Brought to you by:

Informatica

Responsible AI

for dummies

A Wiley Brand

Develop and deploy responsible AI systems

Navigate AI compliance efficiently

Drive decision-making with trusted data

Informatica Special Edition

Guy Hart-Davis

About Informatica

Informatica (NYSE: INFA) brings data and AI to life by empowering businesses to realize the transformative power of their most critical assets. When properly unlocked, data becomes a living and trusted resource that's democratized across your organization, turning chaos into clarity. Through the Informatica Intelligent Data Management Cloud™, companies are breathing life into their data to drive bigger ideas, create improved processes, and reduce costs. Powered by CLAIRE®, Informatica's AI engine, the cloud is dedicated to managing data of any type, pattern, complexity, or workload across any location — all on a single platform. Informatica. Where data and AI come to life.

For more information, visit www.informatica.com.

Connect with Informatica on social media:


linkedIn.com/company/Informatica


x.com/Informatica


facebook.com/InformaticaLLC

Responsible AI

Informatica Special Edition

by Guy Hart-Davis

for dummies®
A Wiley Brand

Responsible AI For Dummies®, Informatica Special Edition

Published by
John Wiley & Sons, Inc.
111 River St.
Hoboken, NJ 07030-5774
www.wiley.com

ISBN: 978-1-394-31531-4 (pbk); ISBN: 978-1-394-31532-1 (ebk); 978-1-394-31533-8 (ePub).
Some blank pages in the print version may not be included in the ePDF version.

Publisher's Acknowledgments

Some of the people who helped bring this book to market include the following:

Project Manager and Editor: Carrie Burchfield-Leighton

Sr. Managing Editor: Rev Mengle

Acquisitions Editor: Traci Martin

Sr. Client Account Manager: Matt Cox

Special Help from Informatica: Donal Dunne, Mark Kettles, Kayley Weinbaum, Carmen Coats, Melissa Whitaker, Shino Kizaki

Table of Contents

Introduction

Artificial intelligence (AI) has made such huge advances in the past few years that almost every aspect of computing now features an AI component. AI advances have happened so quickly as to outstrip regulatory guidance. Now, however, regulation is starting to catch up with the technology. For example, the European Union's Artificial Intelligence Act (EU AI Act), which is gradually coming into effect over the two years following August 2024, imposes stringent restrictions on any company that creates, offers, or uses AI systems anywhere in the EU.

Other countries and bodies are also regulating AI. China has implemented strict regulations on AI applications, particularly around data protection and cybersecurity. Canada has released guidelines for ethical and accountable AI development and use. Singapore has established a national AI strategy, while India has published a draft AI ethics code. The United Nations is developing international standards for AI guidance.

To comply with such regulation, companies must use AI ethically, respect the privacy of individuals, and avoid bias or unfairness in AI results. The company must also be able to track its data sources and usage, ensure its AI systems' decision-making is transparent, explainable, and accountable, and evaluate the societal impact of its AI usage.

These requirements add to the challenges that companies already face in integrating disparate data from the proliferation of data sources driven by generative AI, cleansing and standardizing that data, and keeping the data private and secure.

A movement called *responsible AI* can solve these issues.

About This Book

Responsible AI For Dummies, Informatica Special Edition, is your guide to what responsible AI is and how to implement it effectively in your company or organization. Responsible AI

- Gives you a conceptual framework for ensuring that AI systems are developed, deployed, and used in ways that are ethical, fair, and beneficial to society

- Includes measures to protect the privacy of individuals and to enable humans to understand how each AI system works and makes decisions
- Makes provisions for holding individuals and organizations accountable for the actions and decisions of their AI systems

Icons Used in This Book

This book displays icons in the margin to point you to two specific types of information:


REMEMBER

The Remember icon highlights information you should make sure you don't forget.


TIP

The Tip icon points out practical advice and key information you're likely to find helpful.

Beyond the Book

This book gives you a firm grounding in responsible AI, but this topic is huge. If you're eager to know more, follow these links:

- **www.informatica.com/blogs.html:** Visit this blog from data-management experts at Informatica.
- **www.informatica.com/ResponsibleAIBlog:** Accelerate responsible AI outcomes by harnessing automated inferred data lineage.
- **www.informatica.com/ResponsibleAI:** Lay the data foundation for responsible AI governance.
- **www.informatica.com/ResponsibleAIWebinar:** Discover how to harness responsible AI data governance for better business results.
- **partnershiponai.org:** Find responsible AI research, best practices, and policy analysis.

IN THIS CHAPTER

- Understanding AI's rise
- Highlighting why responsible AI is needed
- Introducing solutions to implement responsible AI

Chapter 1
Grasping the Need for Responsible AI

Artificial intelligence (AI) is the hottest topic in computing right now. Consumers are taking enthusiastically to chatbots, and businesses are rushing to add AI capabilities to their products and services. Government bodies — from the United States (U.S.) Department of Defense to the Environmental Protection Agency (EPA), the European Environment Agency (EEA) and the United Kingdom's National Health Service (NHS) — are using AI systems for a wide variety of tasks.

