16. Facilitating stakeholder communication around AI-enabled systems and business processes

Matthew Bundas, Chasity Nadeau, Thanh H. Nguyen, Jeannine Shantz, Marcello Balduccini, Edward Griffor, and Tran Cao Son

INTRODUCTION

Business process means the collection of related, structured activities or tasks that serve a particular business goal (Wikimedia Foundation, 2022).

Artificial intelligence (AI) has produced spectacular results across a multitude of domains. Business solutions that seemed impossible are now made possible thanks to AI-enabled components, whose use is often imperative to the overall success of business processes. However, leveraging AI is not trivial. Given the complexity of AI components and their behavior, communication is a major hurdle among stakeholders with different backgrounds and goals. Each group of stakeholders may have its own set of concerns and requirements. Vocabulary can vary depending on each stakeholder's domain of expertise, and each group likely has its own goals, which can conflict with other groups' goals.

For example, within an organization, public relations experts may want to promote transparency surrounding decisions of AI-enabled systems and business processes. Transparency appears to be a good practice when working with your clients. However, cybersecurity experts may argue that excessive transparency would threaten security. Even within a small organization, aligning AI-enabled systems goals may be challenging. Now consider extending this to multiple business processes across various organizations with several divisions within each organization. How does one manage the multi-layered business processes involving AI-enabled systems?

To complicate matters, generally, with AI, data goes in, and decisions come out, yet the processes between input and output lead to decisions that are often difficult to explain. This is frequently the case with machine learning (ML), a problem compounded by its widespread use. Research shows that biases may be fed into and often hidden in ML models, which can lead to unintentional and undesirable results (van Es et al., 2021). Systems in which the decision-making process is not transparent are referred to as "black boxes," and their nature is often problematic for businesses. The complicated and opaque decision-making processes of AI

components increase the communication challenges already faced by a diverse group of stakeholders.

In this chapter, we demonstrate that the cyber-physical systems (CPS) framework, developed by the US National Institute of Standards and Technology (NIST), provides a useful tool for solving these challenges. The NIST CPS framework was created to bring together CPS's stakeholders by providing a common vocabulary and process structure. While the NIST CPS framework was explicitly conceived for CPS, we believe that the underlying approach can be useful to overcome the challenges that emerge from the use of AI components in AI-enabled systems and business processes – especially "black-box" AI components. From the perspective of policy and practice, we believe that this approach can be effective in guiding stakeholders through the development of best practices surrounding systems and processes that involve diverse knowledge domains and goals. We begin the chapter by providing an introduction to the NIST CPS framework and then shift our attention to practical aspects by discussing a use case inspired by recent events.

BACKGROUND

NIST is a United States government organization specializing in developing standardized templates and processes for various uses and applications (NIST, 2022). NIST recognizes the many challenges in designing, constructing, operating, and assuring a CPS, and in response, developed the tool known as the CPS Framework (NIST, 2017).

The CPS Framework provides the basis for designing, building, and assuring a CPS. This structure determines whether the system under development meets the expectations and addresses the concerns identified by stakeholders. The CPS framework creates a "common foundation" on which systems can be "developed, safely and securely combined, verified, and delivered" to a diverse group of stakeholders (NIST, 2017). By design, the scope of the CPS framework is vast so that it may be adopted by a broad range of CPS application domains.

The CPS framework is fundamentally made up of concerns and facets. Concerns are identified through the lenses of multiple stakeholders. They are addressed throughout the CPS facets or the processes and activities that comprise the conceptualization, realization, and assurance of the CPS. Concerns are considered during the activities of all three facets and to every function, from the individual components to the sets of features that deliver function in a realized CPS. Thus, concerns form the basis of the CPS framework. The CPS framework organizes related concerns into higher level concerns and, ultimately, into one of the ten highest level concerns, called aspects, including the functional, human, business, and trustworthiness aspects and six others. The CPS framework's list of concerns resulted from a consensus between more than 500 stakeholders, including government, industry,

and academic stakeholders. It is, nonetheless, not a completed structure and may be modified or extended based on the needs of the application and the changing operating environment (NIST, 2017).

