# Group 8 – PAI’s Moonshot

|  |
| --- |
| *Assignment:*   * *Answer the following questions :*   + *Slide 1: Is “PAI moonshot” a moonshot?*   + *Slide 2: Why are the tech companies doing the PAI moonshot? Why is it counter-intuitive?*   + *Slide 3: What are the Open Innovation practices?*   *Remark : use the articles to construct the answer but you can also develop answer we ideas that are not in the articles*   * *Each group has to present their answers in 15 minutes max* |

## Falon Fatemi (2020) “How Companies Should Answer The Call For Responsible AI” < https://www.forbes.com/sites/falonfatemi/2020/02/28/how-companies-should-answer-the-call-for-responsible-ai/?sh=1fc1a48d13f5>

There’s widespread consensus that we’re in the throes of the fourth industrial revolution Artificial intelligence and its sister technologies is transforming virtually every business. Yet with AI’s enormous potential comes great responsibility. The majority ([77%](https://knowledge.wharton.upenn.edu/article/want-responsible-ai-think-business-outcomes/)) of CEOs say that AI threatens to increase vulnerability and disruption to the ways they do business. Unfortunately, the call for responsible AI has taken a backseat for many companies. Only 25% of companies say that they definitely prioritize considering the ethical implications of an AI solution before investing in it, according to [research](https://www.pwc.com/gx/en/issues/data-and-analytics/artificial-intelligence/what-is-responsible-ai/responsible-ai-practical-guide.pdf) by PwC.

In 2016, Amazon, Facebook, Google, DeepMind, Microsoft, and IBM came together to found the Partnership on Artificial Intelligence to Benefit People and Society (Partnership on AI). Since it was founded, the nonprofit coalition has amassed more than 100 partners, including members from industry, academia, and civil society. The partnership marks an important shift towards prioritizing responsible AI. But it’s merely a small step in the right direction.

Here are three important steps that companies need to embrace in order to progress towards responsible AI.

**1. Prioritize explainability and interpretability**

An important prerequisite for responsible AI is explainability and interpretability. According to PwC’s research, 84% of CEOs agree that AI-based decisions need to be explainable in order to be trusted. The proverbial “black box” of AI needs to be opened. Black box models should be supplanted with models that are interpretable. As PwC warns, a lack of interpretability can “expose an organization to operational, reputational, and financial risks. To instill trust in AI systems, people must be enabled to look “under the hood” at their underlying models, explore the data used to train them, expose the reasoning behind each decision, and provide coherent explanations to all stakeholders in a timely manner.” As IBM’s AI ethics policy [states](https://www.ibm.com/watson/assets/duo/pdf/everydayethics.pdf), “Allow for questions. A user should be able to ask why an AI is doing what it’s doing on an ongoing basis. This should be clear and up front in the user interface at all times.”

Explainability goes hand in hand with documentation throughout the entire AI lifecycle, from model design to implementation and use. Design and decision-making processes should be documented. Additionally, it should be clear when and why AI systems make mistakes. AI is inevitably going to fail. By making all aspects of AI development transparent, we can empower humans’ judgment to kick in and avoid many negative fallouts.

**2. Make responsibility concrete and pervasive**

It’s easy to pay lip service to the importance of responsible AI. This tendency is so pervasive that we now have a dedicated term to encompass the practice—“ethics washing”, also known as “ethics theater”. Ethics washing involves falsifying or exaggerating a company’s promotion of “AI for good” initiatives. As a recent [report](https://ainowinstitute.org/AI_Now_2018_Report.pdf) from research institute AI Now highlights, “While we have seen a rush to adopt such codes, in many instances offered as a means to address the growing controversy surrounding the design and implementation of AI systems, we have not seen strong oversight and accountability to backstop these ethical commitments.”

While corporate governance bodies are becoming increasingly common, the public is, oftentimes, afforded little insight into the operations and, especially, the interventions enacted by these governance bodies. We don’t know how they are holding themselves accountable. It’s not enough to simply set up an advisory committee. Companies need to be clear on how they are evaluating and debating ethical issues, embrace a more open and encompassing discussion related to these issues, and communicate revelations and insights to the broader community.