These early days of AI adoption by consumers and businesses have had an anything-goes, Wild-West atmosphere, with change and progress happening at breakneck speed and governments and regulators scrambling to catch up. Some uses of AI have been biased, unethical, unfair, or simply dangerous. But now a framework called *responsible AI* is gradually bringing order to the chaos, ensuring that AI technologies are used in ethical, transparent, and safe ways.

In this chapter, you discover what responsible AI is, why it's needed, and how companies are implementing it.

Tracking AI's Rapid Rise and Future Trajectory

In July and August 1956, the Dartmouth Summer Research Project on Artificial Intelligence held a workshop to explore the possibility of creating machines able to think, reason, and learn. This workshop, held at Dartmouth College in New Hampshire, USA, launched the field of AI studies. But despite high-profile successes, such as IBM's Deep Blue computer beating world champion Garry Kasparov at chess in 1997, overall progress was slow enough for the public to keep seeing AI as science fiction rather than reality.

That all changed when the research organization OpenAI unleashed its ChatGPT chatbot on a largely unsuspecting public in November 2022. ChatGPT is a large language model (LLM). An LLM is an AI system trained on huge amounts of text data that enable it to understand human language with natural phrasing and to respond to queries.

Usage exploded. Within two months, ChatGPT had 100 million users. In early 2023, Google released its Bard LLM, and Meta introduced its LLaMA LLM; in late 2023, Anthropic introduced Claude, another LLM. Many other LLMs followed, with hundreds — perhaps even thousands — of LLMs available to the public.

But huge numbers of LLMs are just the tip of the AI iceberg. Businesses have now integrated AI systems into their operations to sharpen their decision-making, boost their efficiency, and outmaneuver their competitors. Business tasks enhanced by AI include the following:

- **Customer service:** Chatbots handle straightforward customer queries, passing on complex queries to humans.
- **Sales and marketing:** AI tools identify potential sales leads and mine customer behavior and preferences to develop personalized marketing campaigns.
- **Human resources:** AI tools perform functions from screening resumes to preparing performance evaluations.
- **Data analysis:** AI tools analyze large datasets to extract business insights.

What lies ahead for AI in business? The future looks bright, if not yet fully understood, because greater usage of AI is all but certain. The details are hazy, but developments such as the following seem promising:

- **AI-powered healthcare diagnostics:** Real-time monitoring of patients and personalized treatment plans could improve healthcare outcomes and reduce costs.
- **AI personal assistants:** Smart and flexible personal-assistant systems could streamline task management and scheduling.
- **Autonomous transportation:** AI systems may enable the long-held dream of self-driving vehicles or drones.
- **Smart manufacturing:** AI tools can optimize manufacturing processes and reduce downtime by scheduling predictive maintenance.

Understanding the Need for Responsible AI

Companies looking to add AI to their businesses face potential problems ranging from a shortage of AI specialists and skilled AI staff to the cost and infrastructure needed to implement AI systems, train them, and keep them running well.

One of the biggest AI implementation problems for businesses at present is ethical and regulatory concerns. Now that the regulatory regime has caught up with technology enough to dispel the virtual Wild-West atmosphere, businesses need to do the following:

- Use AI ethically.
- Comply with all relevant privacy regulations.
- Avoid bias and unfairness in AI results.
- Implement AI explainability, enabling people to understand and explain how AI systems work.
- Ensure AI fluency in staff, boosting employees' AI skills to drive AI adoption and usage.

TIP

The best way to meet these requirements is to establish a sound framework of data governance. Check out Chapter 3 for more information.

A second major problem is ensuring that the AI system has enough high-quality data available. The data needs to be accurate, consistent, complete, and structured suitably for analysis. Preferably, the dataset should be refreshed regularly with new data. For regulatory compliance, the data lineage should be traceable, so analysts can see where data has come from and how it's being used.

REMEMBER

Responsible AI provides a conceptual framework intended to ensure that AI systems are developed, deployed, and used in ways that are ethical, fair, and beneficial to society. The goals of responsible AI include the following:

- Protecting the privacy of individuals
- Making the workings and decisions of AI systems easily understandable by humans
- Holding individuals and companies accountable for the actions and decisions of their AI systems

Moving toward Effective Solutions for Responsible AI

By implementing a data governance strategy, a company puts in place a framework that outlines policies, guidelines, and processes for ensuring data quality, data security, and compliance with applicable regulations. This strategy works together with technological solutions, such as Intelligent Data Management Cloud (IDMC) from data experts Informatica, to drive adoption and automate data governance tasks.

For example, AI-powered tools can identify and address data quality issues, classify data automatically, and enforce data privacy regulations. By integrating technology with a sound data governance framework, a company can streamline its data management processes, reduce risks, and improve its decision-making.

