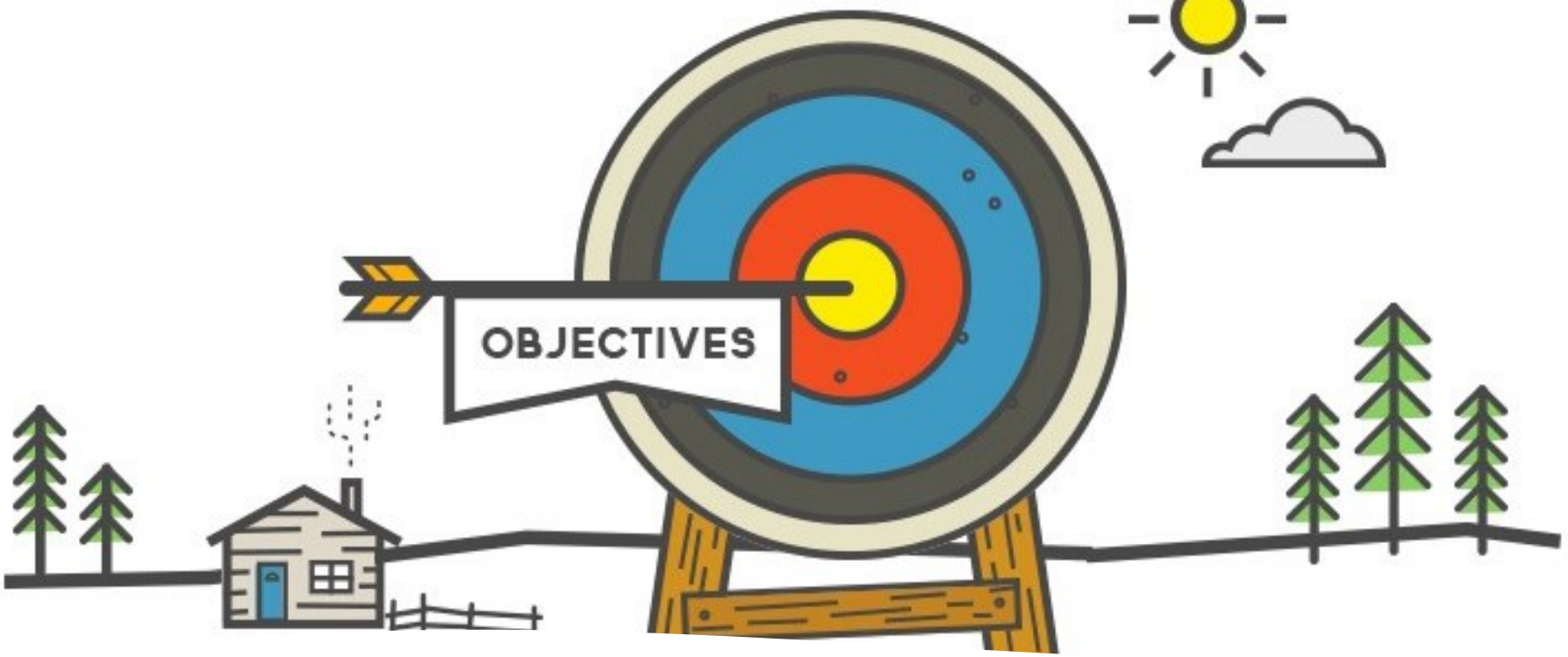
Future Trends in Big Data for Sustainability

Carlos J. Costa



Learning Goals

By the end of this session, you should be able to:

- Reconnect Big Data concepts to real-world systems
- Understand how AI, IoT, and blockchain complement analytics
- Analyze how data-driven systems support sustainability
- Identify the role of analytics in future business operations

Session Overview

- Big Data (Review)
- From Analytics to Systems
- IoT – Data Generation
- AI – Data Exploitation
- Blockchain – Data Trust
- Integrated Systems & Sustainability
- Discussion

Big Data and Analytics

What is Big Data?

