White paper

Building AI into your applications

Staying on track and choosing the right goals for your AI efforts is crucial to its success

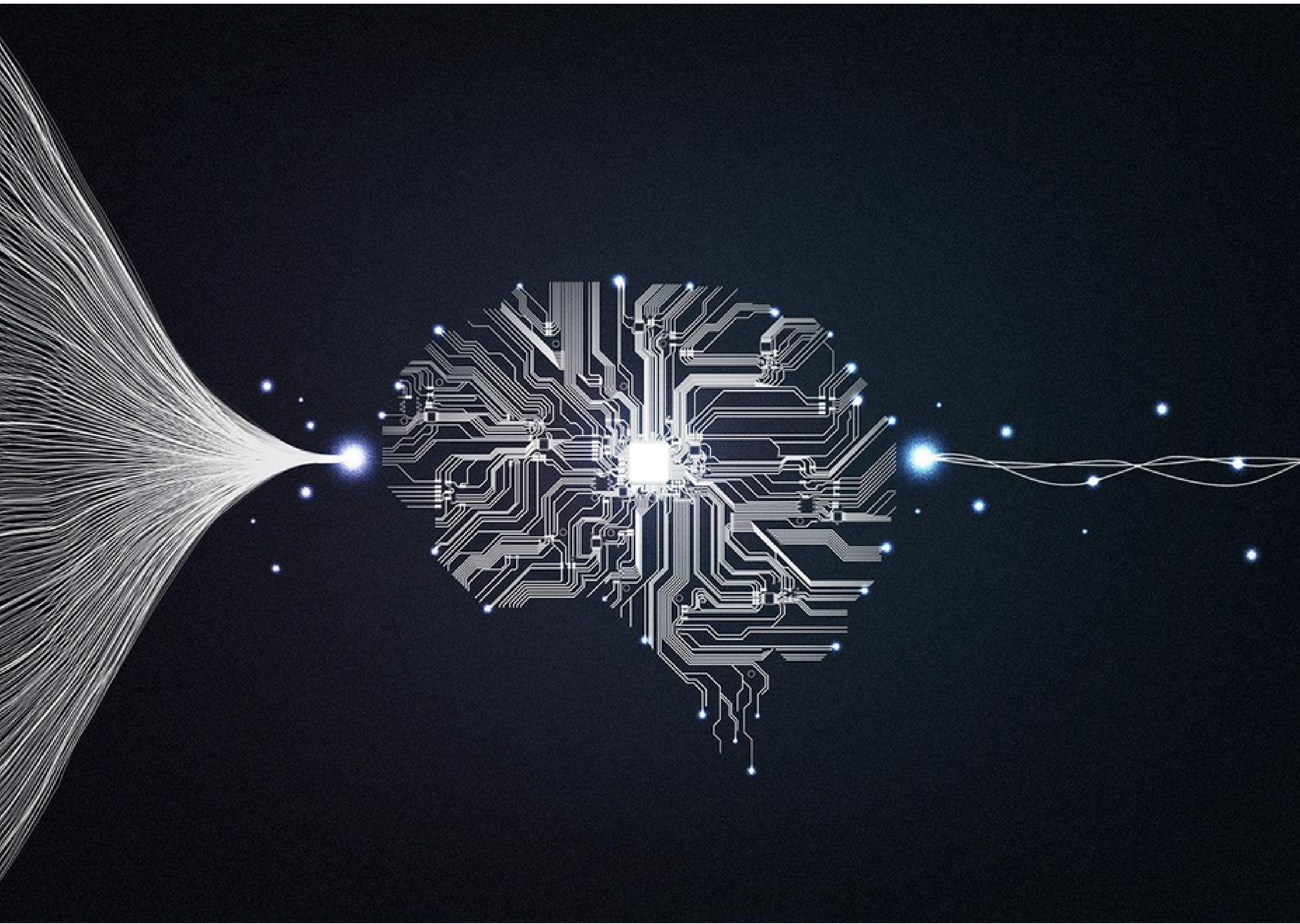
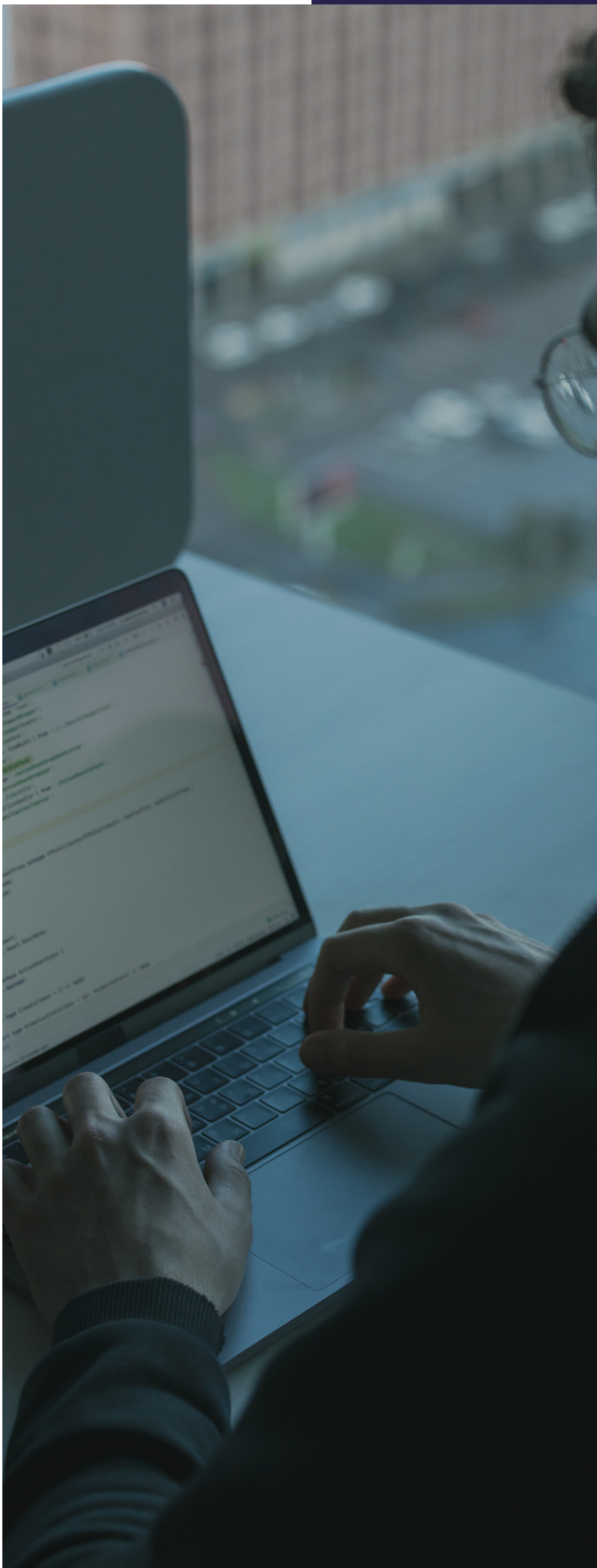