The companies that truly place responsible AI on a pedestal will make responsible AI an organization-wide imperative. The Wharton School of the University of Pennsylvania has [emphasized](https://knowledge.wharton.upenn.edu/article/want-responsible-ai-think-business-outcomes/) that responsible AI requires engagement from the entire C-Suite. Ideally, AI engagement goes deeper and pervade an entire organization. Kush Saxena, Chief Technology Officer at Mastercard for example, has called for mandatory ethics training for all AI programmers. Only with organization-wide accountability can we truly move the needle towards responsible AI.

**3. Minimizing bias**

Bias pervades AI tools. Take, for example, software to predict future criminals that is [biased against blacks](https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing) or a recruiting tool used by Amazon riddled with [bias against women](https://www.reuters.com/article/us-amazon-com-jobs-automation-insight/amazon-scraps-secret-ai-recruiting-tool-that-showed-bias-against-women-idUSKCN1MK08G). Companies must commit to mitigating bias. This mitigation shouldn’t be an afterthought—it should guide AI development. Before developing algorithms, AI designers and developers should come together to discuss and identify potential biases in data, identify the potential ramifications of these biases, and then proactively take steps to minimize them.

Minimizing bias also involves constructive dissent, a phrase embraced by Rumman Chowdhury, the Responsible AI lead for Accenture. Chowdhury has [explained](https://venturebeat.com/2019/07/17/how-ai-companies-can-avoid-ethics-washing/), “Successful governance of AI systems need to allow ‘constructive dissent’ — that is, a culture where individuals, from the bottom up, are empowered to speak and protected if they do so. It is self-defeating to create rules of ethical use without the institutional incentives and protections for workers engaged in these projects to speak up.” If people feel empowered with high levels of psychological safety to voice concerns, this will fuel more productive conversations pertaining to responsible AI and, especially, mitigating bias.

Finally, minimizing bias requires diversity. A [recent report](https://ainowinstitute.org/discriminatingsystems.pdf) from the AI Now Institute revealed that 80% of AI professors, 85% of AI research staff at Facebook, and 90% of AI employees at Google are male. Racial diversity among AI researchers and industry leaders is minimal. Only by bringing diverse perspectives and backgrounds to the table will we be able to mitigate bias and build responsible AI tools.

Elon Musk has dubbed AI [more dangerous than nuclear weapons](https://www.cnbc.com/2018/03/13/elon-musk-at-sxsw-a-i-is-more-dangerous-than-nuclear-weapons.html). Safiya Umoja Noble, an associate professor at UCLA, has [called](https://safiyaunoble.com/role-artificial-intelligence-society-next-human-rights-issue-21st-century/?utm_source=VentureBeat&utm_campaign=16d608696c-AIWeekly&utm_medium=email&utm_term=0_89d8059242-16d608696c-4659801) AI one of the major human rights issues of the 21st century. The discussion of AI should be at the forefront of our public and corporate discourse. By prioritizing explainability and interpretability, making responsibility concrete and pervasive, and minimizing bias, we can ensure that the potentially disastrous and fatal implications of AI don’t overshadow the benefits and promises.

## [PAI](https://packagingeurope.com/topics/victoria-hattersley/) (2016),” Industry Leaders Establish Partnership on AI Best Practices”< https://www.partnershiponai.org/industry-leaders-establish-partnership-on-ai-best-practices/>

Amazon, DeepMind/Google, Facebook, IBM, and Microsoft today announced that they will

create a non-profit organization that will work to advance public understanding of artificial

intelligence technologies (AI) and formulate best practices on the challenges and

opportunities within the field. Academics, non-profits, and specialists in policy and ethics

will be invited to join the Board of the organization, named the Partnership on Artificial

Intelligence to Benefit People and Society (Partnership on AI).

