

A Human-Centered Methodology for Creating AI FactSheets

John Richards, David Piorkowski, Michael Hind,
Stephanie Houde, Aleksandra Mojsilović, Kush R. Varshney
IBM Research AI
ajtr@us.ibm.com, djp@ibm.com, hindm@us.ibm.com,
stephanie.houde@ibm.com, aleksand@us.ibm.com, krvarshn@us.ibm.com

Abstract

As artificial intelligence (AI) models and services are used in a growing number of high-stakes areas, a consensus is forming around the need for a clearer record of how these models and services are developed to increase trust. Several proposals for higher quality and more consistent AI documentation have emerged to address ethical and legal concerns and general social impacts of such systems. However, there is little published work on how to create this documentation. In this paper we describe a methodology for creating the form of AI documentation we call FactSheets. This paper describes the methodology and shares the insights we have gathered while creating nearly two dozen FactSheets. Within each step of the methodology, we describe the issues to consider and the questions to explore with the relevant people in an organization who will be creating and consuming AI facts. This methodology may help foster the creation of transparent AI documentation.

1 Introduction

Recent work has outlined the need for increased transparency in artificial intelligence (AI) for data sets [7, 2, 9, 3], models [15], and services [1]. Proposals in support of ethical and trusted AI are also emerging [21, 19, 11]. Although the specifics differ, all are motivated by the desire to define a set of attributes that capture essential details of how an AI model or service was developed and tested to better understand technical, ethical, and regulatory concerns.

Despite this work on transparent reporting mechanisms, there is little consideration of how to create this documentation. Determining *what information* to include and *how to collect* that information is not straightforward. To our knowledge this is the first work outlining a methodology for creating this documentation. We believe this methodology can promote the creation of useful AI documentation.

We have proposed a mechanism for AI documentation called FactSheets [1]. FactSheets take a more general approach to AI transparency than previous proposals [7, 2, 9, 15, 21, 4] in several ways:

- FactSheets are tailored to the particular AI model or service being documented, and thus can vary in content
- FactSheets are tailored to the needs of their target audience or consumer, and thus can vary in content and format, even for the same model or service

Copyright 2021 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Bulletin of the IEEE Computer Society Technical Committee on Data Engineering

- FactSheets capture model or service facts from the entire AI lifecycle
- FactSheets are compiled from information generated by multiple contributors as they perform their actions throughout this lifecycle, thereby increasing the accuracy of these facts

FactSheets can document AI services in addition to individual models. We think this is important for three reasons:

- AI services are the building blocks for many AI applications. Developers call the service API and consume its output. An AI service can be an amalgam of many models trained on many datasets. The models and associated datasets are (direct and indirect) components of an AI service, but they are not themselves the interface to the developer.
- An expertise gap often exists between the producer and consumer of an AI service. The production team leverages the creation of one or more AI models and thus will mostly contain data scientists. The consumers of the API services tend to be developers. When such an expertise gap exists, it becomes more crucial to communicate the attributes of the service in a consumable way.
- Services composed of trusted models may not necessarily be trustworthy, so it is prudent to also consider transparency and accountability of services in addition to datasets and models. In doing so, we take a functional perspective on the overall service and can test for performance, safety, and security aspects that may go beyond what is relevant for a model in isolation.

Our methodology is motivated by user-centered design principles [14], where input from multiple stakeholders is collected to inform design. Although this takes more time than a single person designing the documentation, it is significantly more likely to meet the needs of FactSheet consumers [13]. This paper focuses on a specific form of AI documentation, FactSheets, but the techniques can be applied to other forms of AI (or even non-AI) documentation. Note, also, that our discussion centers on business applications of AI but the techniques can be applied to creating documentation for AI outside of this setting.

Before we describe our methodology in detail, we first highlight a few key concepts. Section 2 describes the AI lifecycle, summarizing the relevant roles and workflow for the construction and deployment of an AI model or service. Section 3 describes the concept of a FactSheet and motivates the need for a FactSheet Template. Section 4 presents our seven-step methodology for constructing useful FactSheets. Section 5 presents further guidance for those organizations planning to create FactSheets. Section 6 discusses how the methodology can help to address the needs of consumers with regards to the potential safety and harm of AI. Finally, Section 7 touches on what we are finding as FactSheets are put into production use.