To better understand the relationship between aspects and concerns, let us look at the aspect of trustworthiness. The trustworthiness aspect is broken down into several individual concerns, including privacy, reliability, resilience, safety, and security (NIST, 2017). For a system to be considered trustworthy, i.e., to satisfy the trustworthiness aspect, each of those concerns must in turn be satisfied. Concerns may be associated with constraints on the system's design or behavior, which are called requirements in the CPS framework. We say that the requirements address the concern they are associated with. For a concern to be satisfied, requirements that address it must be satisfied. The set of these requirements, which are added for the sake of concerns deemed relevant to the system, comprises a CPS model, which is the outcome of the conceptualization facet. They are measurable constraints on measurable parameters or characteristics of the system. We further expand on the relationship between aspects and concerns in the use case section below.

The three facets and their interdependencies during the lifecycle of a system are identified in the CPS framework (NIST, 2017). Once again, these are conceptualization, realization, and assurance. Each facet has its unique set of characteristics, activities, and artifacts documenting whether and how expectations are met. We can think of the artifacts as the end product of the individual facet (Balduccini et al., 2018).

The conceptualization facet focuses on the design and outline of the CPS by creating a blueprint for how the device will function and be constructed, primarily in a logical sense. The conceptualization facet produces the artifact of a blueprint, or model, for the CPS (NIST, 2017).

The realization facet focuses on how the CPS is built and tested against the design requirements developed from the CPS model. The realization facet produces the artifact of the CPS itself (Balduccini et al., 2018).

Lastly, the assurance facet highlights the use of the artifacts of conceptualization and realization as evidence that requirements are met by the CPS, e.g., to assure that the CPS is safe, secure, trustworthy, etc. The artifact of the assurance facet is an assurance case consisting of evidence and test results as well as requirements that have been implemented throughout the design and construction of a CPS. These assurance cases are evidence that a system can function safely to achieve its desired goal (Balduccini et al., 2018).

ILLUSTRATIVE USE CASE

While AI and ML provide a pathway to new and exciting possibilities, these technological solutions are not without challenges that may hinder adoption (Radanliev et al., 2020). We do not always understand the reason for the decisions made by AI

components. Predicting what they will do under new circumstances is also sometimes difficult.

Consider the controversy recently sparked by the release of the Apple Card issued by Goldman Sachs Bank USA. David Heinemeier Hansson, a tech entrepreneur and author of Ruby on Rails, tweeted about alleged gender discrimination in the algorithms used to determine credit limits for the Apple Card. Despite filing joint tax returns and not disclosing income specifics when applying for the card, Hansson received a credit limit 20 times that of his wife. Ironically, his wife has a better credit score. Apple responded by raising Hansson's wife's credit limit. However, the resolution was a one-off response as Hansson was informed that Apple could not change the algorithm's decision (Reuters, 2019; Vincent, 2019).

Hansson was not the only tech leader to report discriminatory issues with the Apple Card. Apple co-founder, Steve Wozniak, was given 10 times the credit limit offered to his wife. Wozniak called on the government to investigate the operation of "black box" algorithms, which experts say are often biased (Bloomberg, 2019).

According to New York state law, any algorithm leading to discriminatory treatment of protected classes of people, including women, violates the law. In November 2019, the New York Department of Financial Services (NYDFS) announced it would formally investigate the Apple Card and claims of gender discrimination. On March 23, 2021, NYDFS (Department of Financial Services, 2021) issued a report of their investigation's findings, stating that after conducting interviews with witnesses, analysis of thousands of pages of records, and examining data concerning more than 400,000 New York State applicants, they did not find evidence of unlawful discrimination under fair lending law. However, the report acknowledges unequal access to credit based on gender in the industry as a whole, suggesting it is a systematic problem in need of a remedy (Department of Financial Services, 2021). Although Apple and Goldman Sachs were cleared in the court of law for any wrongdoing concerning gender discrimination, Apple and Goldman Sachs could still have underlying issues with gender discrimination, and their public image undoubtedly was impacted.

COMMUNICATING AROUND AI-ENABLED BUSINESS PROCESSES VIA NIST CPS FRAMEWORK

As we saw in the investigation into the Apple Card case, the law says that the unintentional nature of a bias is not an excuse for non-compliance. What happens when ML algorithms go awry? Two main challenges associated with ML models are:

1. As mentioned earlier, AI components, especially those based on ML, often act as "black boxes." Results can be difficult or impossible to explain (Gallagher, 2020; Laplante et al., 2020; Hall, 2020).
2. Even interpretable AI-enabled systems may be too complicated to explain, especially for non-experts (Hall, 2020).