Table of contents

Executive Summary

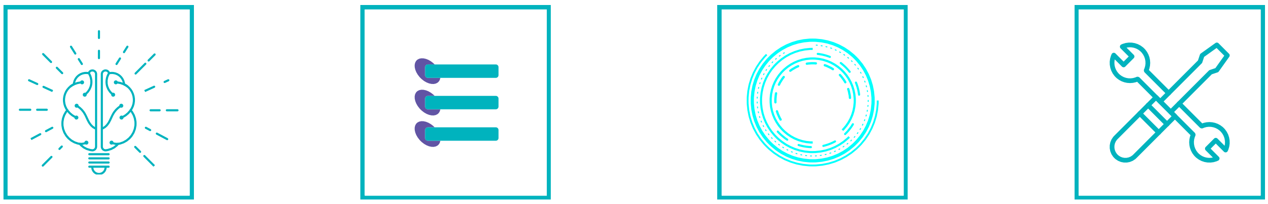
The document is a comprehensive guide on integrating Artificial Intelligence (AI) into applications. It outlines a four-stage AI design process:

- **Intelligence**, where AI's expected behavior is defined
- **Business Process**, involving strategic and operational planning for AI integration
- **AI Technology**, focusing on intellectual property and data strategy
- **Tinkering**, which covers software development and addressing AI challenges. The whitepaper emphasizes the iterative nature of the design process, the importance of precision in each stage, and the necessity of adapting business processes to leverage AI effectively.



Phases of the AI infusion process

Let us dive into the four phases of the design process, each phase representing an activity. These activities can feasibly occur simultaneously or in a order different.



INTELLIGENCE | BUSINESS PROCESS | AI TECHNOLOGY | TINKERING

- **Intelligence:** During this phase, we pinpoint a behavior we anticipate artificial intelligence to exhibit.
- **Business Process**: The second phase deals with finding processes ripe for AI and addressing the strategic and operational issues.
- **AI Technology**: This phase pertains to choosing the right intellectual property approach and data strategy.
- **Tinkering**: Tackle your software development challenges and AI cancers in this phase

Exercise

Identify some AI products you would like to design. For each of these design ambitions, keep a log throughout with you to identify how "what" you're learning relates to each of them. By the end of this whitepaper, you will be able to organize the design effort into four activities.