2 The AI Lifecycle

The AI lifecycle includes a variety of roles, performed by people with different specialized skills and knowledge that collectively produce an AI model or service. Each role contributes in a unique way, using different tools. Figure 1 specifies some common roles.

The canonical process starts with a business owner who requests the construction of an AI model or service. The request includes the purpose of the model or service, how to measure its effectiveness, and any other constraints, such as bias thresholds, appropriate datasets, or the required levels of explainability and robustness.

The data scientist uses this information to construct a candidate model by using, most typically, a machine learning process. This iterative process includes selecting and transforming the dataset, discovering the best machine learning algorithm, tuning algorithm parameters, etc. The goal is to produce a model that best satisfies the requirements set by the business owner.

Before this model is deployed it often must be tested by an independent person, referred to as a model validator

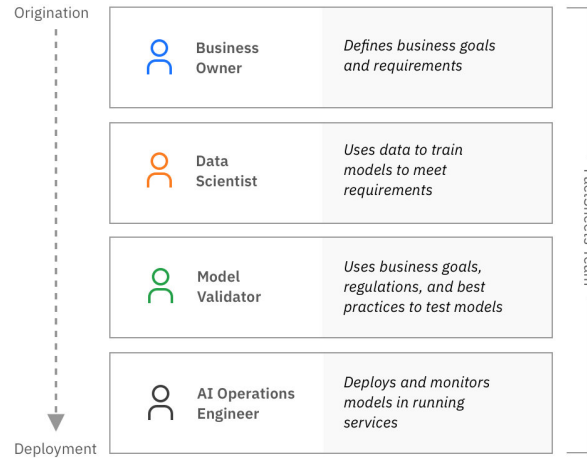


Figure 1: Key roles in a typical AI lifecycle

in Figure 1. This role, often falling within the scope of model risk management [16], third party testing [23, 5], or certification [22, 5], is similar to a testing role in traditional software development. A person in this role may apply a different test dataset to the model and independently measure metrics defined by the business owner. The person may also develop a "challenge" model to see if a simpler, and thus, less risky, solution could solve the same problem. If the validator approves the model, it can be deployed.

The AI operations engineer is responsible for deploying and monitoring the model in production to ensure it operates as expected. This can include monitoring its performance metrics, as defined by the business owner. If some metrics are not meeting expectations, the operations engineer is responsible for taking actions and informing the appropriate roles.

AI lifecycles will include iteration within a role (a data scientist, building many models before passing it to a validator) or between roles (an operations engineer sending a model back to a data scientist because it is performing poorly). More sophisticated lifecycles will likely have additional roles. A common pattern is for a model to be combined with other models or human-written code to form a service. In such a case the validator's role may be extended to also validate the full service.

A model is not a static object in the lifecycle, and thus, a FactSheet must incorporate the facts and lineage from all phases of the "life of the model". This will introduce transparency not only into how the model was built and what it does, but also how it was tested, deployed, and used.

3 FactSheets and Templates

FactSheets [1] are a collection of information about how an AI model or service was developed and deployed. FactSheets summarize the key characteristics of a model or service for use by a variety of stakeholders. We have previously summarized the difficulties developers face when creating FactSheets [8]. This paper describes the best practices we have developed in the process of creating FactSheets for nearly two dozen models. These include FactSheets for standalone models as well as services that encapsulate one or more models. They cover a wide range of application areas including text analysis and generation, language translation, object detection, object classification, audio signal classification, weather forecasting, agricultural crop yield prediction, and facility energy optimization.