The objective of the Partnership on AI is to address opportunities and challenges with AI

technologies to benefit people and society. Together, the organization’s members will

conduct research, recommend best practices, and publish research under an open license

in areas such as ethics, fairness, and inclusivity; transparency, privacy, and interoperability;

collaboration between people and AI systems; and the trustworthiness, reliability, and

robustness of the technology. It does not intend to lobby government or other

policymaking bodies.

The organization’s founding members will each contribute financial and research resources

to the partnership and will share leadership with independent third-parties, including

academics, user group advocates, and industry domain experts. There will be equal

representation of corporate and non-corporate members on the board of this new

organization. The Partnership is in discussions with professional and scientific

organizations, such as the Association for the Advancement of Artificial Intelligence (AAAI),

as well as non-profit research groups including the Allen Institute for Artificial Intelligence

(AI2), and anticipates announcements regarding additional participants in the near future.

AI technologies hold tremendous potential to improve many aspects of life, ranging from

healthcare, education, and manufacturing to home automation and transportation.

Through rigorous research, the development of best practices, and an open and

transparent dialogue, the founding members of the Partnership on AI hope to maximize

this potential and ensure it benefits as many people as possible.

## Sea Matilda Bez and Henry Chesbrough. "Competitor Collaboration Before a Crisis: What the AI Industry Can Learn." *Research-Technology Management* 63.3 (2020): 42-48

Artificial intelligence (AI) technology is promising and controversial. AI is a technological breakthrough in software development—unlocked by the amount of accessible data, increased computation power, and advanced algorithms—that vastly improves the ability of computer systems to make accurate predictions, optimize decision-making, enable voice as an input source for computing, identify and access digital images, and translate between languages. AI is a general-purpose technology able to generate value in applications across many industries (Agrawal, Gans, and Goldfarb 2018). Applications of AI include self-driving cars, cancer diagnoses, robotics, and automated content. Some US jurisdictions have already passed legislation mandating that the US criminal justice system develop and use basic AI tools in sentencing and parole decisions.As Koepke et al. (2018) explain, “In the last five years, legislators in all fifty states have made changes to their pretrial justice systems. Reform efforts aim to shrink jails by incarcerating fewer people—particularly poor, low-risk defendants and racial minorities. Many jurisdictions are embracing pretrial risk assessment instruments—statistical tools that use historical data to forecast which defendants can safely be released—as a centerpiece of reform.”

The AI industry, which includes all the companies that prioritize AI and thus reinvent themselves to benefit from what AI does very well, such as strategic data acquisition or automation, has received several warnings. Tesla CEO Elon Musk likened AI research to “summoning the demon” (Bloomberg 2014, 0:17). Nobel prize–winning physicist Stephen Hawking told the BBC in 2014, “The development of full artificial intelligence could spell the end of the human race” (British Broadcasting Corporation 2014, 3:40). The Army Cyber Institute and Arizona State University’s Threatcasting Lab have identified key future threat areas including weaponized AI, large-scale unemployment from job displacement, and panopticon social control of citizens (Army Cyber Institute at West Point and Arizona State University Threatcasting Lab 2017).

While these possibilities are not imminent, other risks from using AI are already evident. Using AI risk assessment tools in the US criminal justice system to reduce the number of pretrial detentions of suspects may inadvertently increase the disproportionate detention of racial minorities (Partnership on AI 2019; Koepke et al. 2018). Skilled users of AI technology can create “deep fake” video, where a person can be made to say almost anything, and it is impossible to detect the manipulation and fake content (C-SPAN 2019, 42:50). Such risks are the unintentional result of companies like Google or IBM making freely available to researchers, graduate students, and anyone with an interest in machine learning tools like Tensor Flow, an open source software library that makes it easier for any developers to design, build, and train deep learning models, and Apache System ML, a machine learning system that can scale and optimize machine learning projects using big data. Democratizing access to these tools stimulates positive research into AI but also creates possible downsides. AI technology used for a negative purpose by one person or company could damage the entire AI industry and lead to an extreme scenario of AI being forbidden. Tech giants using AI recognize its potential risks and the need for collaboration on a collective response. In 2016, Amazon, Facebook, Google, DeepMind, Microsoft, and IBM created the nonprofit organization Partnership on AI (PAI). Apple joined in 2017. Collectively they committed research resources to enable PAI to study and formulate best practices on AI technologies, to advance the public’s understanding of AI, and to serve as an open platform for discussion and engagement about AI and its influences on people and society (Partnership on AI 2019).

