AMENDED Exhibit 77

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

A Vision for Responsible AI (Was "SAIL 2020")

AI is fundamental to Facebook's mission. It helps us keep our community safe and solve integrity problems at scale, it improves the relevance of the content people see and helps them build meaningful connections, and it enables new product experiences in AR and VR.

AI is a general purpose technology that is transforming every field it enters. The stakes are high. The use of AI is under immense scrutiny because the risks and unintended consequences of its wide deployment are not well understood. Facebook's use of AI is no exception. People expect Facebook to develop and deploy AI in our products responsibly.

Our investments in the responsible development of AI have started later and are smaller compared to other technology companies that are similarly powered by AI. As a result, our main concern is when it comes to responsible AI, people expect and will expect more from us than we are able to deliver. However, we believe we still have an opportunity to catch up and potentially get ahead of the curve.

This note outlines some of the challenges posed by the responsible development of AI at scale, and proposes elements required to meet them. The Society & AI Lab (SAIL) team should play a key role in materializing this.

Outline - REVIST ONCE WE FINALIZE STRUCTURE

Objectives
Problem Areas
Strategy
Elements Needed for Success
Discussion
Appendix A: Example challenges and opportunities
Appendix B: Sources and Competitive Landscape

What Success Looks Like

Facebook becomes a trusted leader in the responsible development and use of AI. Facebook shares best practices with the industry and wider community.

The Challenge

AI offers unprecedented opportunities to benefit people and society. But it also brings new urgency to old social questions and magnifies underlying challenges that have vexed Facebook for years and society since existence.

Many of these questions have no technical solution; they are questions of principles and values. They are questions about fairness, safety, privacy, and accountability. **At their core, they are about our social responsibilities and how decisions related to those responsibilities should be governed.**

Furthermore, we have only just begun to understand the difficult challenges posed by the widespread deployment of AI. Anticipating and addressing these challenges will require fundamental changes to how we conceive and build AI-powered products. And it requires us to take explicit responsibility and to have clear governance mechanisms to resolve the tensions and dilemmas that will inevitably arise.

The challenges we face can be grouped in four broad categories. This list is not comprehensive (or in priority order) but is consistent with the common problem areas industry leaders, civil liberty groups, and the academic world have identified.

Fairness and inclusiveness. Ensure AI systems work for everyone, and that they don't have unfair biases in favor of groups of people. This is challenging because the data AI is trained on can reflect biases that exist in society, and might not represent everyone.

Robustness, reliability and safety. Ensure AI systems work as intended, that they are robust to adversarial attacks and that they do not have unintended harmful consequences. This is a challenge because it's much harder to assure quality in AI systems than in traditional computer systems.

Privacy and security. Guarantee people's data will be used as they expect and not shared without their consent, or be vulnerable to attacks. Empower people with control over their privacy and ensure AI does not accidentally reveal private details about people.

Transparency and accountability. Enable stakeholders of AI systems (maintainers, users, policy makers) to understand and verify the system behavior conforms with the principles and responsibilities we commit to.

Two important remarks:

AI engineering excellence and a focus on quality are necessary in order to address the challenges above, but they are not sufficient. Some of the challenges above still stand after we've built very high quality AI systems that work as intended.

Addressing the challenges above will surface tensions, for example between privacy, profitability, transparency and fairness. This highlights the importance of having clear governance structures and escalation paths.

Strategy and Approach

To help Facebook become a trusted leader in the development and use of responsible artificial intelligence, SAIL must do three things:

[redacted]

Rationale: Why This is the Right Approach to Address the Challenges We Face

When we built the Fairness Assessment Framework we quickly realized **fairness isn't primarily an AI problem**. There is a multitude of implementation decisions outside AI in the product chain where questions of unfair bias arise. An example is whether the policies and review guidelines treat or impact every sensitive group equally. When it comes to fairness, **there is no neutral option**. The very

definition of fairness requires taking a position on the spectrum from equal treatment to equality of outcomes. This is not a choice an AI team can make: this is a question of what the product stands for, and **what responsibilities Facebook will take**, given society itself is marred with historical biases and injustice. Similarly, there are **difficult questions of governance**: who should decide what sensitive groups and whose perspective to protect against unfair bias in the context of a specific product?