This work has demonstrated that although FactSheets will contain some common elements, different FactSheets will generally contain different information, at different levels of specificity, depending on domain and model type. They will also contain different information for different industries and the different regulatory

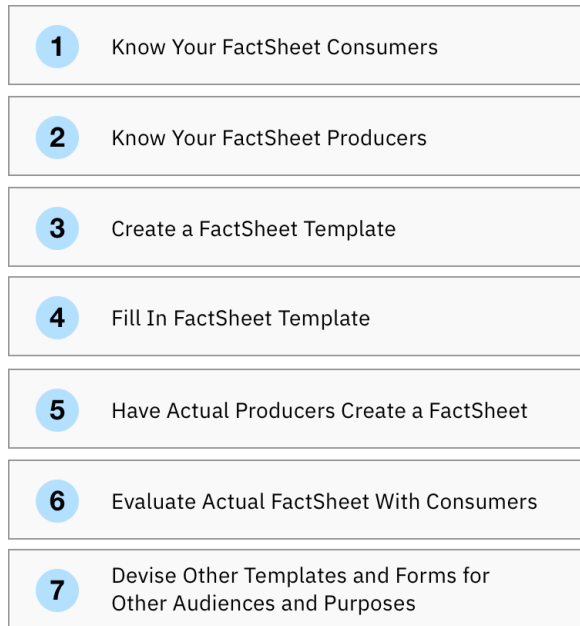


Figure 2: Steps to produce useful FactSheets

schemes within which these industries operate.

Within a particular domain or organization, FactSheets will also take on different forms, and contain different content, for different purposes. Model validators may need detailed information on data selection and cleaning, feature engineering, and accuracy and bias metrics. Business owners may need information on whether a deployed model is meeting business needs. Regulators may need a report detailing how a model complies with established practices and metrics related to safety, bias, and harm. Thus, although there is a strong desire to create a standard template for all FactSheets, we believe this diversity illustrates that for FactSheets, **one size does not fit all**.

We believe that standards will eventually emerge and, like nutrition labels, be useful for some purposes. In the foreseeable future, however, many kinds of FactSheets will be created. We have created the notion of **FactSheet Templates** to manage this diversity. A FactSheet Template can be thought of as specifying the categories or types of information that will be collected and displayed during and after AI development. Any given lifecycle will likely have multiple templates since different people will likely want to see different information, for different purposes, at different points in time. A large part of the job of creating FactSheets is designing the appropriate FactSheet Template(s). This is a prime focus of Section 4.

4 FactSheet Methodology

We now describe our seven-step methodology for the construction of useful FactSheets. For expository purposes, the steps shown in Figure 2 are presented as though they flow in a single stream from beginning to end. The reality is that FactSheet production is highly iterative, especially in the early days of FactSheet adoption within an organization.

Each step lists the key roles involved. In addition to the more typical roles shown in Figure 1, an additional role is identified, namely the “FactSheets Team”. This team is responsible for designing and implementing the FactSheets process within the organization. The first three steps will be driven by this team as they interview potential FactSheet consumers and producers and design the first FactSheet Template. Step 4 will largely be performed by the FactSheets Team but will benefit from the involvement of those with direct knowledge of the model or service being documented. This step may involve several iterations and informal trials with potential

consumers and producers. In Step 5, FactSheet producers will generate an actual FactSheet. In Step 6, FactSheet consumers will assess the quality and usefulness of this FactSheet. The FactSheets Team will be involved in these latter steps as well but will rely heavily on others to produce and attempt to consume actual content. In Step 7, the FactSheets Team repeats the process to increase coverage and value.

To simplify the presentation in the following steps we focus on one fact producer, “Priya”, and one fact consumer “Carmen”. Priya is a data scientist who will generate facts about how she created her model. Carmen is a model validator who will assess the model Priya created on various dimensions including quality, simplicity, and potential risk. Of course, Priya may also be a consumer of facts produced earlier by those who assembled the training data she uses. Similarly, Carmen may be a producer of facts for those who make the final decision on deployment readiness of the model she validates. Although our consumer in this example, Carmen, is part of the AI lifecycle, there are other possible documentation consumers that are outside of the AI lifecycle, such as end users (e.g., a loan officer), affected users (e.g., a loan applicant), or regulators. The same methodology would apply in these cases as well.

This may seem like a lot to think about, especially when there are multiple roles to understand and a desire to sample multiple representative users within each role. But the important thing is to start. Find one person performing each role (some people will be performing more than one role). Spend 30 minutes in conversation with each of them. If needed, find more than one person to explore areas that are still unclear after the first conversation. To speed things up, consider bringing potential producers and consumers together in conversation at any point in this process. They may quickly converge on what information is needed and how it can be produced in a cost-effective way.

