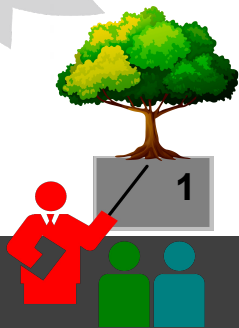Real-World Data Governance

The Data Scientist's Role in Governance Success

Monthly Webinar Series Hosted by DATAVERSITY

Robert S. Seiner – KIK Consulting
August 15, 2024 – 11:00 a.m. PT / 2:00 p.m. ET

Episode #151

The Data Scientist's Role in Governance Success
Introduction

Real-World Data Governance – Monthly Webinar Series

September 19, 2024: Navigating the Data Jungle – Catalogs, Glossaries, and Dictionaries Explored

Third Thursday each Month @ 2pm EST – Register at TDAN.com, KIKconsulting.com, DATAVERSITY.net

Non-Invasive Data Governance / Non-Invasive Data Governance Strikes Again Books

2014: ISBN 9781935504856 / TechnicsPub.com / Amazon.com

2023: ISBN 9781634623599 / TechnicsPub.com / Amazon.com

Non-Invasive Data Governance / Metadata Governance Online Learning Plans

Popular: Business Glossaries, Data Dictionaries and Data Catalogs
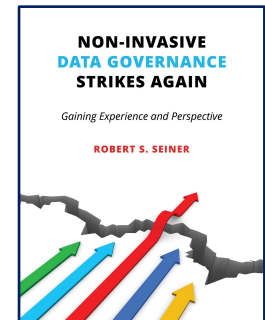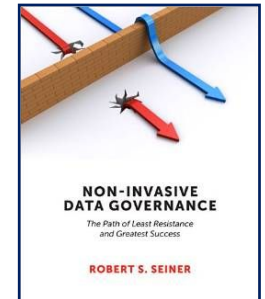
DATAVERSITY Training Center – https://training.dataversity.net

KIK Consulting & Educational Services

KIKConsulting.com

The Home of Non-Invasive Data Governance™

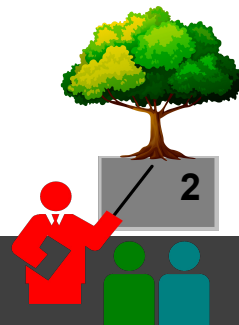### Carnegie Mellon University (CMU)

Adjunct: Heinz College Executive Education – **Chief Data and AI Officer (CDAIO)** Certificate Program

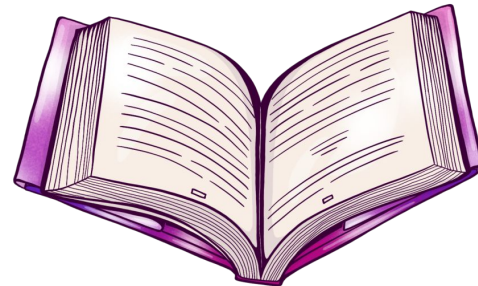# The Data Scientist's Role in Governance Success

Abstract

- In this webinar, I will talk about …
    - Unearthing Governance Opportunities in Data Science Practices
    - Empowering Data Scientists as Data Stewards
    - Integrating Governance into the Data Science Lifecycle
    - Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements
    - Real-World Success Stories of Data Scientists in Governance Success

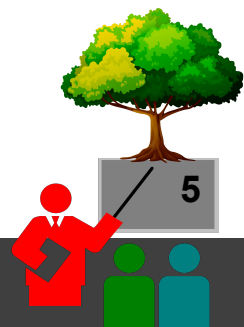# The Data Scientist's Role in Governance Success
Definitions

- **Data Governance** – The execution and enforcement of authority over data.

- **AI Governance** – The framework of policies, procedures, and standards established to guide the ethical development, deployment, and use of artificial intelligence systems.

- **Data Stewardship** – Formal accountability for data.

- **Data Steward** – A person held formally accountable for their relationship to the data.

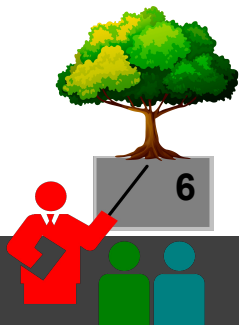# The Data Scientist's Role in Governance Success
Definitions

- Data Science – Data science is an interdisciplinary field that uses scientific methods, processes, algorithms, and systems to extract knowledge and insights from structured and unstructured data. It involves techniques from statistics, computer science, and information theory to analyze and interpret complex data. Data science aims to understand and leverage data to solve real-world problems and support decision-making processes.

