PROTOFECT

Re: **Request for Comments on Patenting Artificial Intelligence Inventions**

Protofect Response
Measuring the AI Stack and a new AI Patent type

To citizens, leaders, scientists and whom it may concern:

The United States Patent and Trademark Office, Department of Commerce has requested information regarding patent-related issues concerning artificial intelligence (AI) inventions for purposes of evaluating whether further examination guidance is needed to promote the reliability and predictability of patenting artificial intelligence inventions.

I am Dr. Suman Deb Roy, the Founder and CEO of Protofect, a data science and AI firm based in New York City. The expertise of Protofect lies in maximizing the performance of data science and AI technologies by systematically designing and measuring its interacting modules, predicting adversarial situations, mapping data lineage, optimizing model tuning and prioritizing instrumentation and model explainability.

According to PWC, AI is expected to be $15.7 trillion industry by 2030. Software infused with machine learning and AI will control 70% of the global economy. However, in terms of science, the AI field is still young compared to manufacturing or biology. Significant time, care and understanding is needed to comprehend the internal engine driving responsible AI systems.

My intention in this letter is to communicate two aspects of a story:

1. That *AI-infused software behaves exponentially differently compared to traditional software.* Meaning over time, AI will show more and more divergence from machines running traditional software.
2. And second, *the need to measure AI's component elements and plan a new type of patent for AI inventions*, in a similar spirit as plant patents.

In the beginning, we must start with a general question: how different is the operation and working mechanism of AI software compared to traditional software?

Translucent Decision Pathways

When traditional software produces an output, we usually have an excellent logic trail to pin-point exactly why it did what it did. However, for AI software, such clear decision audits are currently tenuous at best. There are several intermingled factors that cause this. Some are scientific, others cultural.

1. *The Black Box:* There is widespread media coverage that AI behaves like a black box, its inner workings and mechanisms being obstructed from easy interrogation, analysis and study. While this is somewhat true, the real problem lies in a step further - that much of AI and machine learning today is probabilistic, unlike traditional software which is deterministic.

   Thus, all predictions of non-AI software (henceforth called NAIS) is exactly computable before the prediction is made, while an AI-infused software (AIIS) predicts based on the model, the current incoming data and subsequent reinforcement - which can make for a unique permutation of predictions everytime the computer is run. In other words, the probabilistic nature of AAIS code means for the same input, model and output - its behavior can vary as it learns, relearns and integrates reinforcement - unlike that of a traditional software, whose behavior is static, predetermined.

2. *Exponential Moves:* The Turing Test, which has so far been used as a benchmark measurement of machines exhibiting intelligence, is overwhelming predicated on machines “mimicking” human actions and behavior. However, one drawback of this benchmark is that it overlooks intelligence that is tangential to human understanding of intelligence (e.g. Move 37 of AlphaGo [1]). At best, it reduces the concept of machine intelligence to facets only observed/ comprehended via the “elements” of human intelligence. But since we still understand very little of how the brain works, such a framework fails to capture the exponential nature of learning ability and expansion of prediction possibility that AIIS can theoretically possess. This is consideration of Question 11 of the RFC.

   Once again, such exponential deduction obfuscates the complex rationale behind it, because at the time of analysis, a human might not possess enough information - which becomes available only on hindsight. Thus, at that time, analyzing AIIS code will not help understand whats happening and why. This behavior and divergence from traditional digital software is cause of concern for many.

3. *No Classification or Standards:* There is a severe lack of standards or classification that defines or calculates what constitutes an “AI” in reality, leading to AI becoming a buzzword. Because everything comes labeled as AI for hype and funding, it is hard to separate the prowess of a chatbot, from a medical device involved in critical diagnostic,

to the software running in self-driving cars. A leveling system needs to be defined so comparisons are clearer. This corresponds to Question 1 in the RFC.

4. *The AI effect:* Another reason why it's hard to measure the improvement involved in calling something AI invention is the presence of “AI Effect” - a queer phenomenon where the goalposts of what the world considers artificial intelligence has been constantly moving. The moment we discover something incredible that machines couldn’t do, but can now do through AI - the world pushes the benchmark for what “AI” is - i.e. the definition of an AI task moves further up the sophistication ladder. Culturally, the most effective AI’s are pervasive and become common tools (and perhaps an utility) to human life. Examples would be Apple’s Face ID or Google Search.