Engineering excellence should also apply to AI. In the same way as checklists save lives in the aviation industry, and we get on airplanes confidently as a result, the impact of AI algorithms in products is so vast it is reasonable to require a similar checklist approach. In our informal assessment of ML engineering excellence we have found basic deficiencies in how data is collected and labeled (e.g. systematic human labeling errors in click bait that go undetected), and in the absence of baseline computations (e.g. text understanding algorithms 1000x cheaper in compute and memory perform equally as a number of production algorithms we compared with). We have observed ML teams in product groups are under pressure to ship key metrics gains within the half, and struggle to set aside dedicated resources to improve ML engineering excellence [redacted]

Elements Needed for Success

For the two reasons above we feel in order to succeed SAIL will need institutional support in two ways:

Company wide commitment to responsible AI
XFN support for "responsible innovation" with broader scope than AI

The details are not comprehensive and open for discussion, the intent is to capture elements of the support we believe will be required:

1/ Company wide commitment to responsible AI
It is very difficult to drive change at scale without a clear mandate or a clear articulation of what aspects of responsible AI Facebook will hold itself accountable to. To succeed we will need:

- Top down executive support that carries through to incentives
- Clarity on what societal impact derived from the use of AI in our products Facebook will take **responsibility** for
- A supporting set of **AI principles** that clearly state what Facebook will hold itself accountable to.
- A clear **governance** model that states who should be consulted and who will decide when ethical tensions arise in the application of AI to our products
- A strong culture of quality in the design and implementation of AI systems in our products

2/ XFN support for "responsible innovation" with broader scope than AI
[redacted]

Discussion

- What is the best way to create a company-wide commitment to responsible AI?
  - Can we get Mark to personally validate the work and commit the company to it?
  - Can we tie this work to the company's 4 priorities and Mark's listening tour?
  - Would it make sense for Facebook Inc to publish AI principles and report progress against them regularly?
- How do we best structure and support the "responsible AI" effort?
  - What role should SAIL play?
    - [redacted]
    - [redacted]
  - How does SAIL connect to the broader ethics and responsible innovation led by Margaret Stewart?

- Better Engineering is rewarded and encouraged but not strictly enforced. Are there precedents of other mechanisms to enforce best practices?
- We recommend having dedicated senior XFN leaders for responsible AI in Policy, Legal and Comms.

- What are our blind spots?
  - Should privacy-preserving AI become a major investment area, both from a technology and ethical and social implications perspective?

Appendix A: Examples challenges and opportunities

Here are some actual and hypothetical example problems and success states in these four areas:

- **Fairness and inclusiveness**
  - The Hate Speech classifier incorrectly over estimates the likelihood that content is violating when it is authored by Black people in the US.
  - The Community Review labeler population was predominantly composed of political liberals who made unfairly biased judgments leading to increased removal of conservative content.
  - The Marketplace ranking algorithm reflects biases that exist in society where people are less likely to transact with people from other groups (e.g. different perceived race), and as result down ranks listings historically disadvantaged minority groups
  - The Portal Smart Camera doesn't reliably detect a person of dark skin and so incorrectly crops the shot to someone of lighter skin sitting next to them
  - Success State
    - People with dyslexia can use an AI-powered Additional Writing Help tool which, unlike a traditional spell-checker, is tailored to the types of writing challenges they face and increases their confidence in authoring content on Facebook products
    - Facebook for the Blind uses AI to describe the content of images and videos for visually impaired people
- **Robustness, safety and well-being**
  - Our ranking algorithm distributes click bait that tricks people into commenting or producing other signals that will cause a post to be ranked high, leading low quality content to go viral.
  - Targeted advertising exacerbates an illness —alcoholism— by serving highly context-optimized alcohol ads.
  - Sophisticated reinforcement learning might increase the problem of tech addiction in Stories, News Feed or Instagram
- **Privacy and security**
  - The ads that you are served are so relevant that they feel privacy-violating because they're for a specific product you viewed on another site, or revealing that Facebook inferred something about you that you didn't intend to share (e.g. you are a privately expecting new parent being served diaper ads).
  - A "friend" uploads a photo of you to Facebook without your consent, it is exposed to your social network, and you don't have the power to remove it.
  - Your data is used to train a machine learning algorithm but you did not explicitly consent to this or intend to agree to your data being used for this purpose.
  - Success state
    - Whenever people **explicitly disclose** personal data (e.g. gender, age, family and relationships, address and places, current and previous jobs, life events, other interests, etc) they have a clear understanding of how this data will be used, including by AI algorithms, to affect their product experience
    - People are informed about the **attributes we infer** about them, are empowered to correct them if they are inaccurate (e.g. ads preferences) and they have a clear understanding of how this data will be used, including by AI algorithms, to affect their product experience
    - People understand what data we have on them from **external sources** (e.g. non-Facebook apps connected to their Facebook profile), and how this data is used, including by AI algorithms
    - People understand which of their data is shared with third parties
    - People have guarantees that their data will be protected from theft or unauthorized or non-consensual access by third parties.
- **Transparency and accountability**
  - Your data is used to train a machine learning algorithm but you didn't understand the implications of the data use policy before consenting to it. You may or may not learn that your data was monetized in this particular way.

- Your profile was deactivated for violating the Community Standards but you don't understand why or how to appeal the decision.
- Success state:
  - People understand why they saw (or didn't see) particular housing, credit or employment ads, and who (humans and AI algorithms) is responsible for every decision
  - Publishers understand when and why an AI takes down an article they've shared on a Facebook surface, and they know what they should do differently in order for this not to happen (e.g. algorithmic transparency).
  - We have implemented transparent and privacy-preserving mechanisms for assessing whether our AI-powered products are unfairly biased against protected classes or members of sensitive groups, and we regularly share such reports
  - People understand what actions Facebook is taking to avoid harm caused by AI (for example the propagation of misinformation that leads to violence or social unrest), and have access to an honest assessment of how effective those steps are
  - Facebook shares an annual "Responsible AI" report that identifies the issues we've been tackling along with steps taken and their effectiveness.

Appendix B: Competitive landscape and other sources

Google

- Google AI Principles:
  - Be socially **beneficial**
  - Avoid creating or reinforcing **unfair bias**
  - Be built and tested for **safety**
  - Be **accountable** to people
  - Incorporate **privacy** design principles
  - Uphold high standards of **scientific excellence**
  - Be made for uses that accord with these principles
- In addition to the principles, there's an explicit call out of the applications of AI Google will not pursue:
  - "Technologies that cause or are likely to cause overall harm. Where there is a material risk of harm, we will proceed only where we believe that the benefits substantially outweigh the risks, and will incorporate appropriate safety constraints."
  - "Weapons or other technologies whose principal purpose or implementation is to cause or directly facilitate injury to people."
  - "Technologies that gather or use information for surveillance violating internationally accepted norms."
  - "Technologies whose purpose contravenes widely accepted principles of international law and human rights."
- **Formal review structure to assess new projects, products and deals**, supported by:
  - A Responsible Innovation team led by Jen Gennai
    - "This group includes user researchers, social scientists, ethicists, human rights specialists, policy and privacy advisors, and legal experts on both a full- and part-time basis, which allows for diversity and inclusion of perspectives and disciplines."
  - A council of senior executives and a group of senior experts from across Alphabet (they claim to have reviewed 100 projects (source).
  - A senior executive from Alphabet told me they're incorporating an ethics checkpoint in the Google launch calendar.
- Ethics in technology and in ML training:
  - Ethics in Technology Practice training (in pilot phase, plan to roll out to the entire company)
  - Ethics speaker series
  - Ethics training module as part of the Google Machine Learning Crash Course