Phase 1: Intelligence

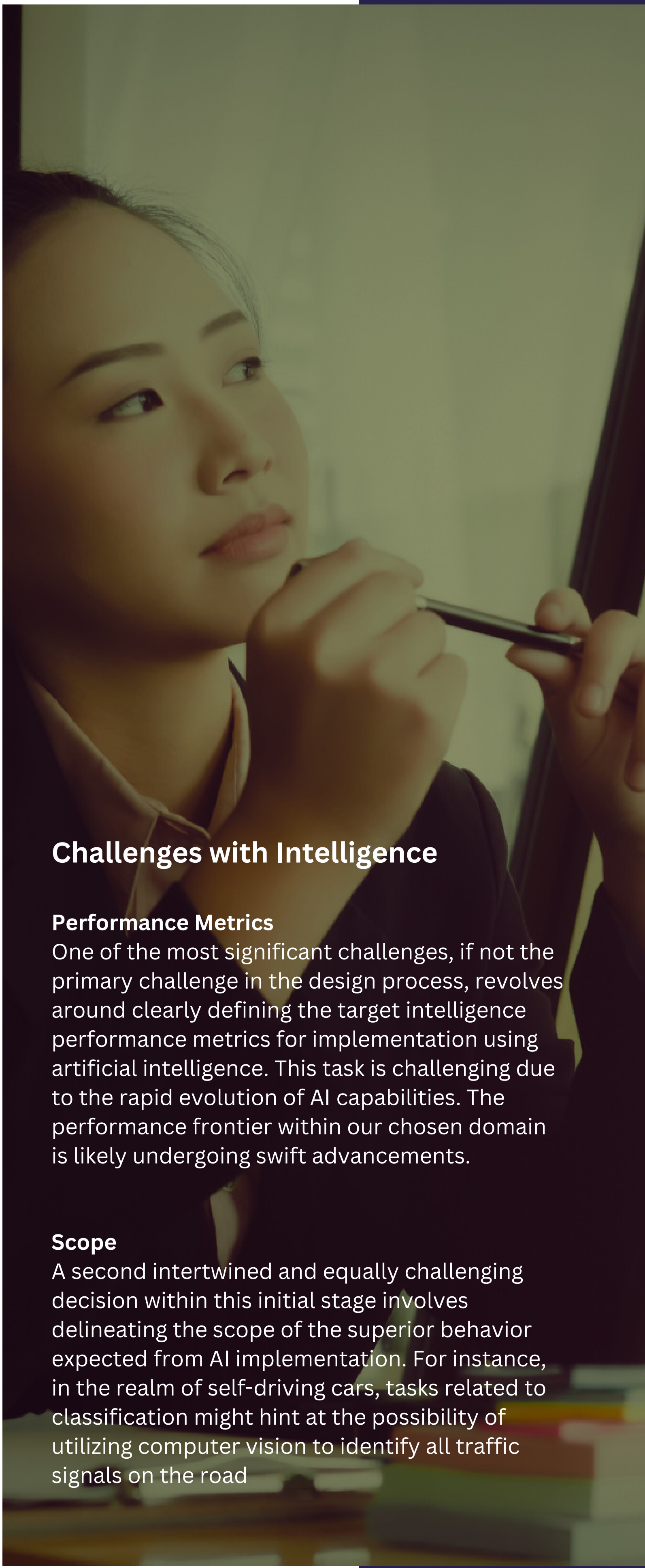
What we expect the AI to do

During this stage, we pinpoint a behavior we anticipate artificial intelligence to exhibit. Essentially, it involves extracting the innate intelligence that we intend to eventually replicate in an artificial manner.

During this first stage, precision regarding the crucial tasks for artificial intelligence is crucial. However, being overly specific might not be necessary until you've navigated through the subsequent stages, particularly if this is your initial engagement with such considerations. In certain scenarios, such as self-driving technology, the models and approaches chosen might limit the scope of the implemented AI.

Choices might shift focus from traffic signals to identifying pedestrians or bicycles, further employing infrared cameras to differentiate warm objects like people from inanimate entities. Integrating these complementary tools can mitigate errors and enhance performance, addressing issues that persist in tasks like human identification in ImageNet cases.

Therefore, this first stage necessitates addressing these two intricate choices: firstly, selecting performance metrics while comprehending the evolving performance frontier in the chosen domain, and secondly, defining the scope of AI's intended accomplishments. Revisiting the four stages in the design process might be required more than once, especially as mentioned earlier. Hence, at the outset of the design process, it's acceptable to be somewhat generic, revisiting and refining after progressing through stages two, three, and four. Subsequent iterations often equip individuals better to tackle these choices of metrics and scope. Moreover, subsequent iterations might prompt a pivot in original thinking based on insights from the initial iteration, akin to how many self-driving companies adjusted their goals by focusing on detecting pedestrians or bicycles in specific environments, especially urban settings where accidents are more prevalent.



Challenges with Intelligence

Performance Metrics
One of the most significant challenges, if not the primary challenge in the design process, revolves around clearly defining the target intelligence performance metrics for implementation using artificial intelligence. This task is challenging due to the rapid evolution of AI capabilities. The performance frontier within our chosen domain is likely undergoing swift advancements.

Scope
A second intertwined and equally challenging decision within this initial stage involves delineating the scope of the superior behavior expected from AI implementation. For instance, in the realm of self-driving cars, tasks related to classification might hint at the possibility of utilizing computer vision to identify all traffic signals on the road

Phase 2:
Business Process

Finding processes ripe for AI

There are two key issues that need to be addressed, the strategic one that involves deciding how will AI play out in our long-term sustainable advantage, and the operational one that is specifying the business process where AI will intervene in setting performance targets for such process, including not just AI but other collateral assets.

01

Strategic
Decide how you will use AI to compete in the marketplace. That is, what is the role that we want AI to play in your design process? there are three ways to compete in the market:

- Being the best player
- Customer solutions player
- Network externalities

Strategies to compete in the market

What role do you want to play (as per the Delta Model by Professor Hax, MIT)

Being the best player
In the case of AI, this means having the best technology possible. For example, the best fingerprint recognition software. Anyone interested in the applications that can be based on fingerprint recognition, we want to use your product, if they can afford it.

Customer solutions player
In the case of AI, this means providing a product that incorporates all AI and all the associated features that will make it a useful product or service. For example, a security and access control solution for buildings that maybe your company already has, that now incorporates fingerprint recognition software and other functionalities required by customers such as locks, intrusion alert mechanisms, connections with the police, 24X7 recovery services, mobile app unlocking, etc. You may not have the best fingerprint AI in the world, but you're offering a service that is useful to a certain segment of customers that therefore will buy it.

Network Externalities
This strategy calls into focus the building of the largest user base possible in a way that results in user getting more benefits as the size of the database increases. For example, you may focus on building a nationwide database of fingerprints so that police can quickly identify people in criminal scenes. The more municipalities and states that participate, the more likely it is that the database is useful. You may need AI for fingerprint recognition, but you don't need to have the best one, and you don't need to provide a full solution to the police.