IN THIS CHAPTER

- Understanding the business challenges of responsible AI
- Recognizing technical hurdles to adopting responsible AI practices

Chapter 2
Identifying the Challenges in Responsible AI

Before you tackle responsible AI, you want to identify the challenges you may face in implementing it. Broadly speaking, these challenges fall into two main categories: business challenges and technical challenges.

In this chapter, you first examine the business challenges, which range from privacy issues to collaboration and risk. After that, you look at the technical challenges that responsible AI can pose. These challenges include data governance and the AI fluency — sufficient understanding of AI coupled with the skill to use AI tools effectively — needed for implementation.

Identifying Business Challenges in Implementing AI Responsibly

When implementing responsible AI, you may face several business challenges:

- **Developing a data governance strategy:** To implement AI responsibly, a business needs to develop a data governance

strategy that ensures data quality, consistency, security, processes and guidelines, and compliance with regulations. See Chapter 4.

- **Getting stakeholder buy-in:** A business needs to get buy-in from stakeholders — from executives to employees, and from partners and customers to the broader community — to align its AI implementation with its business goals, to address concerns, and to encourage adoption and usage.
- **Upskilling the workforce and changing its culture:** A business must boost its employees' skills to the level of AI fluency needed to use AI effectively. Moving to a data-driven culture may involve overcoming resistance to change as well as meeting substantial training costs.
- **Grasping privacy concerns:** Many AI systems need huge amounts of data to function well. AI implementers must balance the need to collect and analyze this data with privacy concerns, especially when some of that data is sensitive, such as personally identifiable information (PII). Implementers must keep the data secure against unauthorized or improper usage. To maintain user trust, a business must not only handle data responsibly but also communicate clearly how it collects, handles, and uses data.
- **Complying with regulations:** When designing and deploying AI systems, businesses must comply with regulations and standards for data privacy. Chapter 3 gives you more info on key regulations and standards.
- **Avoiding bias and flawed results:** Bias can cause problems in AI systems, so managing it is essential. Bias, or even the appearance of bias, can cause dissatisfaction among customers and inflict reputational damage that puts the company at a competitive disadvantage. Flawed results — outputs that are incorrect or misleading — can cause similar problems.

  
  REMEMBER

  In AI systems, bias can result from several factors. Data that isn't sufficiently diverse and representative, complete and accurate, and consistent can produce data bias, while poorly designed algorithms can cause algorithmic bias. AI developers and trainers can introduce human bias, inadvertently or otherwise. Inadequate data, unsuitable algorithms, and inappropriate human input can all produce flawed results.
- **Ensuring fairness:** Fairness is crucial for responsible AI. Lack of fairness raises legal and compliance risks and can harm a company's reputation and destroy customer trust. A

system's operational effectiveness may also suffer from a lack of fairness — for example, if a biased hiring system discriminates against strong candidates or selects suboptimal business partners.

- **Providing transparency and explainability:** In responsible AI, *transparency* means communicating clearly how an AI system works, what it's capable of, what data it uses, and how it makes decisions. *Explainability* means enabling people to understand the reasoning and justification an AI system uses to reach its decisions. Many jurisdictions have laws and regulations that require transparency and explainability in AI systems.


REMEMBER

  Even where they're not legally required, transparency and explainability can benefit companies by building trust with consumers and stakeholders that can lead to competitive advantage. By providing a means to hold companies responsible for their AI systems' decisions and actions, transparency and explainability help companies address issues or problems that arise with those systems.

- **Collaborating effectively:** Collaboration can present a challenge for responsible AI. A company needs to align various groups of stakeholders on a common goal, share data and allocate resources effectively, and build trust and transparency while complying with all relevant regulations. The company may also need to integrate disparate AI systems and protect intellectual property.
- **Reducing risk:** Deploying and using AI systems can result in adverse effects or negative outcomes including legal and regulatory risks from noncompliance, reputational risks from biased or flawed results, or operational risks from poor decision-making or system failures. Companies implementing responsible AI need to reduce such risks as far as possible, minimize the financial risks of creating AI systems, and secure their data against loss or misuse.

Recognizing the Technical Challenges of Responsible AI

Companies also face technical challenges in implementing responsible AI. One challenge is adopting the right data governance application. Which solution offers the right capabilities and

features for the customer? How does it integrate with IT systems, how does it scale, what is compliance like, and so on? You also need to consider security, user experience, complexity of data landscape, support, change management and implementation, and much more. To help ensure that your AI systems are developed and used ethically, you need to implement a data governance framework, such as the National Institute of Standards and Technology (NIST), complemented by a data management tool. This governance means setting up policies, processes, and practices for managing data assets.