5. *Fog of AI:* Finally, another cultural effect that's widespread is the “Fog of AI”. It is hard to pinpoint errors, debug, validate or verify AIIS because of the above points. This leads to people blaming AI as harmful when its action deviates from the intention of its creators, e.g. the MIDAS system [2] or YouTube recommendations [3]. The Fog of AI can be lifted by detailed measurement of the constituent elements in building an AI.

Having asserted that an AIIS operates differently NAIS, both scientifically and culturally as perceived by its operators, we know delve into a method to understand AIIS, and by extension, understand the “delta” of improvement - or innovation - in newly proposed AIIS inventions.

Synthetic Intelligence Analysis

Protofect has developed a system called *Airate* - that focuses on eight different “elements” in the production of AI, which we call the “AI Stacks”. In the same spirit as some of the questions raised by USPTO, e.g. the importance of “Problem framing” or the “Data lake/ Database”, Protofect aims to *measure and quantify the* factors/dimensions in such stacks, together with *recording and monitoring* how AI builders employ these elements within the stacks. While data is in a lower stack, higher stacks include Model and steps for Optimization of that model, concluding in Explainability or AI-human interaction.

With *Airate*, it is possible to break down any AIIS into 8 stacks covering 67+ dimensions, giving us an effective way to measure the components of the AIIS, and by extension - compare the delta of improvement over previous and existing AIIS or NAIS.

In other words, it is in splitting the elements/stack into dimensions that real “value” of an invention can be realized, and it unravels the “lift” compared to prior art. For example, a new invention can improve on how an AI system will automatically detect if its training data is biased. Another invention might describe a system that can deal robustly with adversarial incoming data. Yet another, can describe how an AI system behaves when there is some missing data. All three examples described above come from the database stack of AIIS. This delta of improvement in the proposed AIIS invention, called the *protolift* would also be a recommendation for Question 11 of the RFC.

The governing theory of Airate is our founding principle of Synthetic Intelligence Analysis - observing, unravelling and deducing the threads that hold a AIIS together. When a new thread is proposed as AI invention, we can look at the current state of existing threads, and *measure* the delta of improvement (Fig. 1) - thus identifying if the new proposal beats the threshold. This corresponds to Question 1 of the RFC on the “elements of AI inventions”.