Microsoft

- AI principles in "The Future Computed - Artificial Intelligence and its role in society" book and in the "Our shared responsibility for AI" post:
  - Fairness
  - Reliability & Safety
  - Privacy & Security
  - Inclusiveness

    - Transparency
    - Accountability
- AI and Ethics in Engineering and Research (AETHER) Committee, composed of senior leaders from across Microsoft engineering, research, consulting and legal organizations. Reviews project proposals and has the power to block them, formulates internal policies.
- Disclaimers in earning calls that "New products and services, including those that incorporate or utilize artificial intelligence and machine learning, can raise new or exacerbate existing ethical, technological, legal, and other challenges, which may negatively affect our brands and demand for our products and services and adversely affect our revenues and operating results."

Non-for profit expert organizations

The Partnership in AI

- Composed of 80+ partners (less than 1/3 are for-profit, includes civil liberty groups like the ACLU)
- Has kicked off 6 Working Groups composed of experts from the current 80+ partners:
    - Safety-critical AI
    - Fair, transparent and accountable AI
    - AI, labor and the economy
    - Collaborations between people and AI systems
    - Social and societal influences of AI
    - AI and social good

The AI Now Institute

- Co-founded by Kate Crawford and based at NYU.
- Focuses on four areas:
    - Rights and liberties
    - Labor and automation
    - Bias and inclusion
    - Safety critical infrastructure
- Publishes annual reports; some highlights from the 2018 report:
    - warning AI can amplify already increasing inequality,
    - call to Governments to urgently regulate AI,
    - call for companies to waive trade secrets, be transparent and self-regulate

UK Government Data Ethics Framework

- Part of the National Data Strategy, aimed at ensuring the UK public sector gets ethics right in the use of data to deliver public services.
- Proposes a Data Ethics Workbook that spells out 7 principles
    - Start with clear user need and public benefit
    - Be aware of relevant legislation and codes of practice
    - Use data that is proportionate to the user need
    - Understand the limitations of the data
    - Ensure robust practices and work within your skillset
    - Make your work transparent and be accountable
    - Embed data use responsibly

The Oxford Internet Institute

- We have collaborated with them and solicited their input towards "Developing a Digital Ethical Framework" (the report is available from Norberto Andrade upon request).
- Key recommendations:
    - Adopt and implement five core Digital Ethics Principles:
        - Accountability
        - Empowerment
        - Fairness

- Transparency
- Trust

- Take the following set of actions:
  - Commit publicly to adopting and implementing the five core Principles
  - Establish a Digital Ethics Advisory Board
  - Create an internal cross functional Digital Ethics Team
  - Embrace community, collective interest and public good as a foundational goal
  - Create and keep updated a Digital Ethics Training Curriculum
  - Acknowledge, respect and maximize the value of diversity
  - Foster digital ethics in design
  - Create a Facebook Digital Ethics Lab
  - Publish a Digital Ethics Annual Report

Previous version: SAIL 2020 (V0)

Contact Support

Conversation History

Quipbot#22 left the thread *Nov 15 at 11:16 pm*

Quipbot#22 *Nov 15 at 11:11 pm*

======IMPORTANT======

This Quip file has been copied to Google Workspace (https://docs.google.com/document...) as part of the Quip2Google migration efforts. Quip is moving to **Read-Only**. Please work in the new Google file moving forward.

Learn more

https://fb.workplace.com/groups...

Questions?

Visit the Quip2Google Migration Feedback Group (https://fb.workplace.com/groups...).

Note

New edits in this Quip file will NOT be automatically synced to Google Workspace.

Quipbot#22 added Quipbot#22 to the thread *Nov 15 at 11:16 pm*

Quipbot#22 added you to a document *Nov 15 at 11:11 pm*

[redacted] enabled link sharing · View/Edit *May 9, 2019 at 11:08 pm*

[redacted] opened your conversation with [redacted],