Data governance policies typically focus on the following five aspects of data:

- **Quality:** Ensuring that data used to train your AI models is accurate, relevant, complete, and consistent
- **Privacy:** Protecting the privacy of the individuals whose data your AI systems use
- **Security:** Securing data against unauthorized access or usage, disclosure, modification, destruction, or loss
- **Lineage:** Tracking the data's sources, its movement, and its usage through its life cycle
- **Retention and deletion:** Retaining data to comply with legal and ethical requirements and deleting data when its retention is no longer required or needed

Another challenge is building AI fluency in your workforce. To get the most out of AI, your workforce needs to understand the essentials of AI and to be able to work effectively with it. Such abilities are often referred to using the umbrella term *AI fluency*.

AI fluency is a person's ability to understand AI systems well enough to work effectively with them and make informed decisions about them. The degree of AI fluency required depends on the specific position. Some staff members — such as AI developers, deployers, and ethicists — will need expert-level skills in their fields, while others will need more general comprehension and skills.

REMEMBER

To build AI fluency in your workforce, you may need to provide reskilling training for some workers. This reskilling can take various forms, such as on-the-job training sessions, online courses, workshops, or bootcamps.

IN THIS CHAPTER

- Learning about current regulations
- Understanding why data management is essential

Chapter 3
Driving Successful AI Adoption with Data Management

Data management involves collecting, storing, processing, protecting, and using data for business purposes. Effective data management is critical for successfully adopting artificial intelligence (AI) systems for two key reasons: Your business needs to comply fully with the laws and regulations covering AI usage, and your business needs consistent, high-quality data that enables your AI systems to reliably produce the results you need.

Understanding Current AI Regulations and Their Impact

AI is covered by a patchwork of regulations around the world, and complying with all applicable regulations can be a complex business challenge. In this section, you find out about two major pieces of AI regulation.

Examining the EU AI Act

The European Union Artificial Intelligence Act (EU AI Act) became law in August 2024, with its provisions gradually taking effect over the next two years, after which it will be fully applicable. This act applies to any AI system developed, offered, or used in the EU.

It aims to protect the safety and fundamental rights of citizens and businesses while promoting lawful and ethical AI usage, and classifies AI systems into four risk categories:

- **Unacceptable risk:** The Act prohibits or tightly restricts AI systems that pose a high risk to people's safety or rights. Examples of unacceptable-risk AI systems include critical infrastructure or social scoring by governments.
- **High risk:** The Act imposes strict requirements on AI systems that pose significant risk to people's safety or rights. High-risk AI systems include ones that assess whether someone is able to receive a certain medical treatment, to get a certain job or loan, or to to buy an apartment. Requirements include accountability, transparency, and human oversight.
- **Limited risk:** The Act imposes general requirements, such as accountability and transparency, on AI systems that pose a limited risk. Limited-risk AI systems in this category include customer service and marketing.
- **Minimal risk:** The Act imposes no specific requirements on minimal-risk AI systems. Customer service and marketing are also examples of minimal-risk AI systems.

The Act includes provisions intended to make AI systems transparent — understandable and accountable; to subject them to human oversight; and to require data governance, ensuring the lawful and ethical use of data. The Act establishes a framework for enforcing its provisions, including fines for non-compliance and banning unacceptable AI systems. The Act's emphasis on keeping data use lawful and ethical highlights the importance of data management and data governance.

REMEMBER

Data management is the handling of data. *Data governance* means establishing standards, policies, and procedures for managing data. Data governance makes sure your company uses data legally and ethically to align with your business objectives.

Meeting the U.S. Executive Order on Artificial Intelligence

Issued by President Joe Biden on October 30, 2023, the Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence contains wide-ranging measures concerning AI. This Order is called the U.S. Executive Order on Artificial Intelligence for short, and it focuses on three key themes:

- **Regulate AI:** Reduce the risk of AI being abused, and encourage responsible use of AI.
- **Protect Americans' privacy and rights:** Protect privacy by developing privacy-preserving techniques; advance equity and civil rights, and ensure fairness; and support workers affected by AI deployment and usage.
- **Boost innovation:** Promote AI innovation and competition, advance American leadership abroad, and encourage responsible government deployment of AI.

Responsible use of AI and protecting privacy are major themes of the Executive Order. Complying with the Order requires effective data management and data governance.

Clarifying the Need for Data Governance

The EU AI Act and the U.S. Executive Order take different approaches to data governance, with the former being more prescriptive and regulatory — including penalties — and the Executive Order being principle based and flexible. But both highlight the importance of managing data tightly to protect privacy and prevent data misuse.

REMEMBER

Effective data governance and clear communication can build customer trust in a few ways:

- If your company shows it's committed to data privacy and security, customers will be more willing to trust you with their personal information.
- You can use that personal information to give customers personalized experiences, which can increase customer satisfaction and loyalty.

- Accurate and reliable data helps you deliver better outcomes to customers.