Phase 2: Business Process

Finding processes ripe for AI

02

Operational
What is the business process that will benefit from AI, and what are sensible improvement targets for that process, and what are things that you need to address within the process, not just in AI? The idea here is to focus on the process that will be impacted by AI so that it can give the engineers adequate information on what is expected of the AI implementation. We need to define a business process and how AI will improve its performance targets.

If you apply it to answering questions in a call center, you want to lower uncompleted calls by 80% in a call center at the same cost so that the responses that the system give have improved and there's less issues with those answers. Or you may want to lower legal translation costs by 80%, maybe because you produce a first version of the translation and then the lawyers only have to review it and they focus on that tricky places. Or you may want to add voice commands to a webpage. Very few people talk to webpages, but you may want to add voice recognition in the webpage so that you can navigate with voice.

In summary, what we need to do in the second stage is to address these two challenging choices. First, the strategic use of AI as a best product, full customer solution or as creating network externalities. And second, the operational description of a process, including non-AI requirements and specific process performance metrics where AI will help. That is the AI without the AI.

Pegasus One

Phase 3:
AI Technology

In the phase dedicated to AI technology decisions - there are pivotal choices to navigate.

In this third phase, the initial decision pertains to choosing the appropriate intellectual property approach, while the second revolves around selecting the right data strategy. Let's delve into these choices, centered on IP concerning technology and data integration.

01

Intellectual property approach
Considering the insights gleaned from stages one and two, you've acquired an understanding of how AI will augment your business strategy. Now, the task is to hone in and select a specific AI technology for product integration.

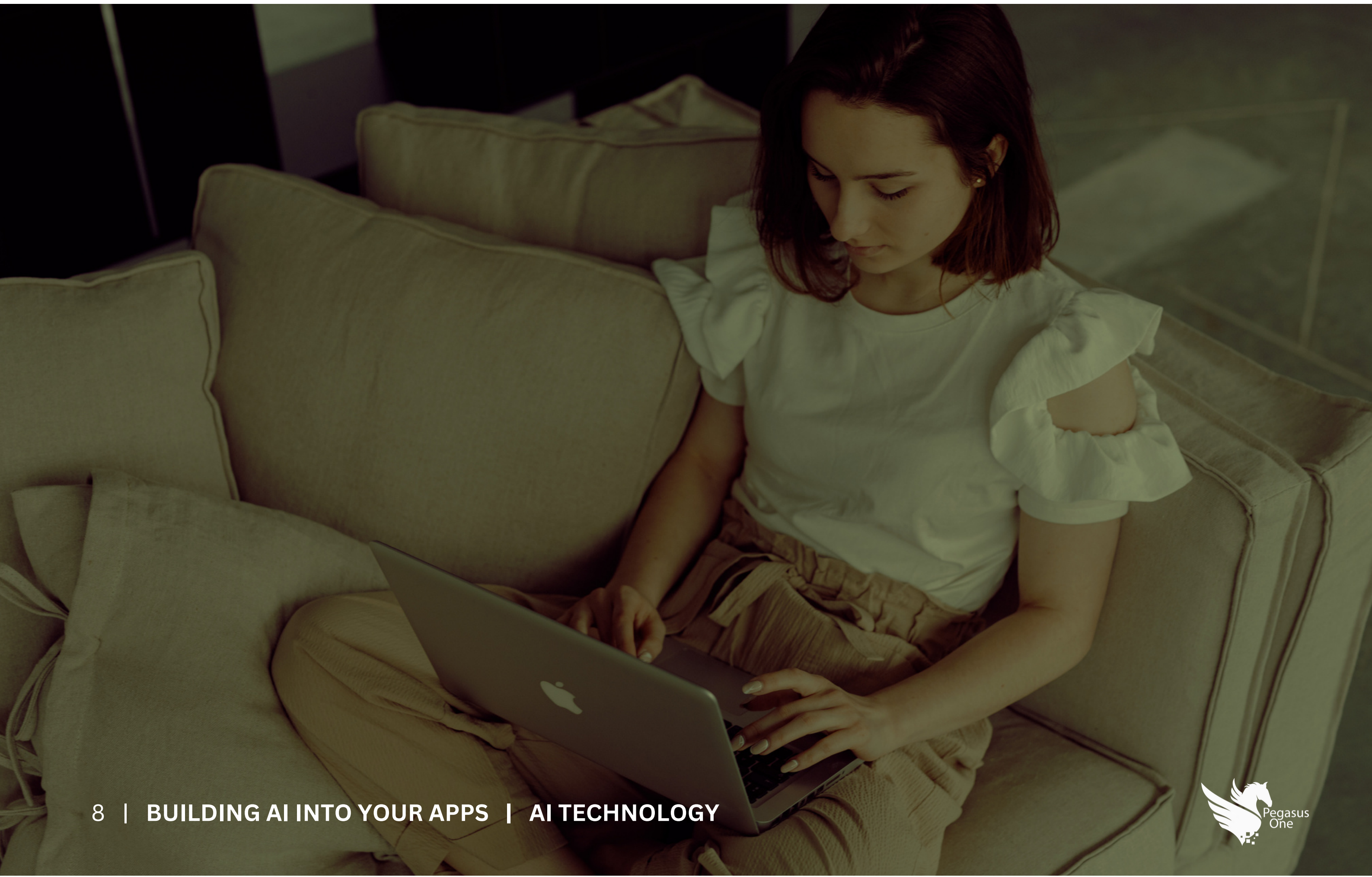
Choosing a specific AI technology presents challenges due to its broad spectrum, offering countless options - possibly thousands or tens of thousands.