4.1 Step 1: Know Your FactSheet Consumers

- Who: FactSheets Team (with potential consumers)
- What: Gather the information needs of potential FactSheet consumers

FactSheets are produced so that they can be consumed. Understanding the information needs of FactSheet consumers is the first and most important task. Here are some of the questions to consider in this first step (with Carmen, a model validator, as the illustrative consumer):

1. What does Carmen currently do when she performs her role?
2. What is Carmen going to be asking for when looking at a FactSheet?
3. What decisions will she be making based on the information presented?
4. How is the FactSheet going to help her do her job more effectively?
5. What are the most important pieces of information that Carmen needs to know?
6. What is Carmen’s level of expertise in general data science?
7. How is Carmen’s expertise going to affect the information presented?
8. Will there need to be additional definitions for terms that Carmen is unfamiliar with?
9. What is Carmen’s level of expertise with respect to the model algorithms being used?
10. What explanations about the model’s algorithm or results is Carmen going to need?
11. What is Carmen’s level of expertise in the problem domain?
12. How is that going to affect the information presented?
13. Will Carmen need help in mapping general knowledge of the problem domain to the particular inputs, outputs, or performance indicators associated with this model?
14. Is Carmen aware of issues related to model risk, potential harm, and regulatory compliance?
15. What information is needed to assess these issues?

4.2 Step 2: Know Your FactSheet Producers

- Who: FactSheets Team (with potential producers)
- What: Gather the kinds of information FactSheet producers might generate

Some facts can be automatically generated by tooling. Some facts can only be produced by a knowledgeable human. Both kinds of facts will be considered during this step. Here are some of the questions we might explore with Priya (a data scientist) about the facts she could usefully generate during the creation of a model:

1. What facts does Priya wish she could conveniently record about the models she develops? It is often helpful to ask about the most recent model, or a model that was particularly important, or a model that was exceptionally difficult to produce, rather than discussing models in general.
2. What did Priya do during the creation of this model that is otherwise unknown to others?
3. Are there general facts about the data, the features, the model algorithm, or the training and testing Priya performs that are important to note? Why?
4. What model-specific knowledge does she have that may not be obvious to others?
5. What domain-specific knowledge does Priya have that may not be obvious to others?
6. Does Priya know who will be consuming the facts she produces? We will assume it is Carmen in this particular case. Does Priya know Carmen? Have they talked about what Carmen needs to know?
7. Is Priya aware of issues related to model risk, potential harm, and regulatory compliance?
8. What information will be needed by others to assess these issues?

4.3 Step 3: Create a FactSheet Template

- Who: FactSheets Team
- What: Define the topics and questions to be included in FactSheets

What is learned in the first two steps leads directly to the most important part of creating FactSheets, namely the creation of a FactSheet Template. As discussed in Section 3, a FactSheet Template will contain questions. Each individual FactSheet will contain the answers to these questions. For example a template may start with the question “What is this model for?”. It may then expand on that question by asking where the model is well-suited and where the model is ill-suited.

The information gathered in the first two steps will inform the creation of this FactSheet Template. You may find that details about how a model is created are much less important in your organization than information about risk assessments and regulatory compliance. Or you may find that detailed questions about robustness against adversarial attacks are needed because of the nature of the models you create or the high-stakes domains within which they are used.

Here are some of the questions to consider in creating the first iteration of a FactSheet Template. Again, this is cast in terms of Carmen’s needs for information and Priya’s ability to produce that information, but similar questions will apply to many of the roles in the AI lifecycle or external consumers of the AI documentation.

1. What are the topics or categories of information needed?
2. Do some of these categories have subcategories?
3. What is a meaningful name for each category or subcategory?
4. What kinds of information should be included in each category? For example, Carmen may want to group all the model performance metrics within a category called “Model Performance”. Information about the representativeness of the training data might be grouped with information on the sensitivity of the model to

drift in a category called, “Potential Sources of Error”.