To help manage data, create a data strategy and policy:

- **Data strategy:** This high-level plan describes your company's goals and objectives for data. The strategy normally includes a data vision that explains the company's long-term objectives for data and its usage, data goals (for quality, accessibility, or decision-making), and data initiatives (plans for reaching the data goals).
- **Data policy:** These rules and guidelines explain how your company actually manages the data — how you collect the data, store and process the data, and use the data. A data policy normally covers data ownership (who is responsible for the data), data access (who can access the data, and when), data security (preventing unauthorized access, use, modification, or deletion), data quality (standards and procedures for ensuring data is accurate, complete, and consistent), and data retention (how long you retain data).

TIP

Make sure your data strategy aligns with your policy objectives. Follow these steps:

1. **Identifying your policy objectives for data, such as ensuring policy and fairness and implementing security and accountability.**
2. **Assess how well your current data practices meet the policy objectives.**
3. **Derive the goals, actions, and metrics needed to create a data strategy that fully meets the objectives.**
4. **Create data governance policies that cover data collection, storage and processing, and use.**
5. **Boost AI fluency by fostering a data-driven culture and improving data literacy.**

   These policies support your policy objectives. With the policies in place, evaluate how effectively your data practices satisfy your policy objectives, and adjust the practices as needed.

IN THIS CHAPTER

- Integrating data for your company's AI systems
- Building a single source of truth
- Governing AI models
- Meeting the Informatica IDMC platform

Chapter 4
Implementing Data Management and AI Model Governance

In this chapter, you examine the challenges involved in integrating data into your company's systems. This data will likely come from disparate sources, such as various business functions and external sources, and may include both structured and unstructured data. You explore how to manage data to create a single source of truth (SSOT) to provide accurate, up-to-date, and consistent data to your AI systems. You also see how to create a robust plan for AI model governance, and you meet a one-stop solution for data management and data governance.

Integrating Data for Your AI Systems

Before your AI systems can use your company's data, you may need to integrate the data — collecting the data, cleaning it, and transforming it into a format suitable for manipulation by AI algorithms. This section explains three challenges you may face when integrating data.

Managing the flow of data

When integrating data, you have to manage the flow of data from its intake point through to the point at which it's actually used. If the data comes from various sources, as is often the case, you may need to standardize the data and clean it up to make it useful.


TIP

If you can target automation for this process of ingesting data, managing the flow of data runs more efficiently while you also maintain data integrity and quality.

Ensuring data scope and variety

AI systems thrive on diverse data — data that's broad in scope and varied in type. Getting such data can be a challenge. If you're drawing together data from disparate sources, you may need to use sophisticated tools and techniques, especially if some data sources are unstructured or semi-structured rather than structured.


TIP

To keep your AI system current, update the dataset with new data from its existing sources. You may also need to identify new data sources that you can add to the dataset.

Reducing or eliminating bias

Human error in designing and training AI systems can produce modeling and systemic biases that lead to inaccurate or unfair outcomes. Another source of bias is data that contains imbalances and anomalies that can skew an AI system's predictions.


TIP

To avoid these problems and comply with applicable regulations (such as the EU AI Act), observe your data for such imbalances and anomalies, and use techniques such as resampling or reweighting to eliminate them.

Managing Data to Create a Single Source of Truth


REMEMBER

An SSOT is an authoritative and centralized data repository, such as a database, that contains your company's most up-to-date, accurate, and verified data. The SSOT acts as a definitive reference point for your company's data usage, giving consistency and reliability for your company's data activities.

Creating an SSOT is a best practice for any company that handles large datasets originating from hundreds or thousands of heterogeneous sources, but when your company needs to comply with governance requirements such as those in the EU AI Act, an SSOT becomes a necessity. The SSOT enables analysts to follow the lineage of data and verify its quality, which helps AI systems produce accurate and reliable insights.

When creating an SSOT, follow these key steps:

1. **Identify your data sources.**

   List the data sources you will include in the SSOT, such as databases, spreadsheets, applications, and external data feeds. Identify the data fields you'll use from the sources.

2. **Create the SSOT database or data repository and establish master data management (MDM) practices.**

   Populate the database or data repository with the data fields. Identify and prioritize key data elements in the SSOT. Assign data stewards responsible for managing the quality, access, and governance of specific data elements.

3. **Integrate the data.**

   Extract the data from the sources, transform it into a consistent format, and load it into the SSOT database. These moves are sometimes called extract, transform, load (ETL) processes.

4. **Cleanse, validate, and standardize the data.**

   Cleanse the data to remove errors, conflicts, and duplicates; validate the data to verify its quality; and standardize it with your chosen formats.

5. **Implement data discovery and classification.**

   Analyze your data and its quality. Create classification categories or labels and apply them to the data.

6. **Create a data catalog.**

   The catalog is a centralized repository containing metadata about the documents in the SSOT. The catalog consolidates information about all data sources, enabling you to manage and access data from a single point and helping data discovery, data lineage, metadata management, and access control.

7. **Implement data observability.**

   Check for quality assurance by monitoring data continuously against predefined quality metrics such as accuracy, completeness, and consistency. Implement advanced anomaly detection to identify and address unexpected problems in your data, establish what caused the problem, and assess its impact.