  Harvard Business Review, Wikipedia, Data Science Central

- Data Scientist – A data scientist is a professional who uses expertise in statistical analysis, machine learning, and programming to analyze, process, and model data, ultimately extracting valuable insights that can inform decision-making. Data scientists apply their skills to identify patterns, predict trends, and develop data-driven solutions across various industries.

  HBR, IBM, Wikipedia
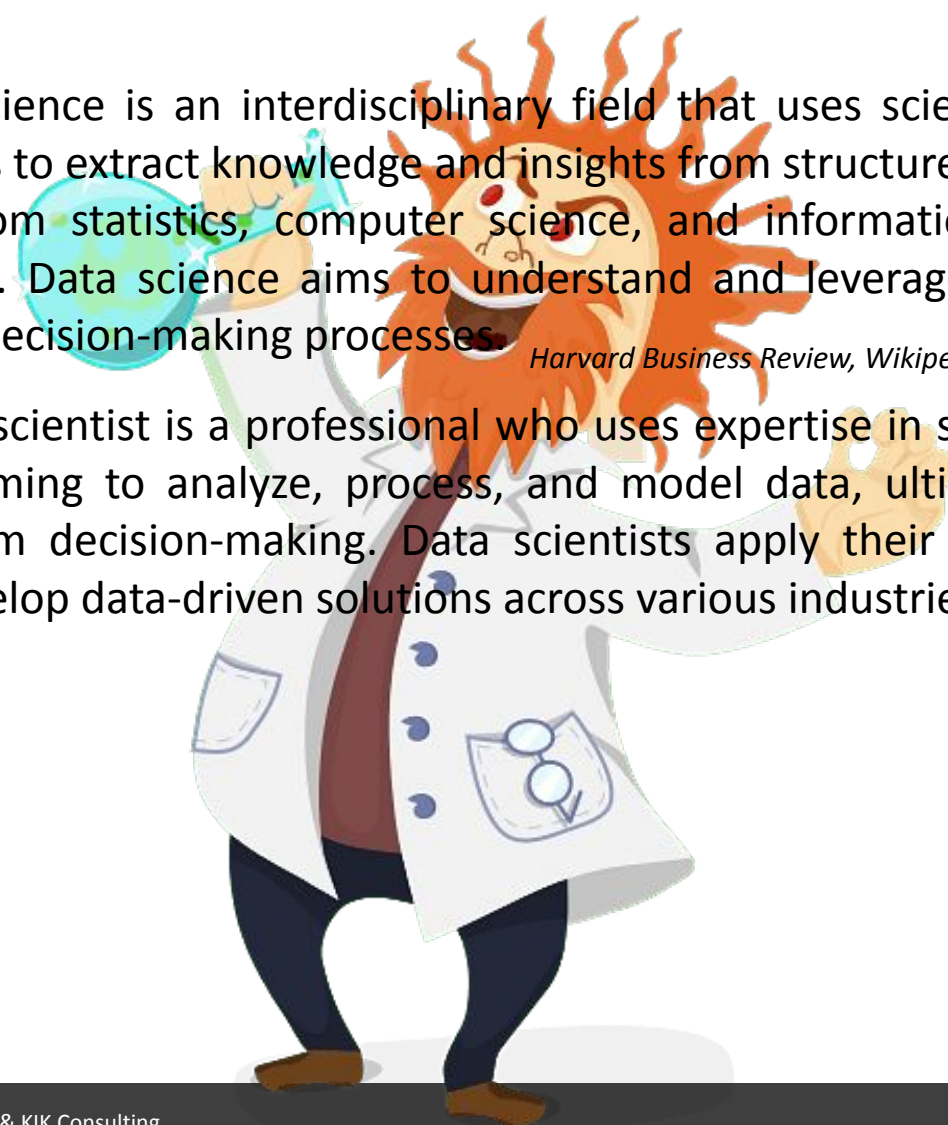
The Data Scientist's Role in Governance Success
Definitions

- Like a scientist – Let's dissect these definitions …

  - Data Science – Data science is an <u>interdisciplinary field</u> that uses <u>scientific methods, processes, algorithms, and systems</u> to <u>extract knowledge and insights from structured and unstructured data</u>. It involves techniques from statistics, computer science, and information theory to analyze and interpret complex data. Data science aims to understand and leverage data to solve real-world problems and support decision-making processes.