5. How should each question in a category be worded so as to be both understandable and evocative for Priya? The goal here is to encourage fact producers to answer in ways that are concise, germane, and understandable.
6. Where will the answer to a question come from? Will it be generated automatically by a tool or entered by a knowledgeable human? If the former, will Priya have some control over the frequency of fact generation or the granularity of recorded facts? If the latter, will Priya be given hints or examples of the kind of answer that would be satisfactory?
7. Are there any regulatory, legal, or business concerns that need to be considered when answering the questions in this template?
8. Are there different presentation formats needed for this information (for example, a short tabular summary of just key facts, or a slide format for presentations to review boards)? AI FactSheets 360 [10] shows three different formats that might be useful.
9. In addition to the human-readable content, is there a need for machine-readable content that Priya might generate?

4.4 Step 4: Fill In FactSheet Template

- Who: FactSheets Team
- What: Informally assess FactSheet Template by trying to fill it in

This step is where you will attempt to fill in your FactSheet Template for the first time. As you do this, informally assess the quality of the template itself. While this assessment is not a substitute for further work with Priya and Carmen (to follow), it may quickly highlight where improvements are needed. In doing this assessment, try to reflect on the template and the FactSheets it will generate from Carmen’s and Priya’s points of view. Ask yourself, or other members of your FactSheets Team, the following questions:

1. Knowing what Carmen knows, will she be able to understand the information that filled-in FactSheets will include?
2. Are there details needed by Carmen that will be missing in these FactSheets?
3. Is there specialized language that Carmen will be unfamiliar with?
4. Will the information allow Carmen to make the decisions she needs to make?
5. How are these FactSheets going to help Carmen do her job more effectively?
6. What might we do to encourage Priya to answer questions in ways that provide what Carmen needs?

4.5 Step 5: Have Actual Producers Create a FactSheet

- Who: Business Owner, Data Scientist, Model Validator, AI Operations Engineer (and others as defined within your organization’s AI lifecycle)
- What: Populate a FactSheet Template with actual facts

At this point you have a solid template and a good sense of how it might be used to create FactSheets. The next step is to have actual fact producers fill in the template for their part of the lifecycle. If there is a question in the template about model purpose, find someone who would actually be entering that information and have them answer the question. Ask a data scientist to answer the questions related to the development and testing of an actual model. If this model was validated, ask the model validator to enter information about that process. Similarly, have a person responsible for model deployment answer those questions. If the lifecycle is not that structured, have the person responsible for most of the work create this FactSheet.

We have found this step to be highly iterative. You can expect sections of your template to be expanded, compressed, or eliminated altogether. Individual questions will be refined within these sections. Stay alert for ideas or helpful hints about other fact producers that may surface. Follow these leads later. The goal here is to create a FactSheet that is ready for evaluation by consumers in the next step. Take the time to get this FactSheet to a level of quality and completeness that will make this next evaluation meaningful.

4.6 Step 6: Evaluate Actual FactSheet with Consumers

- Who: Business Owner, Data Scientist, Model Validator, AI Operations Engineer (and others as defined within your organization's AI lifecycle)
- What: Assess FactSheet quality with those who will be consuming FactSheets in production

In this step we conduct an assessment of the quality and completeness of the actual FactSheet produced in the previous step. If the FactSheet is intended to be used by multiple roles (not uncommon), evaluate it separately for each role. To make each evaluation meaningful, ensure you have agreement with respect to the purpose of the FactSheet. Ask the consumer to imagine using this FactSheet to actually perform their work.

Each evaluation consists of two parts. The first focuses on the content in the FactSheet. The second focuses on the way in which information is presented.

Content Evaluation: The goal of this part of the evaluation is to see how well the content of the FactSheet meets the specifically-designed-for information needs of the consumer. Ask your consumer to go through the FactSheet item by item with their information needs in mind and identify the following:

1. What information is missing?
2. Why is that missing information important to include?
3. How would they like this information presented?
4. Can they give an example?
5. What information is extraneous?
6. Why is that information extraneous?
7. What information is confusing or hard to understand?
8. Why is that information hard to understand?
9. How can that information be made more understandable?
10. Can they give an example?
11. Was the organization of information sensible?
12. If not, what would they change?