To contextualize the abundance of technology choices, until 2012, the volume of AI-related patents remained relatively modest. However, since then, there has been an explosive increase, with the U.S. alone granting thousands of patents annually. This dynamism exhilarates AI product designers with a constant influx of innovations. Yet, this also poses risks, as a chosen technology might become obsolete within months.

Amidst this proliferation of choices, the absence of off-the-shelf alternatives further complicates decisions. However, the upside lies in the myriad of new AI opportunities, potentially enabling you to patent your designs. Incorporating this aspect into your IP strategy is crucial.

Hot take

If your AI application is novel, useful, and non-obvious, we strongly advocate considering patenting.



Phase 3:
AI Technology

In the phase dedicated to AI technology decisions - there are pivotal choices to navigate.

02

Data strategy
The third stage necessitates a critical decision regarding the chosen data strategy. A profound link exists between the available data and the potential algorithms that can be employed. While algorithms correspond to your intellectual property (IP), the data aligns with the data strategy, with data labeling emerging as a pivotal concern. Labeling, also known as metadata, encompasses vital information such as date, time, collection method, and content description. Insufficiently labeled data or a lack of associated metadata limits your capabilities, potentially confining you solely to unsupervised learning.

The data strategy's dynamics often diverge from machine learning algorithms. The abundance of data and metadata broadens the scope of available options. Consider scenarios where AI algorithms didn't previously exist for certain datasets; this changed with developments like facial recognition in photo libraries, transforming their value by enabling automatic organization based on individuals. Similarly, vast amounts of text, previously untapped by AI, were utilized by GPT-3, repurposing seldom-read text collections.

Notably, FAANG companies (Facebook, Apple, Amazon, Netflix, Google) possess data strategies that leverage network externalities, providing them with ample resources to explore machine learning methodologies. Their shared approach results in a competitive edge due to the wealth of data. The saying "data is king" or "data is gold" aptly captures its significance.

The value derived from data stems from three interconnected forces nested within each other. It affords companies adapting their AI strategies a sustainable advantage through superior data utilization.

Network externalities, elucidated in phase two, demonstrate their impact well.

For example, let us consider facebook. The user's perception of a social network's value correlates with the user population. The network with a majority of a user's social connections, like Facebook with 95%, *holds more appeal*.

This concept extends beyond AI, analogous to the value of a phone network increasing with more users.

In Facebook's case, the extensive data pool augments their AI algorithms, particularly in advertising, elucidating their swift acquisition of networks like WhatsApp and Instagram.

Phase 3:
AI Technology

In the phase dedicated to AI technology decisions - there are pivotal choices to navigate.

Path towards AI data monopolies comprises three components:

- **Externalities:** Channeling users towards larger networks
- **System lock-in:** Enhancing value per user
- **Economies of scale:** Amplifying the asset's worth for entities targeting network users

System lock-in constitutes the second force, elevating switching costs within FAANG services. Users' extensive history within applications, encompassing contacts, media, and configurations, complicates transitioning. This effect extends beyond end-users, impacting advertisers, developers, and third-party applications tied to the ecosystem. It solidifies dependence, making transition costly and cumbersome.

Externalities represent the initial of these three interconnected dynamics. We previously addressed this while discussing phase two as a potential approach of utilizing AI. Consider Facebook as an illustration. The significance of a social network to a particular user depends on the network's user base. If two networks offer comparable services, yet Facebook boasts 95% of a user's social connections while the other has only 1%, the user is considerably more inclined to gravitate towards Facebook, assuming all other factors remain constant. This preference stems from Facebook encompassing a larger portion of their social circle.

Economies of scale, the third force, interlaces with system lock-in and network externalities. Companies capitalizing on network effects experience exponential revenue growth. For instance, Facebook's advertising revenue surged from less than a billion to tens of billions between 2009 and 2019. This substantial revenue enables larger R&D budgets for user base expansion or competitor acquisitions. Moreover, a larger user base equates to more data, enhancing AI algorithm proficiency.

Phase 4: Tinkering

Upon entering the fourth phase, you've amassed your AI metrics, scope, strategy, operational targets, IP, and data approaches. It's time for action.