One of the main goals of explainable AI (XAI) is to explain AI-driven systems humans can understand (Muncke, 2021).[1] According to Matt Turek, "new machine-learning systems will have the ability to explain their rationale, characterize their strengths and weaknesses, and convey an understanding of how they will behave in the future" (Turek, n.d.). As such, XAI is positioned to become a critical component in addressing the above challenges.

The NIST CPS framework can be extremely useful in meeting the goals outlined by Matt Turek by providing stakeholders with better control and more informed insights into the behavior of AI-enabled systems as well as by providing stakeholders with ways to discuss the requirements of AI-enabled business processes. The CPS framework can help bridge gaps by simplifying major aspects and concerns of systems into easily understandable components. Moving toward more explainable and trusted models is necessary, especially in highly regulated fields such as health care, insurance, and finance.

CHARACTERIZING THE BEHAVIOR OF AI-ENABLED SYSTEMS AND BUSINESS PROCESSES

While the CPS framework was designed to provide a refined, comprehensive set of concerns that can guide the engineering process of arbitrary systems, the CPS framework can also be extended easily should a company face challenges that require dedicated concerns. To illustrate this, in the remainder of this chapter, we demonstrate multiple ways in which the NIST CPS framework can be leveraged to characterize the behavior of AI-enabled systems and business processes. We begin by eliciting a number of important considerations related to such characterization from a business perspective.

Let us begin by assembling a possible set of considerations from the perspective of business users. We take inspiration from observations found in the literature and contributed by businesses and users leveraging AI for business processes. In CognitiveScale (2019), George Lawton highlights the importance of explainability in AI and identifies four ways of making AI more explainable:

1. *Understand the data*. In addition to having a deep understanding of what the data offers, be sure that training data mirrors the expected data for which the model is developed.
2. *Balance explainability, accuracy, and risk*. Be sure the decisions based on the AI output reflect the company's mission and goals.
3. *Focus on the user.* Explanations must be appropriate for each stakeholder population. Technical explanations should be reserved for only those groups who understand the language. Understanding is important for promoting end-user trust and adoption.

[1] Explainable artificial intelligence (XAI) is a remarkable attempt at making AI components more transparent and will likely improve the chances of success of AI-enabled systems and processes while keeping expectations reasonable (Casey, 2020).

4. *Use key performance indicators (KPIs) for AI risk*. Components of AI risk may include: bias, compliance, comprehensiveness, data privacy, explainability, and fairness. Relevant metrics can be generated for each group of stakeholders.

The first item suggests that one should provide ways of identifying the data features used by the system or process. The second item suggests that methods should be provided to identify and discuss notions of explainability, accuracy, and risk. Remarkably, the third item reiterates a foundational notion already present in the NIST CPS framework: the vocabulary chosen should be hierarchically organized in such a way that concepts at higher levels of the hierarchy are understandable by all stakeholders, regardless of their specific backgrounds and interests. Concepts at lower levels of the hierarchy should, instead, focus on particular expert classes. The fourth item suggests a potential source for a vocabulary describing possible concerns in this area, especially risk-related ones: KPIs. KPIs have already been successfully used in several domains. One example of KPI adoption related to AI is AI Global's AI Trust Index. This index is defined as a FICO-like risk score for AI. The tool allows companies to define their best practices and compares AI practices against industry benchmarks (CognitiveScale, 2019). Because many companies use KPIs, this is a widely accepted strategy. Joydeep Ghosh, chief scientific officer at AI vendor CognitiveScale, claims that companies should first "establish a set of criteria for KPIs for AI risks, including comprehensiveness, data privacy, bias, fairness, explainability and compliance" (CognitiveScale, 2019).

This information indicates useful terms related to potential considerations by business users regarding AI-enabled systems and especially AI-enabled business processes. A possible hierarchical organization of the relevant terms is:

- Rationality
  - Compliance
    - Bias
    - Ethics
    - Fairness
  - Comprehensiveness
  - Data privacy
  - Explainability

As the reader may notice, rationality is chosen as the root of the AI-related hierarchy. This is aligned with the view shared by parts of the AI community that one of the most salient features of successful AI is rational behavior. We also find it to be a better choice for the root concept than AI itself because AI is viewed sometimes as a collection of technologies. In the following section, we present a business-related use case that leverages the above concepts. In a later section, we demonstrate how the CPS framework may be used in that domain and how the above concepts can be associated with the elements of the framework.