Earlier examples exist where industries realized the need for collective action to respond to serious downside risks from technology (e.g. Browning, Beyer and Shetler 1995. King and Lenox 2000. Macher, Mowery and Hodges 1998; Moffet, Bregha, and Middelkoop 2004; Pronovost and Hudson 2012; Rea, Brooks, Burger, and LaScala 1997). The US nuclear and chemical industries had to collaborate to improve the safety and security of their technology after huge accidents. The US semiconductor industry collaborated to avoid being wiped out by a foreign competitor that could produce higher quality chips for a lower cost. These industries, respectively, have collaborated for more than 30 years. We applied the Dynamic Capabilities Framework of “sensing, seizing, and transforming” to assess how leading competitors in the US nuclear, chemical, and semiconductor industries collaborated to mitigate serious technological downside risks. The lessons learned from these three examples can help the Partnership on AI and the AI industry, as well as any industry with a shared sense of fate and with potentially serious risks from technology.

**Dynamic Capabilities Framework**

The Dynamic Capabilities Framework (DCF) refers to a firm’s ability to integrate, build, and reconfigure competencies to address rapidly changing environments (Teece 2007; Teece, Pisano, and Shuen 1997). More precisely, a firm creates the capabilities it needs to respond to external changes by recognizing a meaningful pattern of opportunities and threats (the sensing capability), making timely decisions to tackle these opportunities or threats (the seizing capability), and upgrading to perform the tasks required to address these opportunities and threats (the transforming capability) (Barreto 2010). David Teece developed the DCF to target the firm level; here we apply it to companies trying to create an overarching organization that responds collectively to technological threats. Due to concerns and technological risks that evolve over time, overarching organizations need dynamic capabilities to sense threats, seize the opportunity to take action, and transform so they can continuously take necessary precautions or actions as new opportunities and threats arise. This three-part framework is a useful way to analyse previous examples of industry organizations and deduce important lessons for new organizations like PAI that aim to collectively raise an industry’s environmental and societal standards. We use the Dynamic Capabilities Framework to analyse INPO, the organization formed by the US nuclear industry, Responsible Care, an initiative used by US chemical manufacturers, and SEMATECH, a consortium created by US semiconductor manufacturers

***Sensing***

Sensing is the first step in building dynamic capabilities. It entails recognizing a meaningful pattern of opportunities and threats related to a technology. INPO, Responsible Care, and SEMATECH realized that not all of their respective industry members perceived the technological threat or recognized the need for collective action that warranted their creation. They sensed the need to provide industry members with education and training about the technology threat and collaboration, although each industry took a different approach.

INPO began by educating US nuclear plant owners about the overlap between the public interests in safety and environmental protection and their private interests. The overlap was not intuitive for all plant owners, and for years environment nongovernmental organizations and nuclear plant owners fought vigorously. INPO advised industry counterparts about how developing safety technology fostered a positive public image, while poor environmental records visible to the public were damaging. INPO also emphasized the importance of maintaining good employee relations: poor behavior by the company could lead to employee turnover, whistleblowers publicly reporting internal problems, and difficulties hiring future talent. A third initiative was to forestall government regulation by committing themselves to the highest levels of safety and reliability, since the government could shut down plants it considered unsafe.

