Debevoise & Plimpton

Comparative Approaches to AI Transparency

Lessons from Recent Regulatory Efforts

Anna Gressel, Debevoise & Plimpton LLP

FDA Virtual Public Workshop – Transparency of Artificial Intelligence/Machine Learning-Enabled Medical Devices

October 14, 2021

AI Regulatory Considerations

Data and Privacy Rights

Transparency and Explainability

Consumer Protection and Disclosures

Bias and Discrimination

Model Security and Resilience

Model Risk Management and Oversight

Board and Senior Management Supervision

Third-Party Oversight

Automated Decision-making

Validation and Testing

European Commission Draft AI Act

Unacceptable Risk, *e.g.*:

- Real-time remote biometrics in public places for law enforcement
- Exploitative systems
- Social scoring

Prohibited

High-Risk, *e.g.*:

- Consumer creditworthiness
- Recruiting and workplace management
- Biometric identification
- Education/vocational training

Conformity Assessment

Limited Risk, *e.g.*:

- Emotion recognition or biometric characterization
- Chatbots and interactive AI
- Deepfakes

Heightened Transparency

Minimal Risk

Codes of Conduct

U.S. Financial Regulators on AI Transparency

Debevoise & Plimpton

Federal Trade Commission
Blog post: "Using Artificial Intelligence and Algorithms"

"[W]hen using AI tools to interact with customers (think chatbots), be careful **not to mislead consumers about the nature of the interaction**."

"If you **deny consumers something of value** based on algorithmic decision-making, **explain why**."

"If you might change the terms of a deal based on automated tools, make sure to tell consumers."

"[B]e careful **not to overpromise what your algorithm can deliver**."

Blog post: "Aiming for truth, fairness, and equity in your company's use of AI"

Securities and Exchange Commission
IM Guidance Update on Robo-Advisors

Robo-advisers should **disclose information regarding their business practices and related risks**, including:
- Use of an algorithm;
- Algorithm's functions, assumptions, limitations, and risks
- Circumstances that might cause the robo-adviser to override the algorithm
- Third-party involvement
- Degree of human involvement
- Data used by the algorithm, and how to update or correct that data

Consumer Financial Protection Bureau
Innovation spotlight: "Providing adverse action notices when using AI/ML models"

"The existing regulatory framework has **built-in flexibility that can be compatible with AI algorithms.**"

Federal Reserve Board
Speech: "Supporting Responsible Use of AI and Equitable Outcomes in Financial Services"

"[T]here **need not be a single principle or one-size-fits-all approach** for explaining machine learning models. Explanations serve a variety of purposes, and what makes a good explanation depends on the context. In particular, for an explanation to 'solve' the black box problem, **it must take into account who is asking the question and what the model is predicting**."

Towards a Transparency Taxonomy

Ability to Contest/Correct Decisions

Interaction or Use of AI

Decision Explanations,
Consequences, and Risks
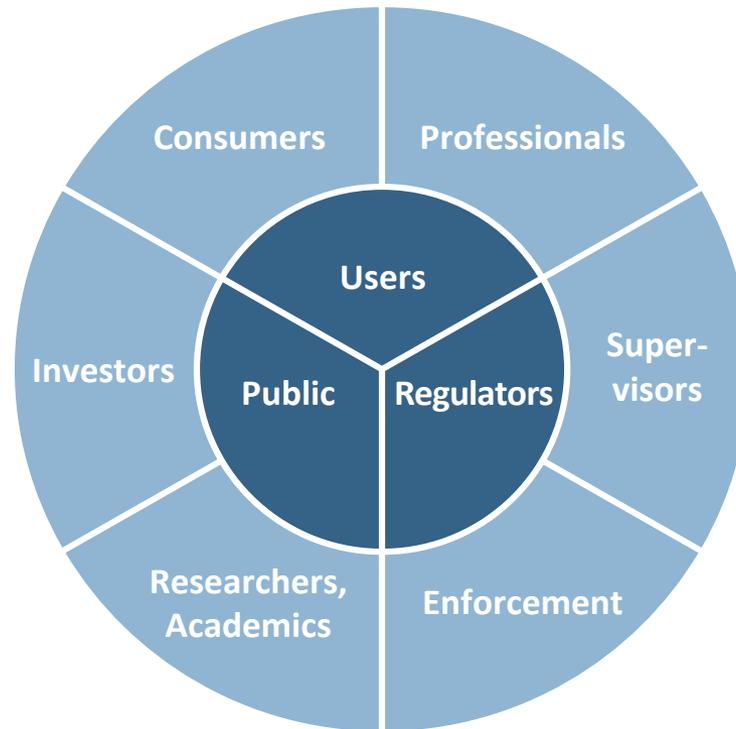
Material Business Risks
or Conflicts of Interest

Model Characteristics,
Capabilitiess & Limitations

List of Approved or
Marketed Systems

Technical
Documentation (Data,
Testing, Validation, etc.)

Use of Standards
and Certifications

Data or Model Logs;
Compliance Records

Consumers | **Professionals** | **Users** | **Super-visors** | **Investors** | **Public** | **Regulators** | **Researchers, Academics** | **Enforcement**

With You Today

Anna Gressel
Associate
+1 212 909 6485
argressel@debevoise.com

Anna R. Gressel is a senior associate in the firm's Data Strategy & Security Group. She actively advises clients on the legal and regulatory implications of artificial intelligence and other emerging technologies. Her practice includes not only representing companies in regulatory inquiries concerning AI, but also assisting companies in developing AI governance, compliance, and oversight mechanisms.