Establishing a Robust AI Model Governance Plan

To support the security and ethical operations of your AI systems, and to help them produce reliable outcomes, create a comprehensive AI model governance plan. This section explains considerations to keep in mind when drawing up your governance plan.

Ensuring privacy and compliance

If your AI systems process sensitive data, as most AI systems do, you must ensure data privacy and comply with the applicable laws and regulations, such as the EU's General Data Protection Regulation (GDPR) law. Include regular audits and monitoring in your governance plan to ensure compliance. Stay up to date as laws and regulations change, and alter your plan to accommodate the changes.

Reducing risk

Given that any AI system runs risks such as exhibiting bias, misinterpreting data, or simply malfunctioning, your governance plan should include strategies for monitoring these risks, identifying when they occur, and mitigating their effects. In your plan, include regular testing and validation to confirm that the AI system is producing safe and accurate results and to reduce the chances of harm if something goes wrong.

Controlling access and sharing

To keep your AI systems secure and reduce the risk of tampering or misuse, limit access so that only authorized staff members and authorized systems can interact with the systems. Establish

robust authentication protocols to prevent unauthorized access, and maintain and regularly review access logs to ensure that only authorized personnel utilize the systems.

Permitting only approved usage

As well as limiting access to your AI systems, you should also specify the purposes for which they can be used. Make a list of approved use cases aligned with your company's goals and ethical standards for the AI systems. Restricting the AI systems to specific purposes helps ensure they produce accurate and reliable outcomes instead of throwing out unpredictable or erroneous results.

Introducing Informatica Intelligent Data Management Cloud

Various companies offer solutions for managing data and implementing governance of AI models. The Informatica Intelligent Data Management Cloud (IDMC) is an enterprise AI-powered cloud data management solution. IDMC offers a comprehensive solution for data management and AI model governance, enabling you to seamlessly integrate responsible AI practices into your business operations.

IDMC provides the following services:

- **Data catalog:** By cataloging your data, this service enables you to identify AI systems, data sources, and data types that are subject to regulation by the EU AI Act or other legislation, while providing data lineage insights for comprehensive compliance reporting.
- **Data integration and engineering:** This service lets you ingest, integrate, and cleanse your data. You can develop and implement data pipelines for AI with trusted data at scale.
- **API and app integration:** This service integrates apps, data sources, and AI systems, facilitating the flow of data between different systems. API and app integration lets you move legacy apps to the cloud, connecting them to your data.

- **Data quality and observability:** Ensure your AI systems receive timely, trusted, and relevant data. The service monitors data quality and system status for fitness and identifies issues and anomalies that can affect the reliability of AI models.
- **MDM and 360 applications:** Identify AI policy-compliant data, and make it available for use in AI projects. This service is particularly helpful for using sensitive data, such as data about customers or employees.
- **Governance and privacy:** IDMC provides data intelligence and reliable insights while complying with data governance and AI governance requirements. *Data intelligence* helps you better understand the fundamentals of your data by deriving its meaningful attributes — the who, what, where, when and how.
- **Data marketplace:** This service exposes data for sharing swiftly and safely while complying with policies for the ethical and appropriate use of AI.

REMEMBER

IDMC provides a one-stop solution capable of handling every step of your data's journey from ingestion to policy-compliant usage in your responsible AI systems.

IN THIS CHAPTER

- Building consumer and stakeholder confidence
- Identifying ten principles of responsible AI
- Understanding Informatica's role in driving responsible AI
- Highlighting two responsible AI pioneers

Chapter 5
Building Trust through Responsible AI

In this chapter, you explore how your company can build consumer and stakeholder confidence by practicing responsible AI data governance. You identify ten key principles of responsible AI and learn how Informatica, a leader in data integration solutions, has helped drive responsible AI practices. Finally, you meet two major international companies that have implemented responsible AI successfully and derived significant benefits from doing so.

Building Customer and Stakeholder Confidence

By adopting responsible AI practices, your company can build the confidence of both customers and stakeholders by

- **Demonstrating transparency and openness:** Communicate clearly with customers and other stakeholders what your AI systems can do, what they can't do, and how they make decisions. Explain the derivation and impact of AI-generated outputs.

- **Adhering to ethical principles:** Develop and use AI systems fairly, avoiding bias so as to give equitable outcomes.
- **Protecting privacy:** Comply with laws and regulations on data privacy. Get informed consent from individuals before collecting and using their data. Handle all customer data securely and responsibly.
- **Implementing accountability and oversight:** Take responsibility for your AI systems' output and decisions. Set up human oversight for your systems, including regular reviews.
- **Committing to continuous improvement:** Monitor your AI systems to ensure they are working effectively and delivering results that are fair and free of bias. Keep up with ethical standards as they evolve, and update your AI systems to meet those standards. Collect feedback and use it to improve your systems.