  - Data Scientist – A data scientist is a <u>professional who uses expertise in statistical analysis, machine learning, and programming</u> to <u>analyze, process, and model data</u>, ultimately <u>extracting valuable insights that can inform decision-making</u>. Data scientists apply their skills to identify patterns, predict trends, and develop data-driven solutions across various industries.
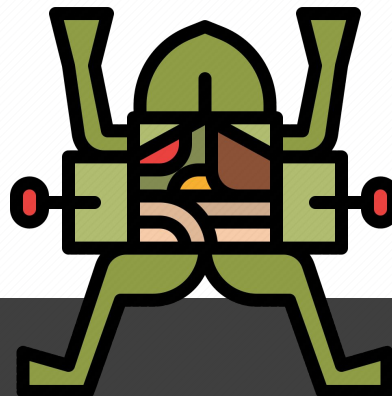
The Data Scientist's Role in Governance Success
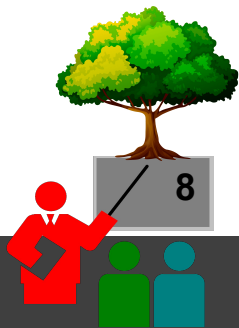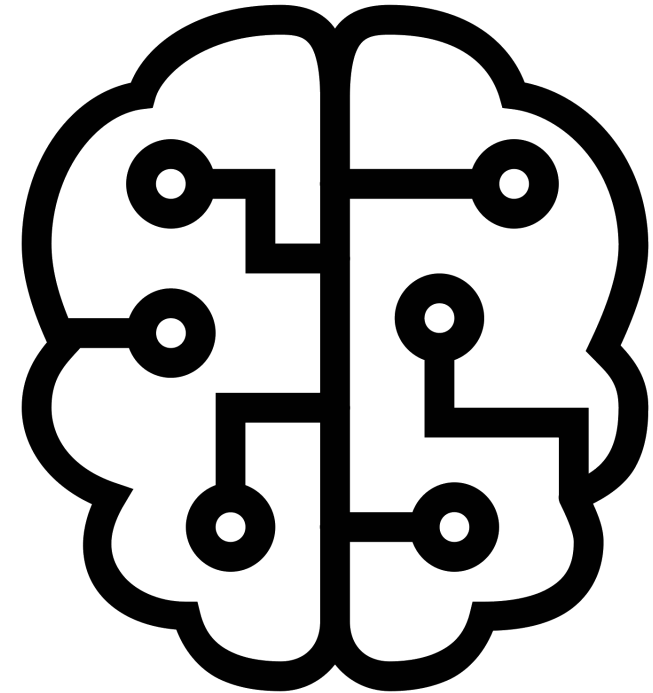## Unearthing Governance Opportunities in Data Science Practices

- Including Both Data Governance and AI Governance
  - **Ensuring Data Quality for AI Models:** Implement robust data governance practices to guarantee the quality and integrity of data, which is critical for training accurate and reliable AI models.
  - **Ethical AI and Data Usage:** Establish AI governance policies that enforce ethical guidelines and data privacy, ensuring responsible use of data in AI applications.
  - **Transparency and Accountability in Data Operations:** Leverage data governance frameworks to document data lineage and usage, enhancing transparency and accountability in AI processes and decision-making.
  - **Cross-Functional Collaboration for Governance Success:** Foster collaboration between data scientists, data stewards, and governance teams to identify and address governance opportunities, driving both data and AI governance success.
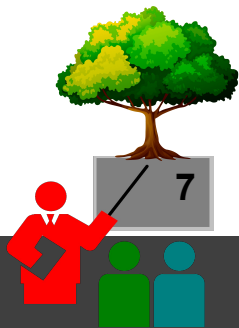
The Data Scientist's Role in Governance Success
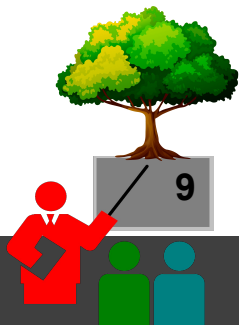## Unearthing Governance Opportunities in Data Science Practices

- Including Both Data Governance and AI Governance
  - Ensuring Data Quality for AI Models
    - Implement rigorous data validation processes.
    - Enforce consistent data standards across all sources.
    - Regularly audit data for accuracy and completeness.
    - Integrate automated data quality monitoring tools.

The Data Scientist's Role in Governance Success
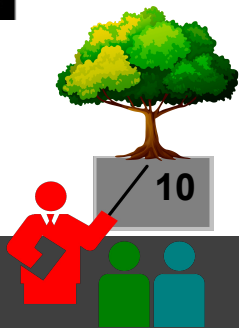## Unearthing Governance Opportunities in Data Science Practices

- Including Both Data Governance and AI Governance

  - Ethical AI and Data Usage

    - Implement fairness and bias mitigation techniques.