Simultaneously, INPO used the worst nuclear power accident in American history at the Three Mile Island in 1979 as an example to make plant owners realize that actions by one company could tarnish the entire industry at the public, business, and government levels. It was not enough for each firm to be conscious and improve its own behavior—watching all plant owners’ behavior was mandatory. INPO aimed to make each plant owner aware of the “community of shared fate,” which increased each member’s involvement and motivated them to keep watch over competitors’ actions. Given the small number of large players in the US nuclear industry, the watchdog mentality proved effective.

US chemical manufacturers implemented Responsible Care because they knew the threat of technology was impacting them negatively. Their image shifted from miracle provider of scientific products enhancing society’s well-being to a societal danger. This distrust arose due to adverse publicity and the tragedy from incidents like Union Carbide’s 1984 Bhopal disaster, which killed more than 4,000 people and injured hundreds of thousands. One opinion poll conducted in the US in 1990 found that the chemical industry's rating of public acceptability had dropped to 25 percent—only the tobacco industry had a lower rating (Gunningham 1995). The accident in India had also confirmed that one company’s action anywhere in the world could tarnish the entire industry. Industry participants issued a high-level official statement of policy through Responsible Care that they hoped all members would commit to. Unfortunately, having the chemical companies commit to a statement of policy was not sufficient to raise the industry’s credibility or regain society’s trust. After some reflection, Responsible Care realized that its value-added was to have industry members that were manufacturing, using, and distributing chemicals use their knowledge to transform the statement of policy into operational terms called Code of Practices. Responsible Care educated its member companies about the unique opportunity they had to create a code of practices that could be a game-changer for the whole industry and avoid an accident that would further tarnish the industry’s reputation.

SEMATECH aimed to revitalize the US chip making industry and reestablish its leadership position in manufacturing globally. To facilitate collective action between competing semiconductor manufacturers unaccustomed to collaborating, SEMATECH spent time educating its members about the gravity of inaction. At that time, they estimated that US market share of the semiconductor industry would shrink from 85 percent to 20 percent, due to the rise of Japanese competitors that had developed superior process and manufacturing technologies resulting in higher quality semiconductor chips. SEMATECH reinforced the shared sense of urgency that the semiconductor industry was vital to American technology preeminence.

SEMATECH spent time developing its culture and came up with a watchword phrase any member could rely on: “If it's not competitive, change it.” The consortium’s intent was to create a shared sense of a “community of fate” that emphasized the need to change in order to survive. SEMATECH’s decision to spend time developing a culture was initially criticized because each company had its own culture and was unlikely to change or adopt a second one.

***Seizing***

Seizing is about making timely decisions to react to the opportunity or threat by co-creating solutions. For INPO, Responsible Care, and SEMATECH, having industry members convinced of the need to improve safety and environmental practices was not sufficient to prompt voluntary disclosure of internal technology risks, engage others to solve these risks, and then use the solution created. This initial absence of seizing capabilities is not surprising. Okura (2008) identified three disadvantages to a firm of disclosing accident information that, to the firm, outweigh the collective benefits: decreased demand when the risk is revealed, fear of losing competitive advantage by having critical information revealed to the public involuntarily, or fear of free-riding behavior whereby firms want to receive others’ accident information but may not want to share their own. Each overarching organization developed different tactics to quell industry members’ concerns about disclosing information about internal accidents or risks due to their technology, and to prompt members to disclose pertinent information needed to monitor each industry member’s performance and effort.

INPO realized a nuclear plant owner considered it counter-intuitive to share its safety weakness and to give competitors access to its plants. Thus, INPO had to find organizational solutions so a company could speak freely about its risks and safety fears. INPO became a neutral third-party organization composed of an independent, international team of peers with extensive expertise and technical skill in nuclear power generation technology. INPO offered members a robust and confidential evaluation of each plant by its independent experts. INPO’s value-added was the rigorous evaluation and the confidentiality of the report, which allowed the plant’s owner to overcome the fear of letting evaluators onsite. The INPO report also offered ready-to-use solutions derived from observations at other plants and, in the case of unique problems, the expert evaluators and the plant owner co-created a specific solution. Providing solutions created an incentive to give the evaluation team access and openly share the current and future safety issues they found.