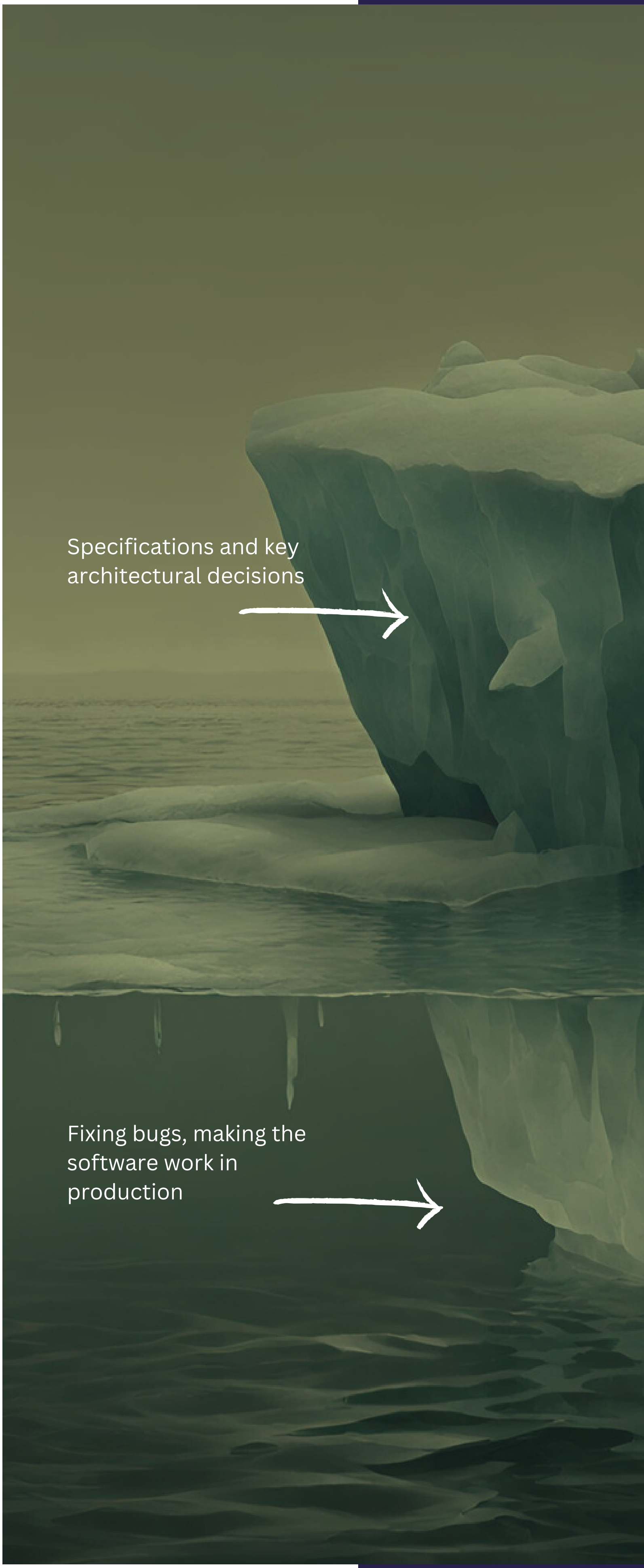
01

Software Development Approach
Engineers need to initiate work on AI algorithms while aligning the rest of the organization with complementary assets.

The technological aspects entail significant effort, akin to an iceberg—the "tinkering iceberg"—with only the more glamorous stages usually visible, like specifications and crucial architectural decisions. Yet, the laborious process of resolving bugs and ensuring software functionality in production constitutes an effort tenfold greater than other technical aspects.

Here, two pivotal choices emerge. The first pertains to your software development approach, while the second revolves around addressing AI issues. The software development strategy mirrors influential choices seen across various businesses. Much like menu selection in a restaurant or franchise locations in retail chains, these choices profoundly impact business outcomes.

Flexibility in pivoting based on outcomes is vital, akin to fine-tuning a business model. Similarly, in AI, adaptability within the software development strategy holds immense significance. Reasons to pivot may range from improved data sources enhancing accuracy or lowering computational costs to adjusting AI responses due to user preferences.



Phase 4: Tinkering

Upon entering the fourth phase, you've amassed your AI metrics, scope, strategy, operational targets, IP, and data approaches. It's time for action.

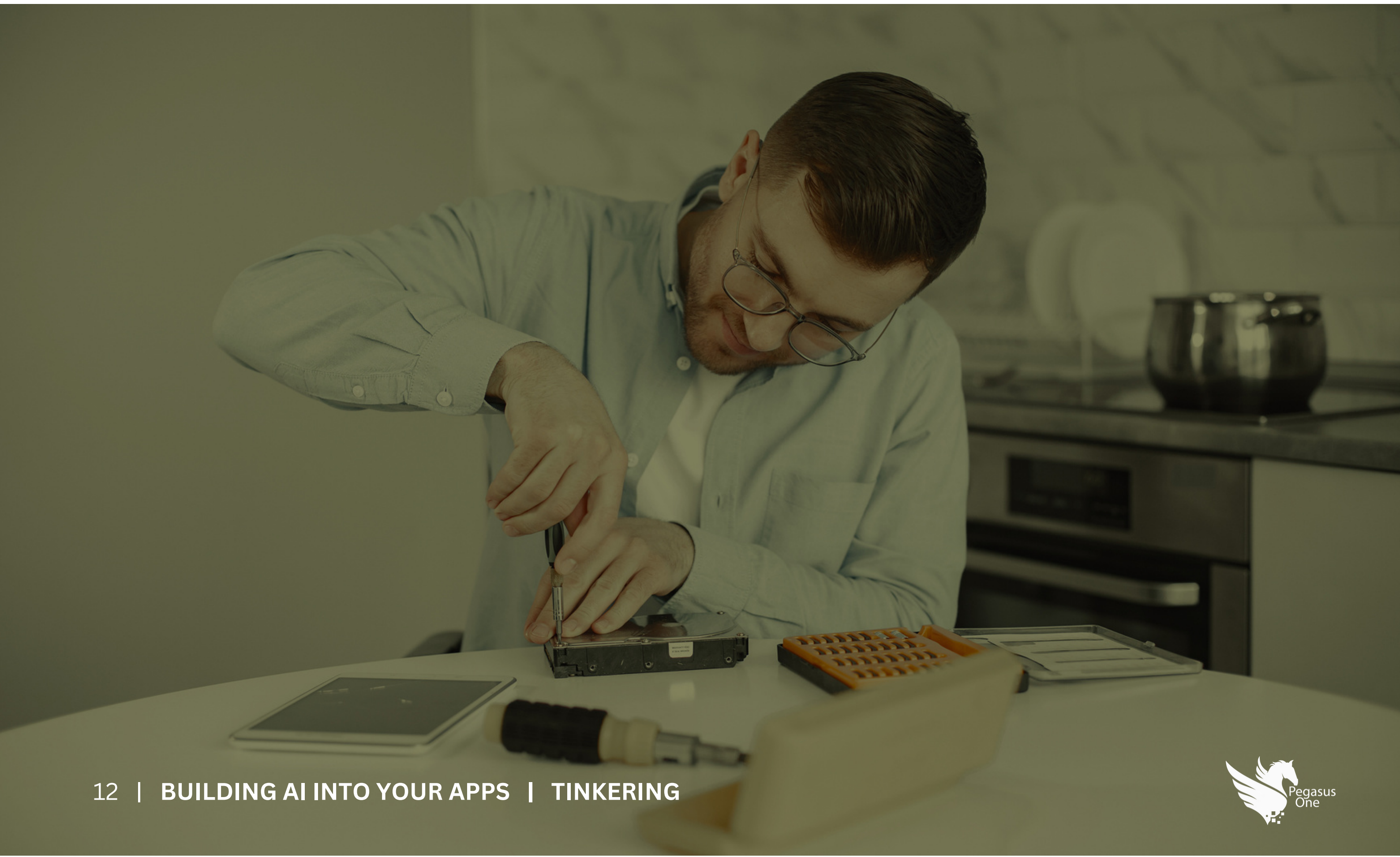
Flexibility must be inherent in the software development process to accommodate the dynamic AI landscape. Incorporating adaptability into the design process remains crucial, given the rapid evolution of AI's frontiers.