    - Ensure transparency in AI decision-making processes.

    - Adhere to privacy and data protection regulations.

    - Develop and follow ethical guidelines for AI deployment.

The Data Scientist's Role in Governance Success
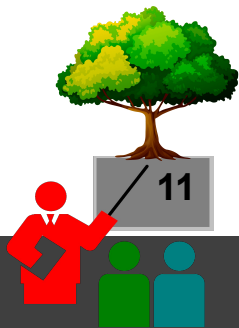## Unearthing Governance Opportunities in Data Science Practices

- Including Both Data Governance and AI Governance
    - Transparency and Accountability in Data Operations
        - Document data sources and processing workflows.
        - Establish clear data stewardship roles.
        - Implement audit trails for data usage and AI models.
        - Regularly review and update governance policies.

The Data Scientist's Role in Governance Success
## Unearthing Governance Opportunities in Data Science Practices

- Including Both Data Governance and AI Governance
  - Cross-Functional Collaboration for Governance Success
    - Foster communication between data scientists and data stewards.
    - Integrate AI governance in cross-departmental projects.
    - Align data and AI governance goals with business objectives.
    - Promote shared responsibility in governance initiatives.

The Data Scientist's Role in Governance Success
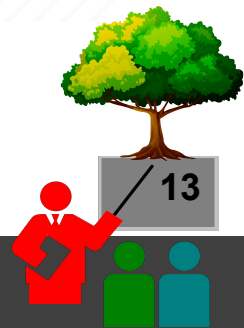## Empowering Data Scientists as Data Stewards

- Including Both Data Governance and AI Governance
  - **Championing Data Quality Initiatives:** Data scientists can lead efforts to maintain high data quality, ensuring reliable inputs for AI models and robust decision-making.

  - **Implementing Ethical AI Practices:** Empower data scientists to enforce AI governance policies that prioritize ethical AI usage, data privacy, and bias mitigation.

  - **Enhancing Data Transparency:** Encourage data scientists to document data lineage and usage comprehensively, promoting transparency and accountability in AI operations.

  - **Promoting Cross-Functional Collaboration:** Facilitate collaboration between data scientists, governance teams, and other stakeholders to align AI initiatives with governance standards and best practices.

  - **Driving Continuous Improvement in Governance:** Data scientists can advocate for ongoing governance improvements by identifying gaps and proposing solutions to enhance data and AI governance frameworks.
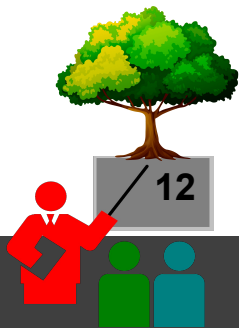
The Data Scientist's Role in Governance Success
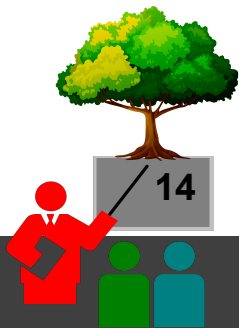## Empowering Data Scientists as Data Stewards

- Including Both Data Governance and AI Governance
  - Championing Data Quality Initiatives
    - Advocate for standardized data quality metrics and practices.
    - Implement rigorous data validation processes for AI models.
    - Promote continuous improvement of data quality standards.
    - Collaborate on data cleansing and enrichment efforts.

The Data Scientist's Role in Governance Success
Empowering Data Scientists as Data Stewards
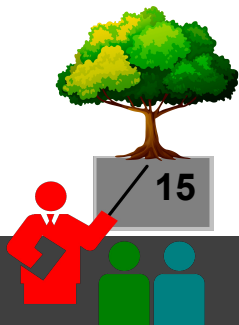
- Including Both Data Governance and AI Governance

  - Implementing Ethical AI Practices

    - Ensure AI models are free from bias and discrimination.

    - Adhere to regulatory and ethical guidelines in AI development.

    - Maintain transparency in AI decision-making processes.

    - Foster accountability through regular ethical reviews and audits.