Responsible Care collaborated with industry members to create the Code of Practices, but creating the code was not enough because it encountered difficulties ensuring members would use them. Members expressed various concerns to explain why they did not want to commit either to the statement of policy or the code. For example, some members worried about the costs involved, or felt that participation in a collective program would undermine the market benefits of proceeding unilaterally. Some refused to participate out of concern that others would be free riders—they would reap the benefits without participating. To address these concerns, Responsible Care created six regional leadership groups comprising CEOs from member companies in each region. These groups met quarterly to compare notes on their progress, or lack thereof, share their difficulties, and offer expertise. Each CEO had to stand up in front of their peers and highlight improvements, which created peer pressure for any CEO with nothing to say. Thus, each CEO asked its employees at least quarterly for feedback on how they were implementing the code. These meetings quickly became a forum for trading advice and reporting on progress. To deal with potential free riders, Responsible Care opted for the possibility of excluding members from the group, which proved to be an efficient sanction tool. Being excluded from Responsible Care would be an argument against a former member in the event of an environmental or societal legal action against them. The value-added of being part of Responsible Care made the potential sanction of exclusion a real threat.

SEMATECH opted to have industry leaders serve as its leader. Robert Noyce, founder of Intel, was the first president and CEO; Charles Sporck, former general manager of Fairchild Semiconductor, was called the “father of SEMATECH.” They devoted time and effort to create the consortium. The opportunity to collaborate with such widely respected individuals in the industry attracted members and resulted in greater commitment. SEMATECH dealt with naysayers strategically by giving them responsibility for key issues. For example, rather than asking technical experts to share knowledge to develop common technical standards, SEMATECH’s leaders created the first version of a standard and asked the technical experts to criticize and improve it.

SEMATECH also had to address members’ fear of sharing knowledge between competing firms. The consortium managers ensured that everyone understood that the meetings had a problem-solving orientation and that participants needed only general knowledge for problem solving. SEMATECH organized meetings to get a consensus on what was generic knowledge and what they could share openly. Conscious of the fact that more strategic knowledge might need to be shared from time to time, SEMATECH instituted breaks during meetings so members could call their headquarters to get permission to share the knowledge.

***Transforming***

Transforming consists of upgrading an organization to a structure that sustains and supports the implementation of responses to new opportunities and threats. Each industry model used different tactics to keep abreast of new risks related to the technology and improve the solutions developed. They all aimed for a process that reinforced practices each time a member shared useful information. Each industry model made different choices to renew best practices continuously.

INPO created a safety management program composed of an independent, international team of peers with extensive expertise and technical skills in nuclear power. Each time this team offers detailed evaluations of nuclear power plants and helps nuclear power plants improve their performance, the in-depth, objective knowledge derived from the evaluations helps identify new risks and new best practices. INPO uses each new risk or best practice to improve standard operations and future evaluations of nuclear operations at each participating company. This process remains active—the Fukushima disaster in 2011 activated extensive sharing and collaboration between INPO member companies.

The Responsible Care initiative developed codes of practice for each step in a chemical plant’s lifecycle: community awareness and emergency response, R&D, manufacturing, transportation, distribution, and hazardous waste management. The National Advisory Panel (NAP) comprises external environmental and labor experts who review the code of practices. Member companies’ CEOs meet quarterly to compare notes on their progress. Responsible Care went further and asked CEOs for their ideas to improve the code. Using the same mechanism of peer pressure, Responsible Care manages to renew its codes continuously.

