



ANNUAL MEETING

November 9-12, 2020 • Online Event

Data Ethics



Speakers

Peter Bothwell, FCAS

Managing Director, Data Science

Actuarial & Data Science

AF Group

Tracy Spadola, CPCU, CIDM, FIDM

Vice President

Strategic Data Operations

ISO / Verisk



Data Ethics

What is Data Ethics?

- Impact from Big Data

- Data Ethics vs Information Ethics

- Other Definitions of Data Ethics

Why Now

- Regulatory focus

- Minimize bias and social inequities

- Corporate and Public Perspective



Data Ethics

Impact on Data Professionals

Data Governance

Specialized Area of Governance

Codes of Behavior

Codes of Conduct exist for various professions

Guidelines

Includes 3rd Party Data

Personally Identifiable Data

Data Handling

Algorithms / Data Models

Business Practices



A Data Science Data Ethics Checklist

- Data Collection
- Data Storage
- Analysis
- Modeling
- Deployment



But does this apply to actuaries?

- Code of Conduct
- Principles
- ASOP 23

ASOP No. 23—Doc. No. 185

sets or **data** sources, if any, to be considered. The actuary should do the following:

- a. consider the **data elements** that are desired and possible alternative **data elements**; and
- b. select the **data** for the analysis with consideration of the following:
 1. whether the **data** constitute **appropriate data**, including whether the **data** are sufficiently current;
 2. whether the **data** are reasonable with particular attention to internal consistency;
 3. whether the **data** are reasonable given relevant external information that is readily available and known to the actuary;
 4. the degree to which the **data** are **sufficient**;
 5. any known significant limitations of the **data**;
 6. the availability of additional or alternative **data** and the benefit to be gained from such additional or alternative **data**, balanced against how practical it is to collect and compile such additional or alternative **data**; and
 7. sampling methods, if used to collect the **data**.



Collection

- Ensure Informed Consent
- Avoid Collection Bias
- Minimize PII
- License agreements
- Cambridge Analytica
- Twitter polls
- Modeling your book of business
- Oculus Quest 2
- Billing Data



- School used key-card to gather detailed behavioral data
- Gave data to a data science vendor
- Used advanced algorithms to identify at-risk students & recommend remediation
- Significant improvement in outcomes, but at what cost?



Storage

- Security
- Privacy
- Retention

- Breaches
- Stalking
- The right to be forgotten
- CCPA requirements
- Facebook
- Texts



Analysis

- Blind Spots & Bias
- Honest Representation
- Auditability
- Gender presuppositions
- Age vs Driving Experience
- Y-axis scope & scale
- Anomalies and transformations
- Documentation
- Repeatability



- Phone app that identifies song title and artist enhanced
- Able to identify both speaker and use internet to attach all available data
- Able to distinguish birth gender from identity gender and provide full history of changes
- How much should be invested to protect that information?



Modeling

- Proxy Discrimination
- Credit Scoring
- Territorial Rating
- Metric Selection
- Target variable choice
- Reason messages
- Communicate Weaknesses
- Potential bias
- Use beyond current book



Deployment

- Redress & Rollback
- Concept Drift
- Unintended Use
- Model fixes & updates
- Regulatory mandates
- Model inputs change behavior
- Environment or process changes
- Model abuse
- Agent gaming



- AI chatbot trained to “sound” like cybercriminal
- Used to troll the dark web offering to buy stolen identities
- Only pursued against individuals with existing criminal investigations
- But does it entice criminal behavior and entrap?



Conclusion

- None of this is new
- But is it part of our Best Practices?



Call to Action

Have the conversation
Review Governance Programs

Final Message



• Thank You!