USE CASE: USING REQUIREMENTS TO SHAPE BEHAVIOR

Let us consider a use case in which a Company A wants to develop the requirements for business processes related to processing credit card applications – for which the business processes use AI. The goal is to minimize bias and ensure the fairness of the credit card application assessment process.

When a consumer applies for a new card through Company A, the application collected by Company A is evaluated in two ways. First, internally, Company A processes the application to determine whether the application should be accepted or denied and, if accepted, with what credit limit and interest rate.

Company A's stakeholders may have many diverse concerns related, for example, to fairness, (financial) risk, explainability, and (cyber)security. To address these concerns, Company A establishes a set of requirements that their processes must follow internally. Company A is committed to eliminating bias in decision-making processes related to the new credit card. Company A's leadership determines that decisions regarding an application's approval or denial, credit limits, and interest rates shall be fair for all applicants. Company A, as a business in the financial industry, also has an interest in minimizing the risk potential applicants may pose to Company A when trusted with a credit card.

Based on substantial risk analysis and prior history, Company A believes that financial risk to Company A is higher among users and applicants who hold many credit cards with other financial institutions. Company A has identified that applicants with five or more active credit cards pose exceptionally high risk. Based on this finding, to help mitigate risk, Company A does not have an interest in accepting applicants with five or more active credit cards. Company A introduces a requirement related only to the applicant's number of active credit cards that if the number of active credit cards is five or greater, the application is declined. Otherwise, the application may still be considered. This requirement, formally denoted as "decline_five_or_more_credit_cards," is used as part of Company A's review process for all applications, helping to address the considerations related to risk.

As part of the same analysis, Company A determines that the risk of default is higher for younger applicants. In particular, those under the age of 28 are more likely to be late or default on payments. Company A defines a requirement based solely on age that is consistent for all applicants: each applicant begins with the same credit limit; those above 28 years of age receive higher credit limits. Company A determines that all approved credit cardholders 28 years of age and older should receive a credit limit 20% higher than those approved credit cardholders younger than 28 years of age. This requirement, labeled "adjusted_credit_limit_above_28_years_of_age," helps to address the considerations related to risk by ensuring applicants who may pose a higher risk to Company A have less of an opportunity to impact Company A negatively.

Company A establishes a similar requirement for assigning interest rates. Because cardholders 28 years of age and older pose less risk, they receive a lower interest rate. Company A determines that all approved credit card holders 28 years of age

and older will receive an interest rate lower than those approved and under 28 years of age. Formally stated, the requirement might be "Company A shall provide a lower interest rate for those who are at least 28 years of age." This requirement, again, helps address the concern about risk, as it helps mitigate risk toward Company A and is formally labeled as "adjusted_interest_rate_above_28_years_of_age."

Company A contracts with several credit score services to obtain applicants' comprehensive financial information. To ensure a sound process and avoid potential future scrutiny, Company A established predictable third-party requirements similar to the internal properties listed above. Each third-party credit score service shall provide evidence demonstrating that the concerns about fairness and cybersecurity are addressed.

For this example, Company A wants to assure predictability and accountability for gender and ethnic-related requirements. Research shows that training data used in producing AI algorithms may be biased if gender representation is unbalanced (Dastin, 2018). Company A requires that "third-party partners shall use gender-balanced data to train their AI algorithms and produce evidence that the dataset has a difference in gender proportion no greater than 5%," which helps address the considerations about bias. We label this requirement "balanced_gender_data."

Similarly, research shows that unintentional bias may be present when gendered job titles are included in the training data, i.e., "stewardess" or "policeman," and women are "devalued when using feminine job titles" (van Es et al., 2021). Company A is concerned that applicants with gendered job titles may receive a lower or higher credit score as part of their application assessment process. To counteract this issue, Company A introduces the requirement "gender_specific_stop_words," expressing that the third-party partners shall produce evidence assuring gender-specific words are not used in determining a credit score.

