Chapter 13: Wrap Up and Future of Data Science

School of International Liberal Studies
Waseda University
Introductory Data Science

Wrap Up and Future of Data Science

1. Wrap Up and Future of Data Science
2. Accountability in AI
3. Data Governance
4. Data Security & Privacy
5. Limitations of Data Science
6. Future Trends in Data Science

Wrap Up and Future of Data Science

- Reflect on the broader role of data science in society
- Discuss ethics, fairness, transparency, and accountability
- Explore real-world cases of algorithmic bias
- Consider the future of data science: opportunities & challenges

What is Data Ethics?

- **Data ethics**: Moral principles for collecting, sharing, and using data
- Key questions in *data ethics*:
  - *Is this use of data fair and respectful?*
  - *Could this cause harm?*
- *Data ethics* covers:
  - Consent and transparency
  - Privacy and protection
  - Bias prevention and accountability

Example of Ethical Data Use: Health Research

- Ethics in medical data: **Informed Consent**: Participants must understand how their data is used
  Privacy Protection: Secure handling of sensitive health information
  Compliance: Follows ethical & legal standards (e.g., HIPAA, GDPR)

HIPAA: Health Insurance Portability and Accountability Act — U.S. law protecting health data privacy and security

GDPR: General Data Protection Regulation — EU law ensuring individual control over personal data

Guidelines for Ethical Data Practice

- Core principles to guide responsible data science: **Consent**: Clear permission before collecting data
  Privacy: Data use must match stated purpose
  Transparency: Be open about what, how, and why
  Accountability: Own your models and their impact
  Fairness: Prevent discriminatory or biased outcomes

How to Implement Ethics in Practice

- Putting ethical principles into action:
  - Privacy notices and opt-in forms
  - Anonymization and de-identification
  - Ethics reviews for sensitive projects
  - Follow professional codes of conduct (e.g., ACM Code of Ethics)

Opt-in forms: Explicit consent mechanisms that let users choose whether to share personal data

ACM Code of Ethics: Guidelines of Association for Computing Machinery for honesty, fairness, and privacy in tech and data use

What Is Algorithmic Bias?

- **Algorithmic bias**: Systematic unfairness in model outcomes
- Affects individuals/groups disproportionately
- Often rooted in:
  - Training data
  - Model design
  - Implicit assumptions

Bias $\rightarrow$ reinforces discrimination, reduces fairness

Source of Bias (1): Biased Training Data

- Training data may underrepresent or misrepresent groups
- Leads to unfair outcomes for minorities or outliers

Example: Amazon Resume Screening (2018)

- AI penalized resumes with "women's" (e.g., "women's chess club")
- Model was trained on 10 years of male-dominated hiring data
- Reproduced past hiring bias

Source of Bias (2): Design Choices and Data Handling

- *Data preprocessing* can introduce bias via:
  - Variable selection
  - Missing data imputation
  - Feature engineering
- **Example: Survey Imputation Bias**
  - Drug-use survey skips more common in high-risk groups
  - Filling missing data with averages $\rightarrow$ underestimates true risk
- **Example: Overfitting in Small Datasets**
  - Too many variables $\rightarrow$ model fits noise
  - Appears accurate in training but fails in real-world application

Source of Bias (3): Feedback Loops

- User behavior feeds future model decisions
- **Example: News Recommender**
  - Initially promotes sensational content
  - Users click $\rightarrow$ system amplifies extreme views
  - Echo chamber grows $\rightarrow$ more biased info

Algorithms shape the environment they learn from

What Is Accountability in AI?

- Accountability = *Who is responsible when AI causes harm?*
- Must remain **human-centered**:
  - People, not machines, are answerable
  - Ensures **oversight** and **corrective actions**

Assign roles, trace decisions, and offer redress for harm

Strategies for Ensuring AI Accountability (1/2)

- **Documentation & Audit Trails**
  - Keep detailed logs of design, data, and decisions
  - Enables diagnostics and assigns responsibility
- **Human-in-the-Loop**
  - Final decisions by humans, not AI
  - Examples: doctors verifying AI diagnoses

Strategies for Ensuring AI Accountability (2/2)

- **Compliance Checks & External Audits**
  - Review for fairness, privacy, and legality
  - Third-party audits increase transparency
- **Ethical Guidelines & Training**
  - Follow frameworks (e.g., *IEEE*, *EU Trustworthy AI*)
  - Train staff with real-world case studies

IEEE: Institute of Electrical and Electronics Engineers — provides global ethical standards like Ethically Aligned Design for responsible AI development

EU Trustworthy AI: A European Commission framework outlining 7 key requirements for AI, including transparency, accountability, and human oversight

What Is Data Governance?