Grasping the Principles of Responsible AI

Responsible AI is a developing area, but it's grounded on the following principles:

- **Fairness:** Ensure equitable outcomes for all users.
- **Bias mitigation:** Detect and reduce (ideally eliminate) biases that may cause unfair treatments of groups or individuals.
- **Transparency:** Make clear the processes used to design and deploy AI; disclose the AI's decision processes and its capabilities.
- **Explainability:** Enable the AI system to explain, in generally understandable language, how it makes decisions.
- **Privacy:** Keep personally identifiable information (PII) secure and use that information only in ways agreed with its providers.
- **Data protection:** Handle individuals' data legally and securely, complying with privacy laws and regulations. Protect the data from unauthorized access and usage. Retain data as required. Delete data securely.
- **Accountability:** Even if an AI system has autonomous capabilities, the system's developers, deployers, and

operators are responsible for the consequences of its operation.

- **Governance:** Companies driving AI development and deployment must establish frameworks, policies, and standards to comply with ethical and legal standards.
- **Safety:** Create AI systems that can handle data reliably without errors that might cause harmful decisions.
- **Security:** Protect AI systems from attacks and build them to withstand tampering and manipulation.
- **Societal impact:** Understand and manage the broad effects AI technology has on society, aiming to maintain ethical standards and provide societal benefits while avoiding having negative effects on social structures.

Driving the Responsible AI Movement with Informatica

Informatica's vision for data management is one of unification and empowerment. It encompasses a foundation that enables self-serve access to valid, trustworthy data, with AI-infused tools that amplify productivity and enrich user experiences.

Informatica Intelligent Data Management Cloud (IDMC) plays a central role in de-risking AI initiatives and enhancing compliance with regulations such as the EU AI Act. The comprehensive solution offers data governance with privacy controls, data quality improvement, and AI-powered data cataloging to ensure the transparency, reliability, and integrity of data. Through automated data management tasks and seamless data integration, IDMC increases operational efficiency and creates a single source of truth.

Complying with regulations

Informatica's solutions include robust compliance measures that help companies comply with a fast-changing regulatory landscape that includes the EU AI Act, the General Data Protection Regulation (GDPR), the U.S. Executive Order on Artificial Intelligence, and the California Consumer Privacy Act (CCPA).

Building ethical frameworks

Ethical frameworks developed by Informatica help companies responsibly develop, deploy, and use AI. By using these frameworks, companies can make sure they use AI in a way that benefits society and respects human rights as well as delivers business insights.

Meeting Two Companies with Responsible AI Success Stories

In this section, I show you two companies that have implemented data governance powered by responsible AI with the help of Informatica IDMC.

Major pharmaceuticals

Based in Europe but operating around the world, a major pharmaceutical company develops, manufactures, and markets both prescription and over-the-counter pharmacological products. The company aims to use AI to develop and roll out drugs more swiftly and to optimize its product line.

The company faced a wide variety of challenges, such as working with more than 150 assorted data models, having only limited visibility into the sources and usage of its data, and using inconsistent data management processes across its multiple departments. To resolve these challenges, the company worked with Informatica. By leveraging the capabilities of IDMC, the company cleansed and standardized its data, enabling it to deliver trusted, high-quality, and compliant data to its AI applications.

REMEMBER

Standardized, governed, and automated data flows make data available in near real-time for analysis and advanced analytics, generating faster and more accurate insights and enabling the company to respond quickly to changing AI regulations while reducing technology costs.

A leading telco

A leading European telecommunications provider turned to Informatica to implement a federated data governance model using the FAIR principles (see the nearby sidebar). The telco faced a challenging data environment complicated by managing more than 1,000 enterprise applications, complying with multiple governance standards, and satisfying growing contractual and regulatory demands.

By partnering with Informatica, the telco has streamlined and standardized its data assets, creating high-quality data products that are both secure and tightly governed, and making these data products accessible to partners and customers around the world.

DECODING THE FAIR PRINCIPLES

Findable, Accessible, Interoperable, and Reusable (FAIR) are four desirable qualities in research data:

- **Findability** involves using metadata and identifiers to ensure data is easily discoverable by humans as well as by machines.
- **Accessibility** specifies data be accessible and retrievable via a standardized protocol rather than a proprietary protocol.
- **Interoperability** uses standardized formats and protocols to make data compatible with other datasets.
- **Reusability** requires documentation, licensing, context, and metadata to enable others to reuse data easily.

IN THIS CHAPTER

- Developing an AI data governance strategy
- Outlining policies, compliance requirements, and ethical guidelines
- Creating an implementation plan
- Building a process of review and improvement

Chapter 6
Ten Key Actions to Implementing Responsible AI

Implementing responsible artificial intelligence (AI) in your company is not only a major undertaking but also one that should become an ongoing, iterative process rather than a once-and-done activity.