While this document doesn't delve into software development, it's imperative to embed a basic software development plan into your AI design process. Establishing a code and data repository methodology, devising a testing and experimental plan, optimizing GPU experimental setups, and outlining a budget justification plan are fundamental preparatory steps before commencing.

The code and data repository methodology entail organizing code versions and documentation, perhaps utilizing tools like GitHub. Equally critical is devising a robust testing and experimental plan encompassing software testing methods and issue management, extending even to factors like business profitability and user acceptance.

For effective number crunching in machine learning, a streamlined GPU experimental setup maximizing computational resources is essential. This setup, whether on cloud platforms or in-house, should be optimized both computationally and financially.

Finally, outlining a budget plan aligned with justifying investments is imperative, considering that software development projects typically exceed estimated costs by two and a half times even with these measures in place.



Phase 4: Tinkering

Upon entering the fourth phase, you've amassed your AI metrics, scope, strategy, operational targets, IP, and data approaches. It's time for action.

02

AI Cancers
The last decision, pivotal within the fourth phase and the overall AI infusion process, revolves around addressing AI limitations and challenges. Despite its prowess, AI grapples with persistent issues that require premeditated solutions. We will delve into three such challenges and propose potential strategies to tackle them.

- Adversial attacks
- Lack of generalization
- Bias within ML algorithms
- Explainability
- Unintended behavior

2.1

Adversarial attacks pose a significant concern. Studies demonstrate the ease with which deep learning AI algorithms can be misled by subtle modifications in input signals.

For instance, a seemingly benign image can be subtly altered to mislead algorithms, leading to potentially severe consequences like misidentifying stop signs or manipulating self-driving cars' interpretations.

Such attacks undermine user security and privacy. Additionally, researchers have showcased activating and controlling devices like Siri with inaudible commands, raising security vulnerabilities further.

Phase 4: Tinkering

Upon entering the fourth phase, you've amassed your AI metrics, scope, strategy, operational targets, IP, and data approaches. It's time for action.

2.2

The second challenge is the **lack of proper generalization** in AI. Often, AI struggles to generalize effectively across various scenarios. For example, translation accuracy between two languages might not ensure similar accuracy when translating to a third language. This issue is evident in face recognition algorithms that underperform with certain skin tones and makeup styles due to inadequate training data. Balancing datasets becomes crucial here to ensure a more comprehensive and unbiased **generalization.**

2.3

Bias within ML algorithms poses a substantial concern. Algorithms frequently mirror biases present in the training data. Past instances, such as biased autocompletions in GPT-2 or recruitment tools reflecting gender bias, exemplify this issue. Ensuring unbiased training data becomes imperative to mitigate algorithmic bias.