- A framework for managing **data quality**, **security**, and **policy compliance**
- Aligns data with organizational goals
- Ensures responsible use across lifecycle

Enables consistent, secure, and trustworthy data handling

Data Governance in Practice

- **Example: Without Governance**
  - Inconsistent formats
  - Outdated data
  - Unsecured sensitive info
- **With Governance**
  - Single source of truth
  - Data stewards assigned
  - Access controls and policies

Principles of Data Governance

- Principles of data governance include:

Data Quality: Accurate, consistent, up-to-date
Security & Privacy: Encryption, access control, audits
Transparency: Documented rules, visible processes
Accountability: Named roles, responsible parties
Access Control: Least privilege & role-based access
Standardization: Clear formats, units, master records
Legal Compliance: Follow laws (e.g., GDPR, HIPAA)

What Are Security & Privacy?

- **Security**: Protects data from unauthorized access, breaches
  - *Tools*: encryption, authentication, access control, backups
  - *Goal*: ensure *confidentiality*, *integrity*
- **Privacy**: Ethical collection, use, and sharing of personal data
  - Focuses on *consent*, *data minimization*, and *purpose limitation*
  - Respects user rights and legal obligations

Security $\not\Rightarrow$ Privacy, but trong security protects data

Example: Online Health Platform

- **Privacy**
  - Collects only consented data
  - Limits use and sharing
- **Security**
  - Encrypts records
  - Uses strong authentication
  - Applies software updates

Key Threats to Data Security & Privacy

- These threats include:
  - **Cyber Attacks:** e.g., Yahoo breaches (2013–14)
  - **Insider Threats:** Staff misuse or errors
  - **Social Engineering:** Phishing, impersonation
  - **Physical Theft:** Stolen laptops (e.g., VA 2006)
  - **Software Bugs:** e.g., Heartbleed (2014)
  - **Unauthorized Sharing:** e.g., Facebook–Cambridge Analytica
  - **Re-identification:** Cross-referencing anonymized data
  - **Surveillance:** GPS, behavior tracking
  - **Data Breaches:** Identity theft, fraud
  - **Over-retention:** Old data = high risk

Ensuring Security & Privacy

- Methods to ensure security and privacy include:
  - **Encryption**: Secure storage and transmission
  - **Access Control**: MFA, least privilege
  - **Anonymization**: Strip identifiers
  - **Firewalls**: Protect networks
  - **Regular Updates**: Patch vulnerabilities
  - **Backups & Incident Response**
  - **Privacy Policies**: Data rules and regulations
  - **Data Minimization**: Only collect what's needed

Limitations of Data Science

- Data science is powerful, but not without limitations: **Dependent on Historical Data**: Cannot predict black swan events (e.g., COVID-19) **Sensitive to Outliers/Assumptions**: May fail on rare or non-conforming cases
  Requires High-Quality Data: "Garbage in, garbage out"
  May Lack Context: Data real-world nuance (e.g., toothbrush misclassified as gun)

AI best used as a decision support tool—not a full replacement for human judgment

Limitations of Data Science: Data Quality and Bias Risks

- **Poor-Quality Data**:
  - Missing values, errors → unreliable models
  - Rare events (e.g., rare diseases) are hard to model
- **Unrepresentative Data**:
  - Bias from unbalanced datasets (e.g., trials on one demographic)
  - Limits generalization to wider populations

Models are only as fair and accurate as the data they learn from.

Limitations of Data Science: Model Interpretability and Trust

- Complex models (e.g., deep learning):
  - high accuracy
  - low transparency
- Hard to explain $\rightarrow$ difficult to trust, audit, or correct errors
- **Bias risk**: Algorithms may replicate social inequalities (e.g., hiring models)

Explainable AI (XAI) is key to accountability and fairness

Limitations of Data Science: No Domain Knowledge

- AI lacks human intuition and context
  - e.g., Misclassifies objects based on superficial features
- Can't replace expert judgment, especially in sensitive domains

Data science works best when paired with human expertise.

Complexity Limits Predictive Power

- Many real-world systems (e.g., economies, ecosystems) are highly complex
- Models cannot capture all variables or feedback loops
- Example: Predicting recessions remains unreliable

Even the best models cannot fully predict chaotic, adaptive systems.

Organizational & Practical Constraints

- Real-world issues limit model deployment:
  - Shifting data (concept drift)
  - Resistance from users or stakeholders
  - High costs for data collection, training, maintenance

Simpler solutions are often preferred over complex models

Future Trends in Data Science

- Data science is evolving rapidly. Key frontiers:
  - **Generative AI** (e.g., ChatGPT, DALL-E)
  - **Sustainable AI** (Green AI)
  - **AI Governance & Regulation**
  - **Explainable AI (XAI)**

Innovation must balance capability with ethics and sustainability

Generative AI

- **Generative AI**: Models to generate text, images, and more:
  - GPT-4 (text)
  - DALL · E / Stable Diffusion (images)
- **Applications**:
  - Synthetic data generation
  - Chatbots and virtual agents
  - Creative prototyping (e.g., design, code, content)
- **Concerns**:
  - Job displacement in creative/analytical fields
  - Spread of misinformation (deepfakes, fake content)

Sustainable AI (Green AI)

- **Problem**: AI consumes significant energy (esp. large models)
- **Possible Solutions**:
  - Efficient algorithms
  - Model pruning (reduce size, preserve performance)
  - Specialized hardware (GPUs, TPUs, neuromorphic chips)

Sustainability is essential for responsible AI development