This same worry also applies to ethnic-specific terms collected on applications or used in training data. With this requirement, named "ethnic_specific_stop_words," third-party partners shall produce evidence assuring ethnic-specific words are not used in determining a credit score. These stop-word requirements help to address the considerations related to bias.

In creating these requirements, Company A has identified the criteria and rules which help guide the credit card application assessment process. These requirements are both helpful for the business in constraining their processes internally and helpful for the applicants. The requirements created by Company A are clearly defined and applied to all applications in the same manner, ensuring that all applications are assessed in the same manner. For these reasons, these requirements help to address the considerations regarding risk and bias and help to address the considerations related to fairness.

In working with third-party partners, and interacting with applicants, Company A would like to ensure that each applicant's sensitive information, such as complete bank account numbers, social security numbers, etc., remains secure during communication between the relevant parties. Inspired by the Federal Trade Commission's (FTC) requirement for merchants to truncate information on receipts (Federal Trade

Commission, 2007), Company A requires that all sensitive information contained in communications is truncated. For example, in communications, a bank account number 0123456789 may be reduced to 56789. This requirement applies internally to Company A and third-party partners, helping to address the considerations of privacy and cybersecurity. It is denoted as "truncate_sensitive_data."

Company A also desires to remain transparent to applicants regarding the factors impacting their application, especially avoiding their decision-making process becoming a "black box." To help ensure transparency, Company A requires that, upon completion of an application assessment, the applicant is sent a letter outlining the factors that went into deciding their application. In particular, Company A requires that each property used as part of the application assessment is presented and that an explanation is provided on how it impacted the specific application. Factors such as the applicant's credit score and other financial metrics used in the application assessment may also be included. This requirement, labeled "application_assessment_explanation," helps address the considerations about fairness and transparency.

Meeting Company A's requirements, third-party partners shall provide predictable and accountable credit scores. Company A obtains a credit score for our application from third-party partners. Ultimately, the third-party results are combined with the internal results discussed above to determine final decisions regarding the terms of the credit card, and an explanation regarding the decisions is provided. The final decision is guaranteed to be fair because each level of the decision process has requirements and evidence that the output is fair. By creating concrete requirements that constrain the implementation of the AI-enabled application assessment, Company A can ensure their practices are in line with their considerations.

CAPTURING AI-RELATED CONCEPTS IN THE CPS FRAMEWORK

In this section, we analyze multiple approaches in which Company A may leverage the CPS framework for capturing the above AI-related considerations. Figures 16.1–16.3 provide a graphical illustration of the three methods applied to the Company A use case just described.

In the figures, aspects are shown at the top of the tree, and their sub-concerns are shown below them. Concerns are shown as ovals with their names inside, and aspects are denoted by rectangles with rounded corners. Requirements are shown as rectangles with their names inside. The sub-concern relation between the two concerns is shown with a solid line between the two concerns. The fact that a given requirement addresses a concern is indicated by a dashed edge labeled "addresses." Note that, for compactness, Figures 16.1–16.3 depict only the relevant portions of the concern tree[2].

[2] The full concern tree contains over 100 concerns and ten aspects.

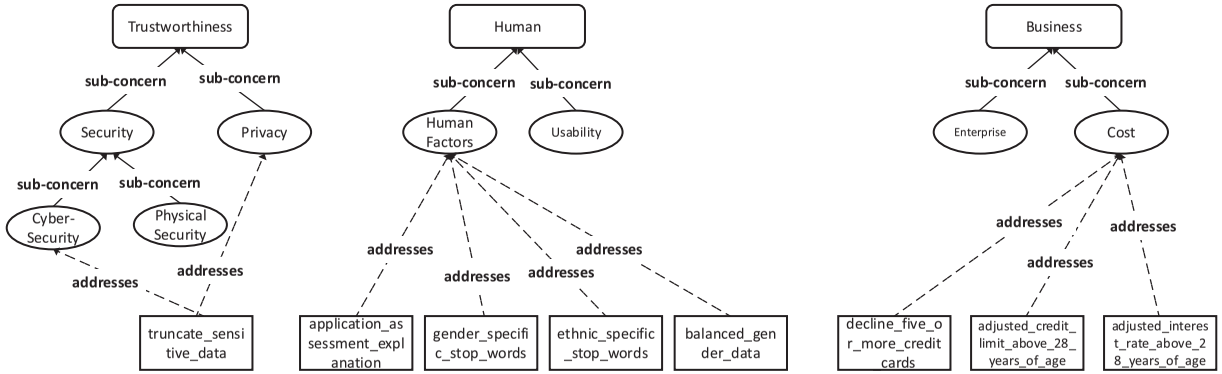

Figure 16.1 Representation of Company A use case with Approach 1, using existing concerns in CPS framework