The Data Scientist's Role in Governance Success
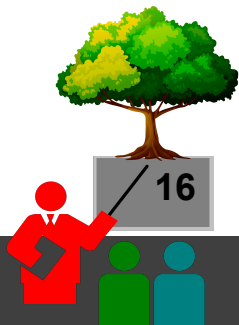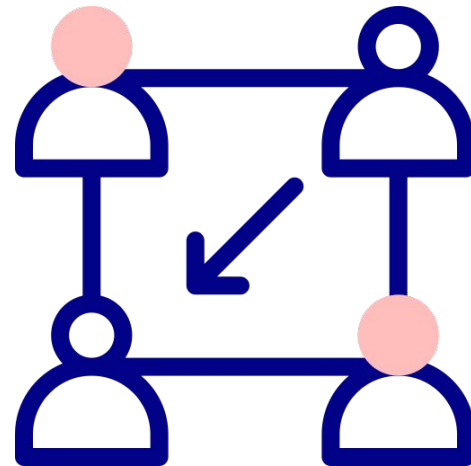## Empowering Data Scientists as Data Stewards

- Including Both Data Governance and AI Governance

  - Enhancing Data Transparency

    - Implement clear documentation practices for data sources and usage.

    - Foster open communication about data collection methods and limitations.

    - Ensure accessibility to metadata for improved data understanding.

    - Maintain detailed data lineage for tracking data transformations.

The Data Scientist's Role in Governance Success
## Empowering Data Scientists as Data Stewards

- Including Both Data Governance and AI Governance

  - Promoting Cross-Functional Collaboration

    - Foster teamwork between data scientists and business units.

    - Encourage joint projects with IT and compliance teams.

    - Facilitate regular cross-departmental meetings on data initiatives.

    - Promote shared goals for data and AI governance success.

The Data Scientist's Role in Governance Success
## Empowering Data Scientists as Data Stewards

- Including Both Data Governance and AI Governance
  - Driving Continuous Improvement in Governance
    - Implement feedback loops for ongoing governance enhancements.
    - Encourage iterative improvements in data and AI governance practices.
    - Monitor and refine governance frameworks regularly.
    - Leverage data insights to inform governance strategy adjustments.
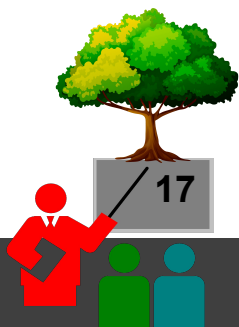
The Data Scientist's Role in Governance Success
Integrating Governance into the Data Science Lifecycle
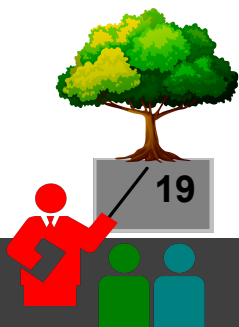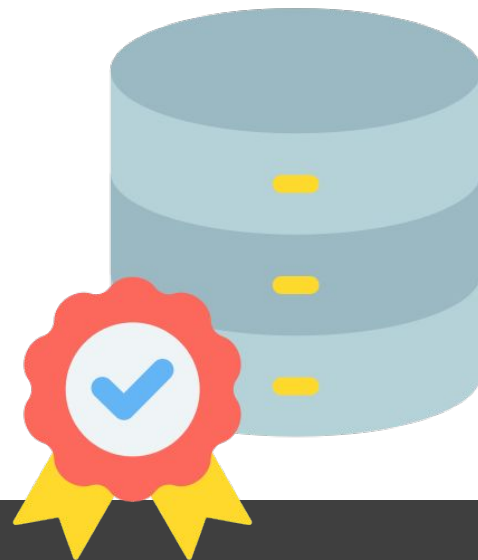
- Including Both Data Governance and AI Governance
  - **Ensuring Data Quality from Ingestion to Deployment:** Implement governance practices at each stage to maintain data accuracy and reliability throughout the data science lifecycle.

  - **Embedding Ethical AI Standards:** Integrate AI governance principles early in the model development process to ensure ethical considerations and compliance are addressed.

  - **Documenting Data Lineage and Model Decisions:** Maintain comprehensive records of data sources, transformations, and model decisions to enhance transparency and traceability.

  - **Facilitating Cross-Departmental Governance Collaboration:** Encourage data scientists to work closely with governance teams to align data and AI governance strategies, ensuring cohesive and compliant practices.
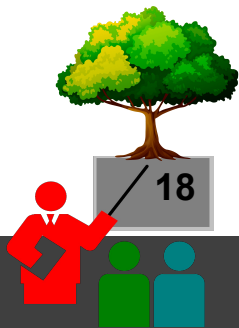
The Data Scientist's Role in Governance Success
Integrating Governance into the Data Science Lifecycle
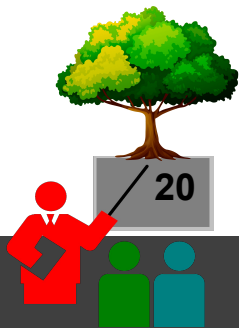
- Including Both Data Governance and AI Governance
  - Ensuring Data Quality from Ingestion to Deployment
    - Implement stringent data quality checks at ingestion.
    - Establish continuous monitoring for data quality during processing.
    - Enforce data validation protocols prior to model deployment.
    - Maintain comprehensive data documentation throughout the lifecycle.