Have the consumer rank the information presented in this FactSheet from most important to least important. Remember to include the information that was noted as missing in this ranking. If time permits, have them share their views about the FactSheet with your larger group. Encourage discussion and ask questions about any unexpected findings, which can often identify gaps in the underlying lifecycle process or confusion about roles. Addressing these gaps can pay large dividends.

Presentation Evaluation: The goal of this part of the evaluation is to see if the way that information is *presented* meets the specifically-designed-for information needs of the consumer. Since some of the information you collect may be visual, make sure to allow for that type of feedback. Ask each consumer to go through the FactSheet item by item with their information needs in mind and identify these things:

1. Is this information presented in an unexpected way?

2. How can the information be presented differently?
3. Why is this alternative a better way to present this information?
4. Can they draw or describe an example?
5. If the information presentation includes interactive elements, are they useful?
6. How can they be made more useful?
7. Why is that more useful?
8. If they could add or change the way that information is presented, how would they?
9. Why is this addition or change an improvement?
10. Is this, overall, the right format for presenting this information?
11. What format would be more suitable?
12. Why is that format more suitable?

4.7 Step 7: Devise Other Templates and Forms for Other Audiences and Purposes

- Who: FactSheets Team (and others as appropriate)
- What: Evolve existing templates and create new ones

By now you will have created a refined FactSheet Template for use by others. They will be able to create useful and consumable FactSheets with that template. But there is more to do. There may be other consumers that need to be supported. Perhaps it is time to turn from an inward focus to an outer one, crafting templates for FactSheets to be consumed by external review boards or regulators. Or it may be time to support other stakeholders not directly involved in the AI lifecycle, such as sales personnel or the ultimate consumers of an AI service. Other formats for the same content may need to be created as well. The above steps can be followed once again. You will have learned a surprising amount about *how* to create FactSheet Templates and FactSheets from having gone through this process once. It will go faster and more smoothly now.

We encourage an ongoing process of reflecting on how well FactSheets support your AI lifecycle once they are fully incorporated and in routine use. Consider how they might be improved. Perhaps a new business opportunity in a new domain has developed or new types of models are being created that capitalize on new algorithmic research. If so, it may be time to refine existing FactSheet Templates or create new ones.

5 Further Guidance

We have observed [8] that producers of FactSheets have a hard time imagining what consumers of FactSheets need to know and how best to provide that information. Model developers, for example, may have a sophisticated understanding of the algorithmic basis for a model, but may describe the model or its performance in ways that assume far too much knowledge on the part of a FactSheet consumer. Consumers may not really know what information they need to support their work without somewhat structured reflection. Our methodology addresses these gaps by applying a user-centered design process [14] to the task of creating useful AI documentation. This process need not be time consuming and expensive. Even talking with a few potential FactSheet consumers and producers will be helpful.

It should be obvious at this point that following this methodology will not lead to a single FactSheet Template across the vast array of organizations creating AI models and services. The methodology *will*, however lead to FactSheets that fit the needs of a particular organization and provide real value to the corresponding AI development, deployment, and monitoring teams.

To put it a different way, one size will not fit all, at least if you dive below a short nutrition-label-like form to something that provides useful detail to all the lifecycle roles in a real organization. Even FactSheets developed with the same template will differ in interesting ways. For example, some models will have FactSheets with

extensive sections on bias and fairness testing with respect to protected populations. Other models will have FactSheets for which fairness and bias considerations are truly not applicable. Within some regulated industries, FactSheets may run to a hundred or more pages whereas the FactSheet produced by a startup company providing an AI component for visual object detection may be little more than a statement of purpose, inputs, and outputs.

An extension beyond user-centered design is *participatory design*, which invites not only the producers and consumers within the organizations’s AI development team to contribute to the process, but also the communities affected by the deployed model or service, such as applicants or patients [12, 18]. Moreover, by including people with lived experience of marginalization, who have an epistemic advantage in spotting potential harms, you will obtain a more comprehensive FactSheet template than if you did not have their participation [6].