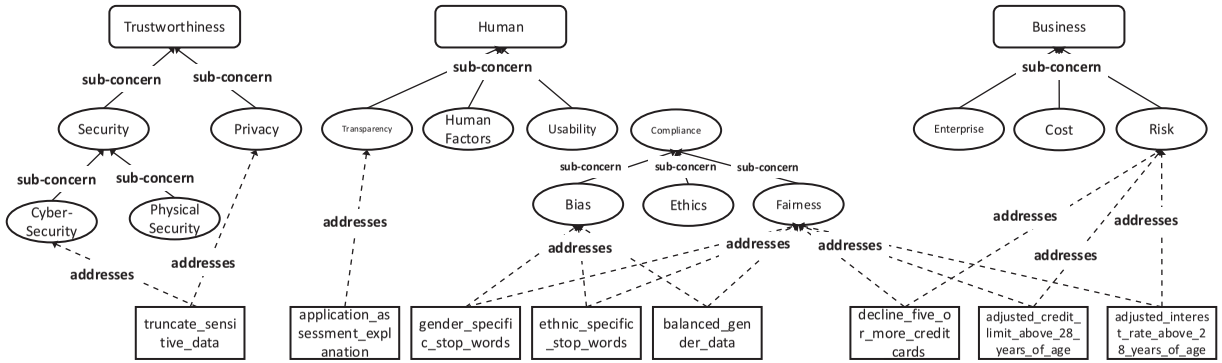

Figure 16.2 Representation of Company A use case using Approach 2, introducing new AI-related concerns

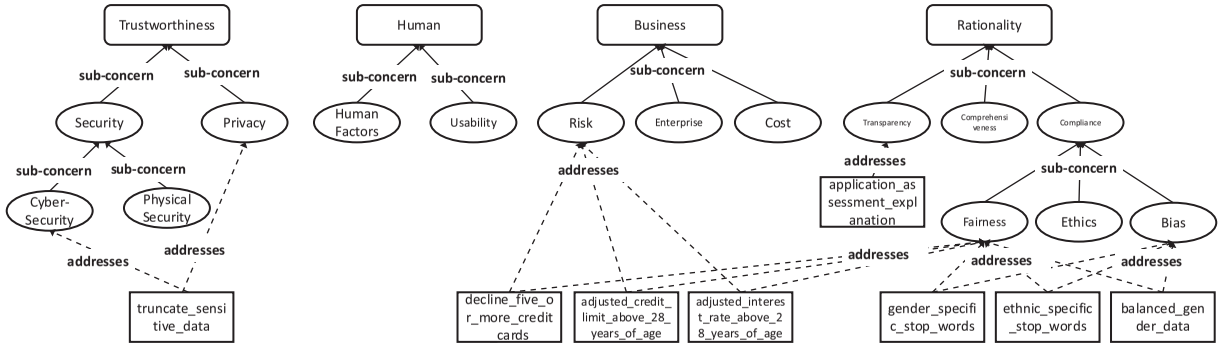

Figure 16.3 Representation of Company A use case with Approach 3, introduction of rationality hierarchy/aspect

The first approach, shown in Figure 16.1, assumes that Company A is satisfied with the set of concerns already present in the NIST CPS framework. The requirements discussed earlier can, thus, be formalized as NIST CPS framework requirements and linked to the concerns that provide the best fit. Notice that associating AI-related concepts to existing concerns may not always be straightforward for Company A. For example, of the available concerns, the best fit for the requirement "decline_five_or_more_active_credit_cards" might be the cost concern. However, one might argue that there is a subtle, yet important, difference between the notion of cost and risk, with which the requirement would be more naturally associated. Despite this slight potential lack of alignment – which might complicate the communication among stakeholders – this approach has the benefit of using the already well-understood existing concern tree without modification. Additionally, each requirement is associated with a single concern, thus simplifying the understanding and use of the diagram.