The Data Scientist's Role in Governance Success
Integrating Governance into the Data Science Lifecycle
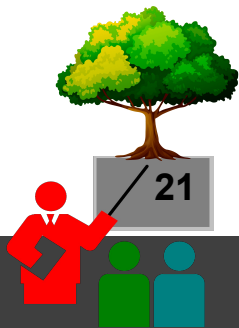
- Including Both Data Governance and AI Governance
  - Embedding Ethical AI Standards
    - Develop and enforce ethical AI guidelines across all projects.
    - Ensure transparency and fairness in AI model development.
    - Regularly audit AI systems for compliance with ethical standards.
    - Promote accountability and responsibility in AI decision-making.

The Data Scientist's Role in Governance Success
Integrating Governance into the Data Science Lifecycle
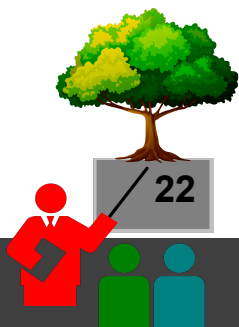
- Including Both Data Governance and AI Governance
  - Documenting Data Lineage and Model Decisions
    - Track and record data provenance throughout the data lifecycle.
    - Maintain comprehensive documentation of AI model development stages.
    - Ensure transparency in data transformations and model decisions.
    - Facilitate audits by providing detailed lineage and decision records.

The Data Scientist's Role in Governance Success
Integrating Governance into the Data Science Lifecycle
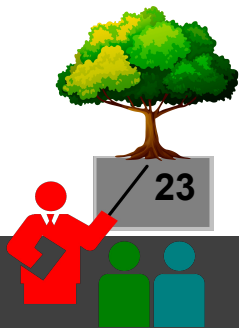
- Including Both Data Governance and AI Governance
  - Facilitating Cross-Departmental Governance Collaboration
    - Encourage joint data governance and AI governance initiatives.
    - Establish clear communication channels between data science and governance teams.
    - Align departmental goals to support unified governance strategies.
    - Share best practices and governance standards across departments.

The Data Scientist's Role in Governance Success
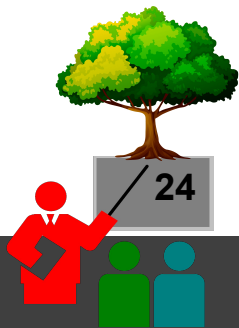## Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements

- Including Both Data Governance and AI Governance
  - **Identifying and Mitigating Risks:** Data scientists can use their expertise to identify potential compliance risks in data and AI models, ensuring adherence to regulatory standards.
  - **Implementing Privacy-Preserving Techniques:** Leverage data governance frameworks to apply privacy-preserving methods such as anonymization and differential privacy in AI models.
  - **Ensuring Ethical AI Practices:** Data scientists can integrate AI governance principles to uphold ethical standards, preventing biases and ensuring fairness in AI applications.
  - **Maintaining Comprehensive Documentation:** Document data sources, processes, and AI model decisions to facilitate audits and ensure transparency for regulatory compliance.
  - **Collaborating with Compliance Teams:** Data scientists can work closely with compliance officers to align AI and data governance strategies with regulatory requirements, enhancing overall governance effectiveness.

The Data Scientist's Role in Governance Success
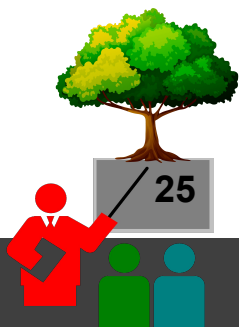## Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements

- Including Both Data Governance and AI Governance
  - Identifying and Mitigating Risks
    - Use data governance frameworks to proactively identify compliance risks.
    - Implement AI governance to detect and mitigate algorithmic biases.
    - Establish continuous monitoring processes to address emerging data risks.
    - Collaborate across departments to develop comprehensive risk mitigation strategies.

The Data Scientist's Role in Governance Success
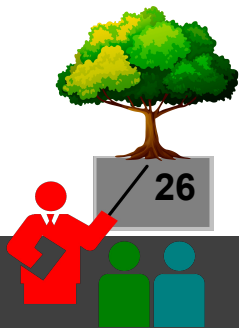## Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements

- Including Both Data Governance and AI Governance
  - Implementing Privacy-Preserving Techniques
    - Apply differential privacy methods to protect individual data points.
    - Utilize data anonymization techniques to maintain data privacy.
    - Incorporate secure multi-party computation for sensitive data analysis.
    - Implement federated learning to ensure data remains decentralized and private.