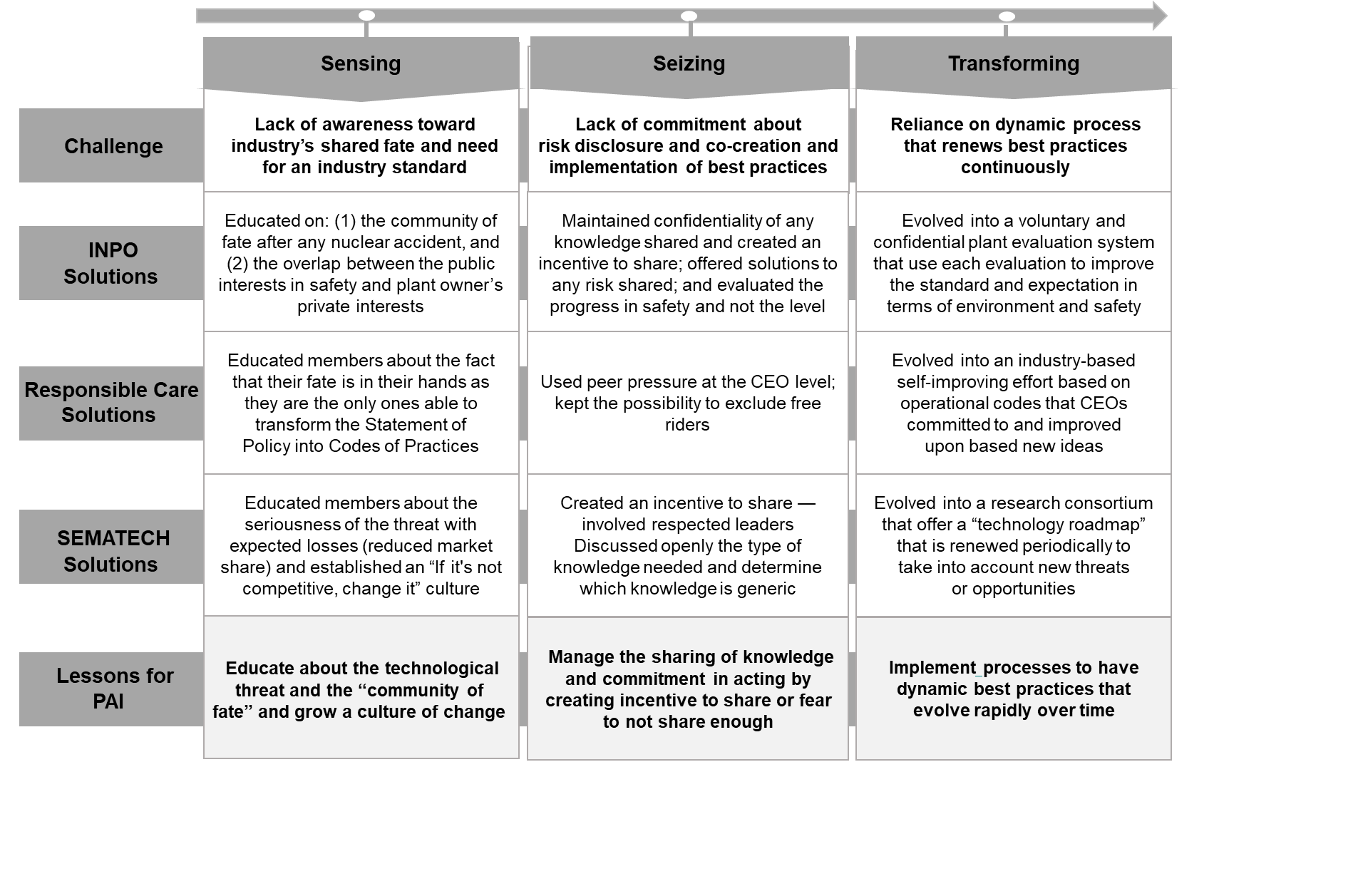
SEMATECH instituted a bottom-up planning process and organized a series of workshops to create a shared roadmap of technology development activities. The roadmap aimed to inform semiconductor companies and their materials and equipment suppliers about any likely new technologies required for manufacturing next-generation chips and any pieces of technology that were likely to be “showstoppers” for which intensive R&D was needed. The roadmap also coordinated the required timing for their introduction. The goal was to create a roadmap of value for all participants, and which any organization could use to set up its own development plans, prioritize investments, or discuss technology trends. Based on this first roadmap, 11 of the original 14-member companies agreed to extend their membership in SEMATECH at the end of the first five-year period. The roadmap was so successful that the Semiconductor Industry Association maintains a similar roadmap that it renews periodically. Renewing the roadmap is a way for SEMATECH to sustain and support the implementation of responses to new industry opportunities and threats.