Figure 16.2 illustrates a second approach for the introduction of AI-related concepts in the concern tree. In this approach, dedicated concerns are added to the trees that represent these concepts and integrated into the existing branches of the concern tree. The new concerns, which in this notional example include transparency, fairness and bias, are shown by ovals with a green border. Introducing dedicated concerns may make it easier to map the requirements of the system to more suitable concerns. Requirements relevant to risk can be linked to the risk concern, those addressing biases are associated with the bias concern, and so on. Compared with the first approach, this may allow for an arguably more meaningful and accurate representation of the behavior of an AI-enabled system or business process and, consequently, more effective communication among stakeholders. On the other hand, this structure lacks an explicit characterization of the role of the AI component in the system or business process.

Figure 16.3 shows the third and final approach for introducing AI-related concepts in the concern tree. In this approach, we extend the concern tree by an entirely new branch on which the AI-related primitives are captured by possibly more fitting concerns rooted in a newly introduced rationality aspect. The idea here is that the new branch makes it possible to gather all AI-relevant concerns in a standard structure. In a similar way to the previous approach, requirements related to these primitives can easily be associated with concerns found in this branch. For example, the requirement "gender_ratio" addresses the bias concern, which is a part of the rationality hierarchy. This representation not only makes it explicit that the requirement addresses bias in the business process but also clearly indicates that the requirement ultimately affects how the AI component of the process makes decisions. This is not as evident in the representation obtained by the previous approaches: in the first approach, the distinct concept of bias is not embedded in a clearly articulated concern; in the second approach, a dedicated concern is present but is not explicitly grouped with the other AI-related concerns. In summary, introducing a separate rationality branch may make for a clearer and more comprehensive representation of the AI-relevant concepts. On the other hand, the introduction of the new branch causes an inevitable increase in the complexity of the concern tree.

CONCLUSION

AI is often critical to the success of business processes. Leveraging it, however, is not trivial. In this chapter, we argued that a major hurdle in the development, use, and maintenance of AI-enabled systems and business processes is communication, specifically the discussion of requirements among stakeholders with different backgrounds and goals.

We introduced the NIST CPS framework, discussed how it can be used in the context of AI-enabled systems and business processes, and advocated that the CPS framework can help overcome the challenges related to them. We introduced an illustrative use case that showcases the challenges stemming from the adoption of AI-enabled systems and business processes. Finally, we demonstrated how the CPS framework may be applied to that use case in order to capture the possible considerations made by stakeholders.

To demonstrate the flexibility of the NIST CPS framework, we showed three approaches for leveraging the NIST CPS framework to capture the complexities of AI-enabled systems and business processes. The approaches were presented in increasing order of the magnitude of the changes to the original structure and of the approaches' ability to clearly link stakeholders' requirements to the relevant elements of AI components.

The use case presented is also intended to highlight how, from the perspective of policy and practice, the NIST CPS framework can provide an effective way of guiding stakeholders through the development of best practices surrounding AI-enabled systems and business processes, a task that is often difficult because it involves diverse knowledge domains and goals.

ACKNOWLEDMENTS

The authors are deeply grateful to Kathleen Campbell Garwood and Virginia Miori for valuable discussions on topics related to this chapter. In addition, we thank Andrew Holmberg, Brandon Andrews, and Bob Shantz for their suggestions and comments on drafts of this chapter. Portions of this publication and research effort are made possible through the help and support of NIST via cooperative agreements 70NANB21H167 and 70NANB22H145. Son Tran acknowledges the partial support of the NSF grants 1914635, 1757207, and 1812628.

DISCLAIMER

best available for the purpose. Portions of this publication and research effort are made possible through the help and support of NIST via cooperative agreements 70NANB21H167 and 70NANB22H145.

REFERENCES

Balduccini, M., Griffor, E., Huth, M., Vishik, C., Burns, M., & Wollman, D. (2018). Ontology-based reasoning about the trustworthiness of cyber-physical systems. *IET Journal of the IoT* – 2018, 2018, pp. 1–10. https://doi.org/10.1049/cp.2018.0012.