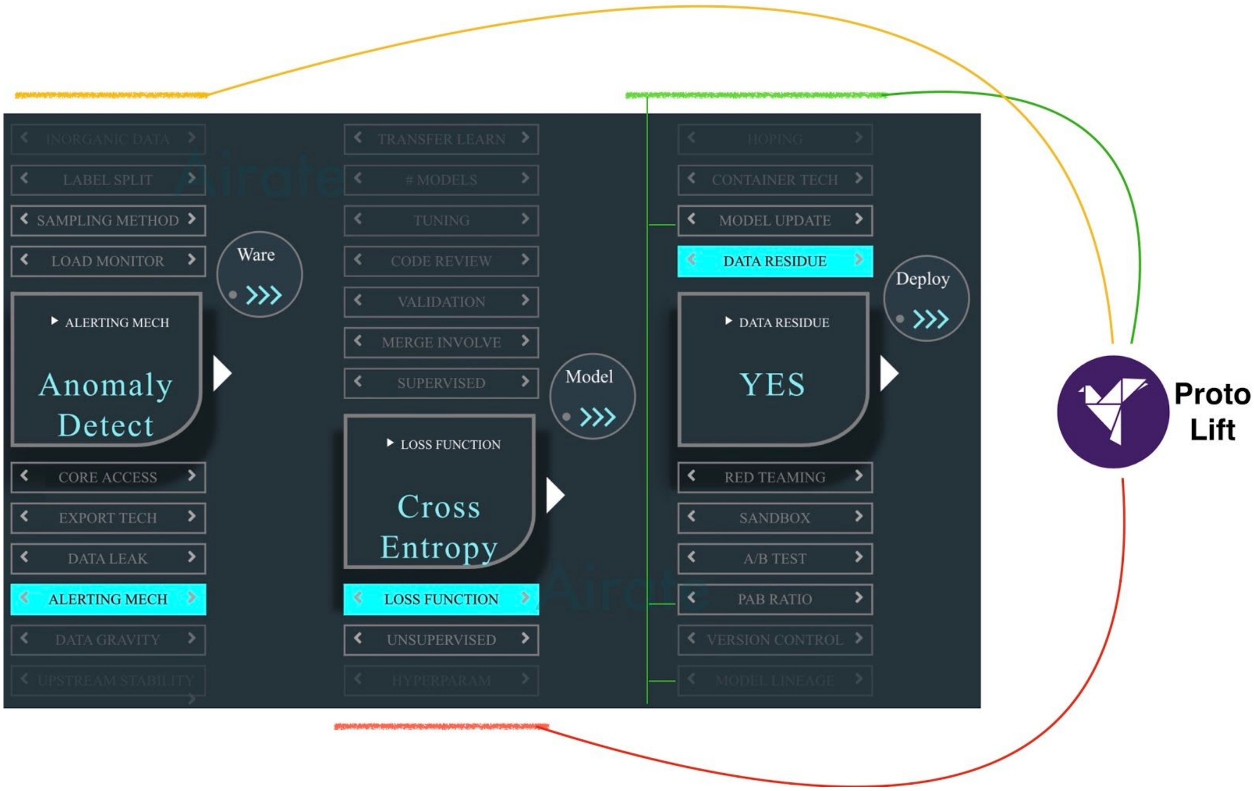

Fig. 1: Description of Airate monitoring AI Stacks and measuring Protolift

What is the Novel? What is Non-obvious?

What stacks or dimensions account as creative expression in AIIS? How can we distinguish if a certain proposed machine learning task is what’s generally the best practice in that layer/stack of AI’s. Synthetic Intelligence Analysis thus helps us quantify a largely subjective phenomenon in the present world - the determination of code that is natural vs. code that is creative [4]. This is crucial for understanding the propensity of the invention from the point of view of the computer, the user and the system architecture itself. It can liberate us from giving away broad patents that hinder the field’s innovation.

We believe that determination of whether an AI invention is natural/obvious vs. creative and artificial *can be measured* by studying the parameters in the AI stack. Because AI is a high-tech gadget, there is a chance that it would become easier to describe an invention which is non-obvious. This might lead to incremental improvements requested to be counted as AI invention. However, my adjusting the benchmark of *protolift,* we can choose the minimum delta of improvement necessary to be counted as "creative invention" - so incremental improvements can be judged accordingly. This corresponds to Question 12 of the RFC.

We thus kill two birds with one stone in *Airate*. We can quantify what is novel with *protolift,* and we can tune what is "non-obvious" in light of prior art by focusing on the right dimensions of the AI stack.

An Interesting Parallel: Plants

We can draw an interesting parallel of building artificial intelligence products to the process by which we judge asexual cultivation of plants. In 1930, the US began granting plant patents to protect distinct, new varieties of asexually reproducible plants (i.e., plants that can be reproduced without seeds, such as by budding or grafting).

Artificial Intelligence is at a similar stage of emergence due to the following parallels:

1. Plants are natural, but they are asexually developed/ evolved. Similarly, intelligence is natural (in humans and many animals), but now we are developing it, artificially.
2. There are key processes in invention of plants, e.g. grafting and budding. These are considered "asexual" methods of propagation, i.e. not via seeds or pollen. Similarly, there are key processes in invention at AI, e.g. data wrangling, feature extraction, modeling, hyperparameter optimization etc. These are spread across different parts of the AI stacks, and are critical in building the "artificial" intelligence.
3. Asexual reproduction is the cornerstone of plant patents because that is what proves that the inventor (or discoverer) can duplicate the plant. The AI tasks mentioned in the previous point are the cornerstone of data science work, and is what proves the builder can replicate the result.
4. Propagation by seeds is not patentable. In AIIP, we can draw an analogy to product-generated data, i.e. the data generated post-deployment. If someone uses the data generated by an AIIP "alpha" and builds a product "beta", you might not be able to claim rights to alpha because you made the beta AI.
5. Finally, and perhaps most importantly - people can still claim a "utility patent" to plants *separately*. To successfully obtain a utility patent, the plant must be made by humans and must fit within the *statutory requirements* (utility, novelty, and nonobviousness). The patent must describe and claim the specific characteristics of the plant for which protection is sought.