The Data Scientist's Role in Governance Success
Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements
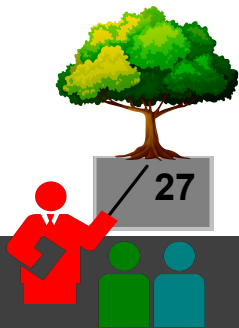
- Including Both Data Governance and AI Governance
  - Ensuring Ethical AI Practices
    - Establish guidelines for unbiased AI model development.
    - Ensure transparency in AI algorithms and decision-making processes.
    - Regularly audit AI models for ethical compliance and fairness.
    - Incorporate ethical considerations into data collection and usage policies.

The Data Scientist's Role in Governance Success
## Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements

- Including Both Data Governance and AI Governance
  - Maintaining Comprehensive Documentation
    - Document data sources and data transformation processes thoroughly.
    - Maintain detailed records of AI model development and training.
    - Ensure comprehensive documentation of data governance policies and procedures.
    - Regularly update documentation to reflect changes in data and AI governance practices.

The Data Scientist's Role in Governance Success
## Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements

- Including Both Data Governance and AI Governance
  - Collaborating with Compliance Teams
    - Align data usage with compliance regulations and standards.
    - Facilitate regular communication between data scientists and compliance officers.
    - Co-develop policies that integrate governance best practices.
    - Ensure AI models comply with regulatory frameworks and ethical guidelines.
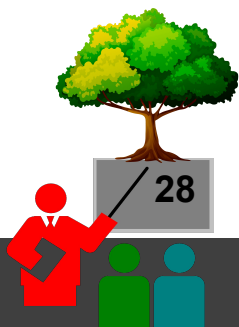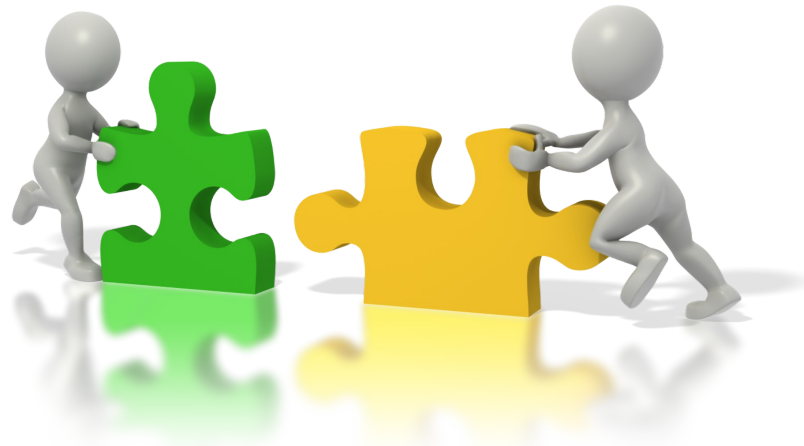
The Data Scientist's Role in Governance Success
Real-World Success Stories of Data Scientists in Governance Success

- Including Both Data Governance and AI Governance

  - **Enhancing Data Quality for AI Models:** A <u>financial</u> services company improved the accuracy of its predictive models by implementing strict data governance practices, ensuring high-quality data inputs.

  - **Ensuring Ethical AI in Healthcare:** A <u>healthcare</u> organization successfully reduced bias in AI diagnostics by employing data governance frameworks to oversee data collection and model training processes.

  - **Achieving Regulatory Compliance in Retail:** A <u>retail</u> business met stringent regulatory requirements by integrating AI governance with data governance, resulting in transparent and accountable data practices.

  - **Promoting Data Literacy and Stewardship:** A <u>tech company</u> empowered its data scientists as data stewards, enhancing data literacy and governance across the organization, leading to innovative and compliant AI solutions.
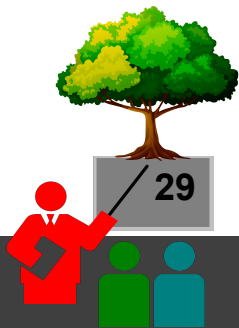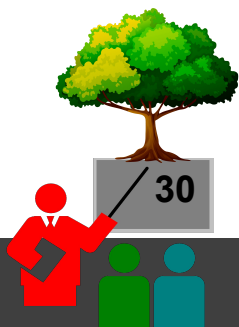
The Data Scientist's Role in Governance Success
Real-World Success Stories of Data Scientists in Governance Success

- Including Both Data Governance and AI Governance – Finance Company
  - Enhanced Data Quality for AI Models
    - Implemented rigorous data validation checks to ensure quality inputs.
    - Used data governance frameworks to standardize and clean data.
    - Integrated AI governance to continuously monitor model performance.
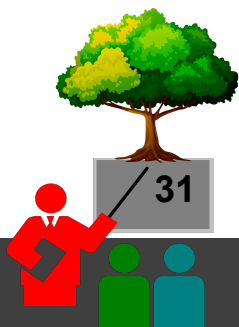   - Collaborated across teams to address data quality issues promptly.