TIP

To streamline the responsible AI implementation process, follow these actions:

- **Develop an AI data governance strategy:** Audit your company's current policies and identify gaps in its data practices. Appoint leadership, allocate roles, and establish a framework for AI governance.
- **Outline policies, compliance, and legal requirements:** Develop and enforce AI policies that align with business regulations, legal standards, and ethical guidelines. Focus these policies on fairness, accountability, transparency, quality, security, and privacy.

- **Create an implementation plan:** Engage stakeholders through structured awareness programs, such as workshops, training sessions, and webinars or online courses. Plan comprehensive training to ensure AI fluency in staff members. Embed sustainable processes and best practices throughout the company.
- **Conduct regular reviews and monitoring:** Implement systems for ongoing audits and monitoring AI model performance. Identify and implement continuous improvements in response to market changes and emerging technologies.
- **Establish an AI ethics committee:** Set up an internal ethics committee including experts on technology, legal matters, and ethics. Charge the committee with assessing your company's AI projects and make sure they're aligned with ethical and social responsibility standards and your company's goals.
- **Incorporate bias detection and mitigation:** Apply advanced tools and techniques, such as counterfactual testing and adversarial training, to detect and mitigate bias in your AI systems. Aim to ensure that your AI systems remain fair and nondiscriminatory no matter which demographics they're dealing with.
- **Develop a risk management framework:** Create a risk management strategy that identifies potential ethical, legal, or operational risks your AI systems may pose. Develop procedures for addressing these risks.
- **Build transparency and explainability into AI decision-making:** Design your AI models with transparency and explainability in mind to ensure that decisions your AI systems make can be clearly understood, validated, and traced.
- **Preserve privacy through AI techniques:** Protect sensitive information by implementing privacy-preserving techniques such as differential privacy, data anonymization, and federated learning.
- **Foster a culture of ethical AI:** Cultivate an organizational culture that gives priority to ethical AI practices. Encourage your company's employees to take an active interest in AI and its usage; raise any concerns or questions they have about AI and its implementation; and make suggestions for improving AI models and their usage.

Making AI Responsible and Work Better with Well-Managed Data

Artificial Intelligence (AI) has moved beyond being a novel concept to becoming a widely adopted element of modern business strategies. As AI continues to evolve, mastering the use of this technology responsibly becomes crucial, harnessing AI in a mindful, respectful and principled manner.

Responsible AI is about fostering ethical, transparent and accountable systems. Informatica is at the forefront, providing the essential tools needed through effective, AI-powered data management to enable trusted data and responsible AI outcomes:

- **Ethical Decision Making and Transparency:** Our tools help ensure data accuracy and reduce biases, making AI decisions more understandable and trustworthy.
- **Accountability and Compliance:** Robust data governance frameworks and audit capabilities help ensure AI systems meet regulatory standards and maintain legal compliance.
- **Safety and Security:** Our data privacy management tools implement stringent security measures to safeguard AI systems and sensitive data.
- **Public Trust and Ethical Commitment:** By promoting ethical, transparent and secure data handling practices, Informatica fosters public trust and demonstrates a strong commitment to ethical AI.

The Informatica Intelligent Data Management Cloud™ (IDMC) helps companies create a trusted data foundation by ensuring the availability and accessibility of high-quality, safe and protected data. This foundation mitigates privacy, security and compliance risks, paving the way for responsible AI applications.

Adopting Informatica's scalable solutions not only reduces risks, but also streamlines your journey towards responsible AI, ensuring better business outcomes and a promising technological future.

Explore Responsible AI Governance: Learn how to leverage these benefits in our latest eBook, *"Chart the Course for Responsible AI Governance:"* **informatica.com/ResponsibleAI**. For more information on how Informatica can help you bring your data to life, visit our website at **informatica.com**.

Where data & AI come to LIFE™

Master responsible AI

Explore how to implement and responsibly govern AI within your business by using expert insights and practical strategies. Learn how adopting a unified approach to data governance, quality, and privacy fosters ethical, transparent, fair, and effective AI use. This book provides leaders with essential knowledge to build trust and optimize business outcomes through responsible AI practices. Secure your technological future by learning how to turn responsible AI from a concept into reality for your enterprise.

Inside...

- AI's evolution and business impact
- Identifying hurdles in responsible AI
- Navigate AI regulation and data
- Elements of an all-in-one solution
- Building trust with ethical practices
- Practical examples of responsible AI
- Ten steps to responsible AI

Informatica

Guy Hart-Davis is the author of several computer books, including *Killer ChatGPT Prompts: Harness the Power of AI for Success and Profit.* Releasing in late 2024 will be *iPhone For Dummies, 2025 Edition; macOS Sequoia For Dummies;* and *Teach Yourself VISUALLY iPhone 16.*

Go to Dummies.com™
for videos, step-by-step photos, how-to articles, or to shop!

ISBN: 978-1-394-31531-4
Not For Resale

for dummies®
A Wiley Brand

WILEY END USER LICENSE AGREEMENT

Go to www.wiley.com/go/eula to access Wiley’s ebook EULA.