Lessons for the AI Industry

The sensing, seizing and transforming portions of the Dynamic Capabilities framework applied to our three past models revealed some important implications for PAI.

Concerning sensing, PAI has created educational material designed to increase public understanding of the potential benefits, costs, and progress of AI. For AI companies specifically, PAI is developing education material on AI best practices; their materials will cover topics such as fairness, inclusivity, transparency, security and privacy, values and ethics, collaboration between people and AI systems, interoperability of systems, and safety, reliability, and robustness of AI technologies (Partnership on AI 2019).

Based on the lessons learned from our three examples, educating on AI risks or best practices is not sufficient to trigger a collective action to respond to serious downside risks from AI technology. Like the organizations we have discussed, PAI will need to expand beyond best practice education to inform participating companies about the overlap of private interests and environmental, safety, and societal technology standards, and cultivate the community of fate for all of the AI industry members (Figure 1). Just as the three industries that faced a catastrophic event needed to educate the industry members, education within the AI industry is imperative because a catastrophic event could occur. AI industry members are unlikely to perceive and accept a shared sense of community responsibility on their own.



Concerning seizing, PAI created several working groups like the Fairness, Transparency, and Accountability Working Group and the [AI and Media Integrity Steering Committee](https://www.partnershiponai.org/ai-and-media-integrity-steering-committee/) that are tasked with making timely decisions to react to any AI opportunity or threat. Each group engages experts from disciplines such as psychology, philosophy, economics, finance, sociology, public policy, and law to discuss, provide guidance, and support objective third-party studies on emerging issues related to the impact of AI on society. PAI’s working groups develop case studies on AI’s impact on labor and its use in criminal justice sentencing.

Based on the three examples presented, PAI may need more specific mechanisms to motivate members to disclose necessary information and to get member company CEOs to commit to implementing best practices developed by PAI regarding safe, socially responsible AI. Motivating industry members to participate in the working groups is a good first step, but PAI must ensure members commit to disclose the AI risks they identify and commit to implement the best practices.

With regard to transforming, PAI needs to evolve into a structure that sustains and supports the implementation of responses to new opportunities and threats. Currently, PAI has workgroups and workshops that it created based on prioritized risks and perceived opportunities. What PAI lacks, but which was present in our three examples, is a process for working groups to continue engaging after they produce results or best practices. Given the rapid pace of AI development and application, PAI’s results cannot be static—they will need to evolve. Past models offered several alternatives to develop dynamic best practices such as a standard-setting organisation that also has the role of evaluating companies’ practices and co-creating solutions, conducting industry-based self-improvement efforts that involve CEOs, and developing and periodically updating technology roadmaps.

**Conclusion**

Cooperation among competitors is sometimes necessary to sustain an industry. The AI industry has not yet experienced a catastrophe and has proactively formed the Partnership in AI. By tackling risks jointly and in advance, PAI and AI companies can co-develop solutions from the design stage, which might reduce the likelihood of a disaster, mitigate one if it occurs, and reduce potential costs for resolving it. With AI industry members’ commitment, PAI can ensure that the industry has three dynamic capabilities: sensing the opportunities inherent in a shared goal of achieving high environmental and societal standards for AI technologies, seizing on opportunities and threats in a timely manner by having companies engaged in AI disclose risks and commit to co-creating and implementing best practices, and transforming their overarching organization to enable a dynamic process that renews and updates industry risks and best practices continuously. The Dynamic Capabilities Framework could be useful for any industry requiring a community of fate in handling the environmental and social sustainability of its technology.