This methodology for creating FactSheets may seem like a lot of work. Following these steps *will* take more time than just having a single person write a FactSheet Template based on a limited understanding of the actions and information needs within your organization. But failing to perform these steps will incur ongoing costs in poor documentation, repeated requests between team members for missing information, insufficient testing based on faulty assumptions about data or model structure, sub-optimal business results, and exposure to unnecessary risk.

We have found that following these steps with even a small number of people, where there is perhaps only one representative for each stage in a lifecycle, will pay dividends. We have also found that iterating quickly, rather than spending substantial time trying to attain perfection within each iteration, will shorten the overall time needed.

6 Harm and Safety

The increasing use of AI systems in high-stakes decision making has underscored the importance of transparent reporting mechanisms. These mechanisms, including FactSheets, can lead to better understanding, and more effective mitigation of any harm or safety issues in the system, such as bias, vulnerabilities to adversarial attacks, or other undesirable societal impacts. For example, a section that describes a detailed analysis of bias in the training dataset can help illuminate if the system is appropriate for a particular use case.

This paper describes a methodology for producing a useful transparent reporting mechanism for AI systems. This methodology can contribute to the identification of potential harm and safety issues. The methodology does this by:

- Explicitly including multiple FactSheet consumers and producers in FactSheet requirements gathering (Steps 1–2)
- Asking questions about their concerns for harm and risk (Steps 1–2)
- Providing a feedback mechanism to allow further input (Step 6)
- Including a broad range of perspectives in the development of FactSheets (Steps 1–7)

This process will increase the likelihood that FactSheets will provide the information needed to understand and mitigate potential harm or safety issues with an AI system.

7 In Practice

We have begun evaluating the FactSheets methodology across three teams within our company and have received strong positive feedback and calls for widespread adoption for both new and existing models and services. One early benefit has become evident from the work carried out by fact producers, who were primarily data scientists. They found that the step of identifying consumers and their documentation-related use cases provided them with a perspective and a sense of purpose that was lacking in their prior documentation efforts. They described

how having a persona (or sometimes a specific person) in mind enabled them to more carefully shape their documentation to meet known needs, a strategy used by data scientists more generally when communicating about their models [17]. By having specific users in mind, data scientists were able to constrain *what* facts to document and *how* to present them, lessening the uncertainty that they reported experiencing in the past.

The benefits of the FactSheet methodology do not come without costs. Our FactSheet creators spent up to 24 working hours crafting a complete FactSheet, with roughly half that time spent gathering feedback from consumers and iterating on content to make it more consumable. These costs can be reduced with better technology to support the creation and curation of facts. But we note that the multidisciplinary nature of AI model development, with each role having their own distinct knowledge, information needs, and preferred tools for accomplishing their work, will continue to require a focus on collaborative activities with their attendant costs and complexities (such as scheduling meetings). Bridging the gap between roles while addressing the back-and-forth, iterative nature of creating FactSheets remains a challenge to be overcome.¹

References

- [1] M. Arnold, R. K. E. Bellamy, M. Hind, S. Houde, S. Mehta, A. Mojsilović, R. Nair, K. N. Ramamurthy, A. Olteanu, D. Piorkowski, D. Reimer, J. Richards, J. Tsay, and K. R. Varshney. FactSheets: Increasing trust in AI services through supplier’s declarations of conformity. *IBM Journal of Research & Development*, 63(4/5), Sept. 2019.
- [2] E. M. Bender and B. Friedman. Data statements for natural language processing: Toward mitigating system bias and enabling better science. *Transactions of the Association of Computational Linguistics*, 2018.
- [3] K. S. Chmielinski, S. Newman, M. Taylor, J. Joseph, K. Thomas, J. Yurkofsky, and Y. C. Qiu. The dataset nutrition label (2nd gen): Leveraging context to mitigate harms in artificial intelligence. In *NeurIPS 2020 Workshop on Dataset Curation and Security*, Dec. 2020.
- [4] European commission. Europe fit for the digital age: Commission proposes new rules and actions for excellence and trust in artificial intelligence. Brussels, Belgium, Apr. 2021.
- [5] European Union. CE Marking. https://europa.eu/youreurope/business/product-requirements/labels-markings/ce-marking/index_en.htm#shortcut-3.
- [6] S. Fazelpour and M. De-Arteaga. Diversity in sociotechnical machine learning systems. arXiv:2107.09163, 2021.
- [7] T. Gebru, J. Morgenstern, B. Vecchione, J. Wortman Vaughan, H. Wallach, H. Daumé, III, and K. Crawford. Datasheets for datasets. *Communications of the ACM*, 64(12):86–92, 2021.
- [8] M. Hind, S. Houde, J. Martino, A. Mojsilovic, D. Piorkowski, J. Richards, and K. R. Varshney. Experiences with improving the transparency of AI models and services. In *Extended Abstracts of the 2020 CHI Conference on Human Factors in Computing Systems*, CHI EA ’20, page 1–8, New York, NY, USA, 2020. Association of Computing Machinery.
- [9] S. Holland, A. Hosny, S. Newman, J. Joseph, and K. Chmielinski. The dataset nutrition label: A framework to drive higher data quality standards. arXiv:1805.03677, May 2018.
- [10] IBM Research. AI FactSheets 360 website. <https://aifs360.mybluemix.net>.