We have little framework to understand the statutory requirements of AI currently. Is it aligning with business goals but sacrificing societal welfare [5]? Is it being unintentionally weaponized? Is it filled with fake inorganic data? We need more research on the negative utility of AI and AI alignment. Therefore, for now, we suggest that the utility of an AI should be a separate category from the AI inventions, just like the utility of a plant is separate from a plant patent.

The AI Inventor

1/ Threshold Inquiry: We must determine whether the combination of the pieces, parts and functionality found within the AI stack can be considered to be within the "common sense" of one of skill in the art such that the invention is merely not a trivial rearrangement of what is already known to exist. One way to combat an obviousness rejection would be to show that *Airate* dimensions of the proposed technology are significantly different than previous stacks.

We have to focus on prior art at both the structural and functional level, e.g., not just whether the invention uses a different dataset, but for the same dataset, does it use a different method of anonymization or a different method of sampling.

2/ Disclosure: What level of detail should AI inventors share? We should look for two overarching themes here: (1) Reproducible / Repeatable without undue experimentation and (2) Best/Suggested mode of operation of the AI. This is crucial to Question 6 of the RFC.

One issue to acknowledge here is that most AI systems are self and life-long learning systems. This means the parameters (weights, layers) are probably going to change over time, with a human in the loop or automatically. However the ecosystem at the time of filing the invention should be described, especially if it claims the system does better by changing a parameter and that is proposed as the significant factor for improvement. Arguably, an AI "instrument" will also change its parameters at will over time. But a time-frozen status of its parameters should be described, not in an academic paper technical level, but enough to measure the delta of improvement over existing solutions.

For example, if every dimension of a data pipeline is the same as some existing invention, but the improvement in performance is due to 10x more layers - then prior information about layers should be made clear. On the other hand, if there's an element that is introduced into the pipeline that is applicable to higher stack layer e.g. a new way for "explainability" or "alignment" - then the number of layers are immaterial, so it can be left out.

The other consideration is while some elements of the invention is made clear, others are held as trade secrets. Tuning a model is quite a profound tasks and often happens after sufficient reinforcement time - and thus I expect pushback from inventors if asked to reveal *all* the model parameters.

3/ Comprehensible Technical Considerations: Following up from question 6, this pertains to Question 7 of the RFC i.e. a requirement that ensures the disclosure must explain enough about the invention so that someone skilled in the art can both make and use the invention.

What are the components of an AI invention. Once again, here we can focus on the very dimensions in which *Airate* allows AI's to be classified in. AI experts are usually knowledgeable in all these dimensions, having previously invented some of them. The invention should focus on whether it improves a certain element/module's working, or whether it provides a substitute solution to transform the module.

For example, a dimension or element could be improving the explainability of the algorithm. Even further, an improvement could be the way in which the AI explains itself - perhaps not just via traditional methods but by question/answer or interrogation. The dimensions should not be extremely low level detail, but not too high level either.

Again, drawing compared to plant patents, The USPTO will only grant a plant patent if the inventor provides a full and complete botanical description that explains how the plant is unique and includes drawings showing the plant's unique features. Airate provides the tools for *synthetic intelligence description*, and could help the inventor provide a full picture of the AI.

4/ The Level of the AI: Currently, we have no way to classify what constitutes a certain level of AI. For example, is the Natural Language Processing (NLP) technique thats used in a shopping site chatbot equivalent to an NLP technique used in an AI that interacts with hospital patients following a diagnosis. They need different levels of explainability, and serve different purposes. How do we compare two identical AI stacks, when one solves a mission critical task with lives on the line, while another serving ads on an internet website?