Bloomberg. (2019, November 10). *Apple co-founder says Goldman's Apple Card algorithm discriminates*. Bloomberg.com. Retrieved March 8, 2022, from https://www.bloomberg.com/news/articles/2019-11-10/apple-co-founder-says-goldman-s-apple-card-algo-discriminates?utm_medium=social&utm_source=twitter&utm_content=business&cmpid=socialflow-twitter-business&utm_campaign=socialflow-organic

Casey, K. (2020, November 19). *How to explain machine learning in plain English*. The Enterprisers Project. Retrieved March 8, 2022, from https://enterprisersproject.com/article/2019/7/machine-learning-explained-plain-english?page=1

CognitiveScale. (2019, September 19). *Using an AI trust index to unblock stalled machine learning & AI projects • Cognitivescale*. CognitiveScale. Retrieved April 15, 2022, from https://blog.cognitivescale.com/using-an-ai-trust-index-to-unblock-stalled-machine-learning-ai-projects

Dastin, J. (2018, October 10). *Amazon scraps secret AI recruiting tool that showed bias against women*. Reuters. https://www.reuters.com/article/us-amazon-com-jobs-automation-insight-idUSKCN1MK08G

Department of Financial Services. (2021, March 23). Press Release - March 23, 2021. *DFS issues findings on the Apple Card and its underwriter Goldman Sachs Bank*. Department of Financial Services. Retrieved March 8, 2022, from https://www.dfs.ny.gov/reports_and_publications/press_releases/pr202103231

Federal Trade Commission. (2007, May 2). *Slip showing? Federal law requires all businesses to truncate credit card information on receipts*. Federal Trade Commission. Retrieved April 15, 2022, from https://www.ftc.gov/tips-advice/business-center/guidance/slip-showing-federal-law-requires-all-businesses-truncate

Gallagher, W. (2020, August 20). *One year later, the Apple Card is a huge but controversial success*. AppleInsider. Retrieved March 8, 2022, from https://appleinsider.com/articles/20/08/20/one-year-later-the-apple-card-is-a-huge-but-controversial-success

Hall, P. (2020, March 19). *Explaining machine learning models to the business*. InfoWorld. Retrieved March 8, 2022, from https://www.infoworld.com/article/3533369/explaining-machine-learning-models-to-the-business.html

Laplante, P., Milojicic, D., Serebryakov, S., & Bennett, D. (2020, November). Artificial intelligence and critical systems: From hype to reality. *Computer*, *53*(11), 45–52. https://doi.org/10.1109/MC.2020.3006177

Muncke, J. (2021, October 24). *The business case for AI safety: Explainability*. Faculty. Retrieved March 8, 2022, from https://faculty.ai/blog/the-business-case-for-explainability/

NIST. (2022, January 11). *About NIST*. NIST. Retrieved April 15, 2022, from https://www.nist.gov/about-nist

NIST. (2017, June). *Framework for cyber-physical systems: Volume 1 - nist* (n.d.). Retrieved March 8, 2022, from https://nvlpubs.nist.gov/nistpubs/SpecialPublications/NIST.SP.1500-201.pdf

Radanliev, P., De Roure, D., Van Kleek, M. et al. (2020). Artificial intelligence in cyber physical systems. AI & Soc. https://doi.org/10.1007/s00146-020-01049-0

Reuters. (2019, November 11). *Goldman faces probe after entrepreneur claims gender bias in Apple Card algorithm*. VentureBeat. Retrieved March 8, 2022, from https://venturebeat.com/2019/11/11/goldman-faces-probe-after-entrepreneur-claims-gender-bias-in-apple-card-algorithm/

Turek, M. (n.d.). *Explainable Artificial Intelligence (XAI)*. DARPA RSS. Retrieved March 8, 2022, from https://www.darpa.mil/program/explainable-artificial-intelligence

van Es, K., Everts, D., & Muis, I. (2021). Gendered language and employment Web sites: How search algorithms can cause allocative harm. *First Monday*, *26*(8). https://doi.org/10.5210/fm.v26i8.11717

Vincent, J. (2019, November 11). *Apple's credit card is being investigated for discriminating against women*. The Verge. Retrieved March 8, 2022, from https://www.theverge.com/2019/11/11/20958953/apple-credit-card-gender-discrimination-algorithms-black-box-investigation

Wikimedia Foundation. (2022, January 25). *Business process*. Wikipedia. Retrieved March 8, 2022, from https://en.wikipedia.org/wiki/Business_process