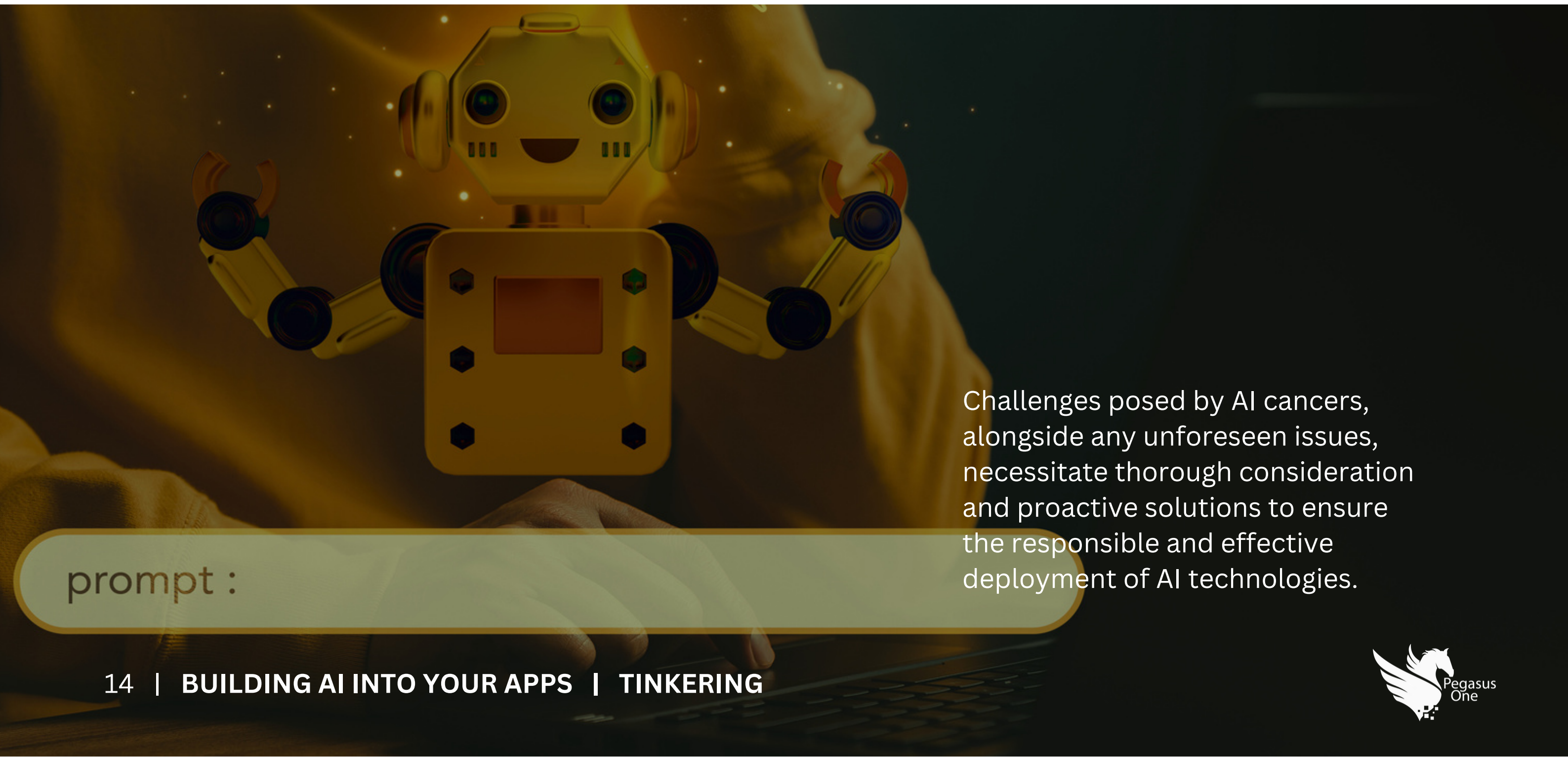
2.4

Another critical concern is the **lack of explainability** in modern deep learning algorithms. These algorithms often arrive at conclusions without providing transparent insights into their decision-making processes. For instance, an algorithm distinguishing wolves from dogs might rely on factors like snow presence in images, rather than the actual subject. The algorithm's inability to articulate its decision-making process underscores the necessity for more explainable AI models.

2.5

Unintended behaviors represent a pressing challenge for any AI system. Instances like a robot causing injury in a shopping mall or an AI chatbot offering alarming advice showcase the unpredictability inherent in AI systems. The complexity arises from scenarios where AI interacts with human actions, requiring meticulous planning to minimize unintended consequences.

Your decision-making process must address these AI challenges and aim to mitigate their impact.



Challenges posed by AI cancers, alongside any unforeseen issues, necessitate thorough consideration and proactive solutions to ensure the responsible and effective deployment of AI technologies.

Importance of AI design process

Rapid evolution of AI possibilities
The swift progression of AI necessitates innovation for practical AI applications. Replicating an AI application proves challenging due to its inherent uniqueness. In contrast to various fields where mirroring a neighbor's choices can yield similar outcomes, such as cars, clothing, phones, education, or even houses, AI stands apart. Attempting to emulate AI strategies without tailored design efforts often leads to suboptimal results.

AI complementary assets
Another key consideration emphasizes the importance of AI's complementary assets. The inherent value doesn't solely stem from AI itself but emerges from the fusion of AI with modifications in business methodologies.

There are instances where a distinctive AI component—such as a custom-made predictive model or a specialized application of computer vision—becomes crucial for standing out in the market. In such cases, formulating a deployment strategy becomes vital.

This strategy should involve the continual improvement of technology through the provision of relevant training data to machine learning algorithms. Simultaneously, it mandates the adaptation of organizational processes and products to capitalize on these ongoing technological advancements. Thus, developing AI products necessitates envisioning how the organization will adapt to the implemented technologies.

Pegasus One

Thank you for reading. You can reach out to us with any queries regarding the document at info@pegasusone.com

Our AI team can help you:

Bring Artificial Intelligence powered Data Analytics to your Cloud

Build custom Generative AI apps or integrate your apps with ChatGPT

Conduct an audit to assess your organisational readiness for AI

Identify opportunities for Generative AI in your organisation

Accelerate your code modernization with AI powered tools

Explore more from Pegasus One

AI Consulting
Case Studies
Whitepapers