The Data Scientist's Role in Governance Success
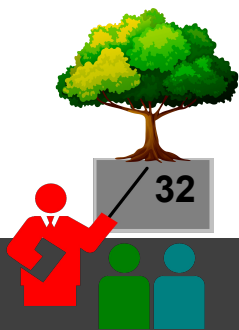## Real-World Success Stories of Data Scientists in Governance Success

- Including Both Data Governance and AI Governance – Healthcare Company
  - Ensured Ethical AI in Healthcare
    - Implemented AI governance to ensure patient data privacy.
    - Used data governance to eliminate bias in healthcare algorithms.
    - Maintained transparency in AI model decisions affecting patient care.
    - Fostered interdisciplinary collaboration to uphold ethical standards in AI.

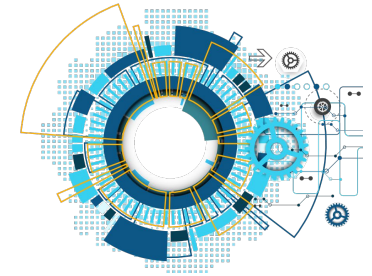# The Data Scientist's Role in Governance Success
Real-World Success Stories of Data Scientists in Governance Success

- Including Both Data Governance and AI Governance – Retail Company
  - Achieved Regulatory Compliance in Retail
    - Implemented data governance to ensure accurate reporting and audits.
    - Used AI governance to maintain compliance with consumer protection laws.
    - Leveraged data quality standards for adherence to industry regulations.
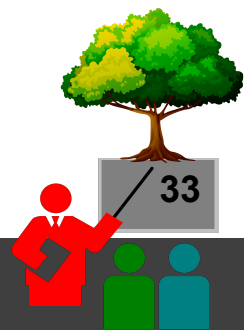   - Integrated AI solutions to enhance fraud detection and prevention.

The Data Scientist's Role in Governance Success
## Real-World Success Stories of Data Scientists in Governance Success

- Including Both Data Governance and AI Governance – Tech Company
  - Promoted Data Literacy and Stewardship
    - Fostered a culture of data literacy through targeted training programs.
    - Encouraged data stewardship by assigning clear data ownership roles.
    - Utilized AI governance to support ethical data usage and understanding.
    - Implemented data governance frameworks to enhance data accessibility and transparency.

The Data Scientist's Role in Governance Success
Summary

- In the webinar, I talked about …
    - Unearthing Governance Opportunities in Data Science Practices
    - Empowering Data Scientists as Data Stewards
    - Integrating Governance into the Data Science Lifecycle
    - Leveraging Data Scientist Expertise for Compliance and Regulatory Requirements
    - Real-World Success Stories of Data Scientists in Governance Success
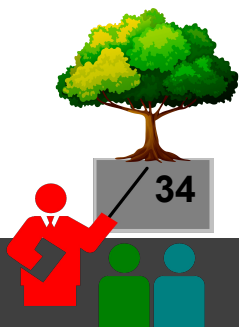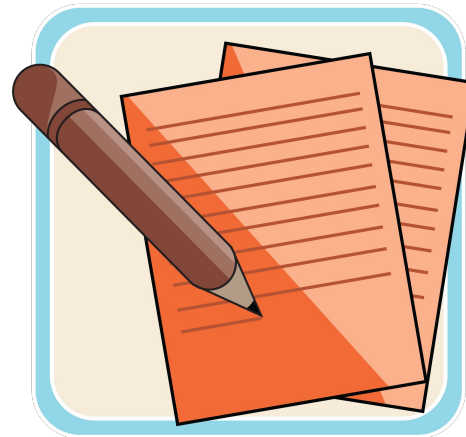
The Data Scientist's Role in Governance Success

Q & A - Contact Information

- Robert S. Seiner

  KIK Consulting & Educational Services – KIKconsulting.com

  rseiner@kikconsulting.com   @RSeiner

thank you!