¹An earlier version of this paper [20] includes two illustrative FactSheet templates created using this methodology. The templates are for the same model, but for two different consumers, illustrating the generality of the approach.

- [11] IEEE. P7006 — standard for personal data artificial intelligence (AI) agent. <https://standards.ieee.org/project/7006.html>, 2017.
- [12] M. K. Lee, D. Kusbit, A. Kahng, J. T. Kim, X. Yuan, A. Chan, D. See, R. Noothigattu, S. Lee, A. Psomas, and A. D. Procaccia. WeBuildAI: Participatory framework for algorithmic governance. *Proceedings of the ACM on Human-Computer Interaction*, 3(CSCW):181, 2019.
- [13] M. A. Madaio, L. Stark, J. Wortman Vaughan, and H. Wallach. Co-designing checklists to understand organizational challenges and opportunities around fairness in AI. In *Proceedings of the CHI Conference on Human Factors in Computing Systems*, page 318, 2020.
- [14] J.-Y. Mao, K. Vredenburg, P. W. Smith, and T. Carey. The state of user-centered design practice. *Communications of the ACM*, 48(3):105–109, 2005.
- [15] M. Mitchell, S. Wu, A. Zaldivar, P. Barnes, L. Vasserman, B. Hutchinson, E. Spitzer, I. D. Raji, and T. Gebru. Model cards for model reporting. In *Proceedings of the ACM Conference on Fairness, Accountability, and Transparency*, Atlanta, USA, Jan. 2019.
- [16] Office of the Comptroller of the Currency. <https://www.occ.gov/news-issuances/bulletins/2011/bulletin-2011-12.html>, Apr. 2011. OCC Bulletin 2011-12.
- [17] D. Piorkowski, S. Park, A. Y. Wang, D. Wang, M. Muller, and F. Portnoy. How AI developers overcome communication challenges in a multidisciplinary team: A case study. *Proceedings of the ACM on Human-Computer Interaction*, 5(CSCW1):1–25, 2021.
- [18] V. Prabhakaran and D. Martin, Jr. Participatory machine learning using community-based system dynamics. *Health and Human Rights Journal*, 22(2):71–74, 2020.
- [19] I. D. Raji and J. Yang. ABOUT ML: Annotation and benchmarking on understanding and transparency of machine learning lifecycles, 2019.
- [20] J. Richards, D. Piorkowski, M. Hind, S. Houde, and A. Mojsilović. A methodology for creating AI FactSheets. <https://arxiv.org/abs/2006.13796>, June 2020.
- [21] The European Commission’s High-Level Expert Group on Artificial Intelligence. Ethics guidelines for trustworthy AI. Brussels, Belgium, Apr. 2020.
- [22] United States Consumer Product Safety Commission. Testing and certification. <https://www.cpsc.gov/Business--Manufacturing/Testing-Certification>.
- [23] United States Consumer Product Safety Commission. Third party testing. <https://www.cpsc.gov/Business--Manufacturing/Testing-Certification/Third-Party-Testing/>.