participation: Does code architecture mitigate free riding in the open source development model? *Management science*, 52(7):1116–1127, 2006.

[15] Jean-Gabriel Young, Amanda Casari, Katie McLaughlin, Milo Z Trujillo, Laurent Hébert-Dufresne, and James P Bagrow. Which contributions count? analysis of attribution in open source. In *2021 IEEE/ACM 18th International Conference on Mining Software Repositories (MSR)*, pages 242–253. IEEE, 2021.

[16] Mikko Valimaki. Dual licensing in open source software industry. *Systemes d'Information et Management*, 8(1):63–75, 2003.

A Why TIG is Complementary to Open Source and not a Competitor to it

An open and collaborative approach to code development offers numerous advantages. Crowdsourced peer review enables early identification of issues and ensures adherence to standards, while access to a diverse pool of skilled contributors yields higher-quality outputs, delivered more quickly and at reduced cost [12]. Open collaboration may lead to a tipping point where, with sufficient participation, the project rapidly outpaces closed alternatives.

The power of the open and collaborative approach is exemplified by the Open Source software movement. By promoting source code sharing while mitigating the free rider problem, Open Source has succeeded in providing world-class public goods in the form of software, without the need for government or charitable funding. The Innovation Game is designed to accelerate science by accelerating the open development of important methods in computational science. These methods are embodied in computer code. As such, the Open Source movement is an important inspiration for The Innovation Game. Nevertheless, key differences exist between the economics of science and Open Source, necessitating a new approach, as we will discuss.

Successful Open Source development is most likely with codebases that possess all of the following properties:

Property 1 (*Large codebase and/or performance-sensitive application*). Copyright is fundamental to open source licensing, granting the original creators exclusive rights to reproduce, distribute, and modify their work, which they leverage to maintain the code's openness (known as "copyleft"). Copyright can be circumvented by reimplementing the code, thus avoiding copyleft. However, reimplementation is difficult for large codebases, and tends to be difficult to achieve while retaining the benefit of code optimisations which are required for performance-sensitive applications [3].

Property 2 (*Immediate utility*). Contributors are incentivised to help create an output of immediate utility, whether commercial or otherwise [13]. In Open Source, access to the resulting software is usually a major part of the incentive to contribute.

Property 3 (*Non-monolithic/modular structure*). A modular project allows for multiple contributions to be made simultaneously with minimal coordination. Modularity also enables individuals to contribute usefully without needing to understand the entire project, making it more accessible. As a result, a highly modular project can thrive without full-time contributors who would normally require monetary compensation [14].

Some or all of these properties are often absent in implementations of algorithms for computational science and machine learning. Absence of these Properties can be compensated for by introducing monetary incentives (at least in principle), contingent on a significant degree of value capture being achievable.

Dual licensing is an established approach to capturing value in open and collaborative software development which aims to balance openness with monetisation by offering both a free (Open Source) licence and a (paid) commercial licence for the same software.[14] In traditional dual licensing the Open Source licence is typically of the copyleft variety: requiring that any modifications to the software be released under the same licensing terms. For developers who don't want to release their modifications under these terms, there is the option to pay for a commercial licence. Under certain circumstances, such as when Property 1 is absent, the copyright on which copyleft depends can be difficult to enforce. In this scenario, patents can be used.[15]

In the context of algorithms for science, traditional dual licensing has various limitations. The open source development model is not suited to monolithic codebases, which are often a feature of these algorithms (Property 3). Effective value capture requires that some property rights be secured, which is challenging if the research area has no immediate commercial utility (Property 2).

A further issue is that modifications or additions to these algorithms might not be distributed in the ordinary course of commercial use—instead— it is the *output data* that is distributed. However, in copyleft licences such as GPL, the requirement to make additions or improvements to the code available (under the same licensing terms) is triggered only by the distribution of the source code. It follows that, for such algorithms, the intended effect of copyleft may be neutralised, making it harder to capture value through dual licensing because there is no clear incentive to pay for a commercial licence.

A.1 Why The Innovation Game Doesn't Require Properties 1-3

In this Section, we discuss how The Innovation Game enables open development from codebases lacking the Properties described in Section A. Regarding Property 1, the greater likelihood of reimplementation due to a smaller codebase is offset by the high performance sensitivity typical of scientific applications. Nevertheless, a well-resourced entity might attempt to reimplement an algorithm to avoid the licence fee and apply their own code optimisation. Although our licensing fees are structured to make this approach financially unattractive, we will also seek patent protection for the algorithm to further reduce this risk.

As we have seen, the demand, generated by Benchmarkers performing OPOW, creates a "synthetic market" for the Methods supplied by the Innovators. This demand exists irrespective of the demand for such algorithms in wider industry. Thus "immediate utility" (Property 2), if not already present, is brought about by the Benchmarkers. The synthetic market distributes rewards in the form of tokens to Innovators, enabling the possibility of full-time engagement. This necessary since codebases for computational methods are often monolithic (Property 3). Currently, applied research projects in computational science (those which are privately funded) tend to be conducted on a proprietary basis, where funding is given in return for ownership of the resulting IP. With The Innovation Game, commercial funding can be provided in return for tokens earned by a Method submitted by an Innovator, with the IP embodied by the Method becoming a public good.

For fundamental research in computational science, the incentives are different to those of applied research. This is because, for fundamental research, Property 2 is also absent: The outcomes from fundamental research tend to be so uncertain that the incentive to participate *in order to use the output* is significantly diminished. This means that private research funding is not normally available, and so fundamental research must be *publicly* funded. For fundamental research, public funding eliminates the commercial case against Open Sourcing, but the lack of Properties 1-3 tends to reduce engagement, limiting the potential benefits of open collaborative development. By offering token rewards, The Innovation Game aims to increase engagement in fundamental research in computational science, and also improve efficiency of resource allocation by utilising the market mechanism tied to OPOW.

[14] In recent years, dual licensing has fallen out of favour, partly due to the perception that it exploits the work of voluntary contributors who do not receive any direct financial compensation for their contributions [15]. The market created by The Innovation Game addresses this issue of contributor compensation.

[15] The developer receives a patent licence if, and only if, the licensing terms are honoured. One example of a project with this licensing approach is MySQL [16].