5/ *Verification and Validation:* The performance of an AI system, in its current form, can sometimes be hard to replicate, and thus verify. With the same inputs, same environment, will the AI system give the same output - depends on whether its deterministic. Thus, it can be complicated to test if the system will always give the same output with exactly the same past of inputs. However, replicability and data residue analysis is slowly edging its way into machine learning - meaning in the future it will be easier to make such corresponding analysis.

6/ *Contribution:* While there are many elements that contribute to the usefulness of an AI product, the contribution of a person or a team, is usually focused on pre-deployment.

This includes people responsible for collecting the data, shaping it for training, choosing a model, optimizing the model, monitoring and hyper-parameter optimization that model, verifying the results, validating the robustness of the model and protecting the model pipeline from mischievous actions. This also includes people who discover unexpected behavior from existing model behavior, and finds a way to improve it. This can help answer question 2 of the RFC.

7/ *Convincing Phostia:* The judgement of non-obviousness is critical in the case of AI inventions. This is because of three reasons: (1) The term AI is often used as an umbrella term - leading to it becoming a buzzword, and there is a significant “Fog of AI” that fails to distinguish what’s really innovative work, from what's merely a different result due to changing some parameters. (2) The depth of technical knowledge is zoned in on specific parts of the stack. For example, some are experts in data sampling, while others in hyperparameter optimization. (3) Further, the final goal of the model is key: a model’s output on NLP episodic memory with 50% accuracy can be a huge step, whereas that's not necessary high enough for finance applications. What this means is the benchmark of “high-fidelity” is different for various sectors. This corresponds to Question 8 of the RFC.

We need a Phosita in Graham Factors of AI which provides flexibility, adaptability and keeping up with such rapidly shifting “benchmarks” for what considers innovative can be challenging. We have to be nimble and fluid in thinking of what passes the grade for “invention”, because the AI effect can push it too.

Define Utility Level AI

There are two obvious possibilities in the future: (1) An AI discovers another AI (e.g. drug discovery). (2) Some external company AIIS was used to build a new AI.

For the first question, could IP be provided to an AI software purely, or must it go to original founders of that software, by the transitive property? Answering Question 4 of the RFC, we do not recommend at this time that synthetic intelligence be considered for an “inventor”. AI’s can run a million simulation of a scenario before choosing the one that optimizes the output function, however, in that case every simulation could possibly become a candidate for “invention” - quickly degenerating into intractability.

For the second issue, consider the question: does a pharma company provide Apple with invention-sharing, purely because the software was run on an Apple computer to discover the particular drug. Similarly, if 90% of AI products are being built on cloud platforms, does it make sense to include the cloud platform as an “inventor” — because it serves as a utility for the AI building companies.

I think this hinges if the AI being used is categorized as utility. If so, it cannot be a contributor to an invention. It is important to measure what makes an AI a utility service - perhaps depending on the scale at which it helps humans and other AI to perform common tasks. This corresponds to Question 2 and 3 of the RFC.

We thank USPTO for encouraging and taking the lead in getting the community together for one of the most important discussions of our times. If we think about it, *intelligence* is the quality that takes us into the future we can't grasp.

It is critical that when we take steps to develop intelligence by artificial means, we deploy proper protections, oversight and monitoring, but at the same time not stifle invention and progress. And we must measure the properties of AI inventions. We believe this is one of the greatest moments in western history, on how we move into the information age powered by artificial intelligence.

For questions/comments and getting in touch, please email roy@protofect.com

References:

[1] https://www.huffpost.com/entry/move-37-or-how-ai-can-change-the-world_b_58399703e4b0a79f7433b675
[2]https://spectrum.ieee.org/riskfactor/computing/software/michigans-midas-unemployment-system-algorithm-alchemy-that-created-lead-not-gold
[3] https://www.wired.co.uk/article/youtube-for-kids-videos-problems-algorithm-recommend
[4] https://majadhondt.wordpress.com/2012/05/16/googles-9-lines/
[5] https://datasociety.net/output/governing-artificial-intelligence/