- Large, complex datasets
- Characterized by:
  - Volume
  - Velocity
  - Variety

What You Already Know

- Data pipelines
- Storage & processing (e.g., distributed systems)
- Analytics types:
  - Descriptive
  - Predictive
  - Prescriptive

What Big Data Actually Does

Big Data systems:

- Store large-scale data
- Process data efficiently
- Enable analysis

They answer:

“What can we learn from data at scale?”

What Big Data Does NOT Do Alone

- Does not generate data
- Does not decide autonomously
- Does not ensure data integrity

- Leads to the need for other technologies

What is Sustainable Business?

- Balancing:
  - Economic performance
  - Environmental protection
  - Social responsibility
- Also known as the **Triple Bottom Line**

From Analytics to Systems

From Tools to Systems

Real-world organizations don’t use isolated tools: They use **data systems**

A system includes:

- Data sources
- Processing infrastructure
- Analytical models
- Governance mechanisms

Functional Roles in Data Systems

- Think in **functions**, not technologies:

Function	Role
Data generation	Where data comes from
Data processing	How data is handled
Data analysis	How value is extracted
Data validation	How trust is ensured

Mapping Technologies to Functions

- IoT - Data generation
- Big Data - Data processing
- AI - Data analysis
- Blockchain - Data validation

Key Concept

These technologies are **complementary components of the same system**

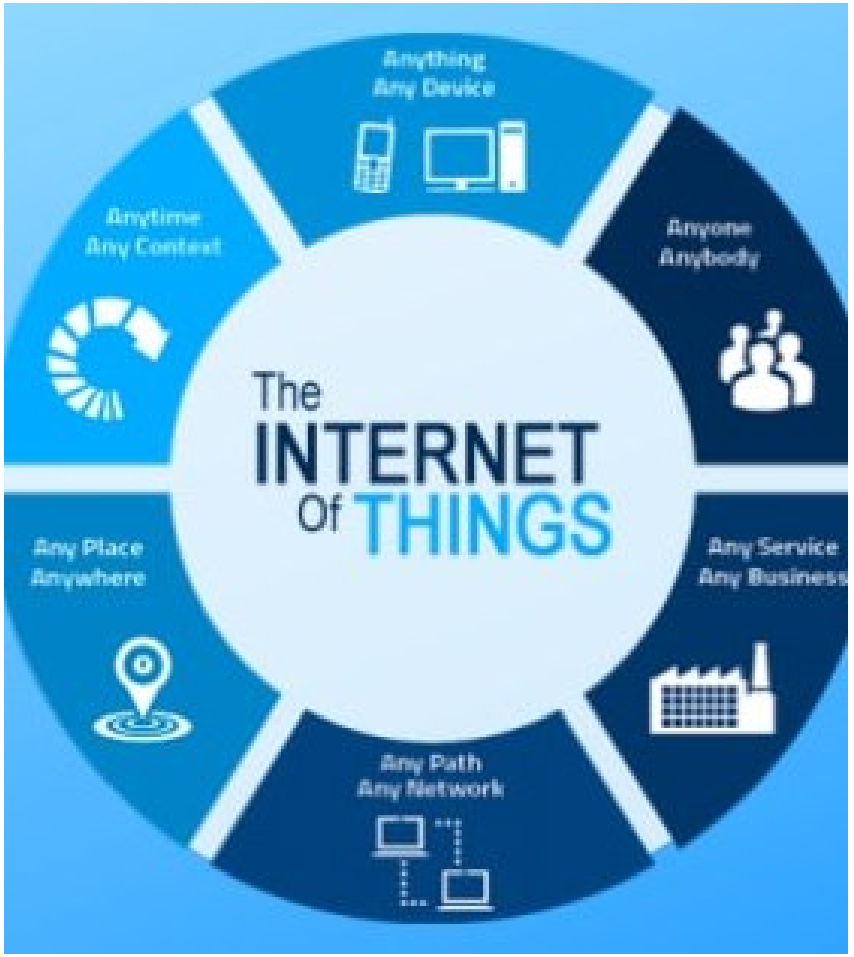
IoT: Data Generation

What is IoT?

- Network of connected devices
- Continuous data collection

Examples:

- Sensors
- Smart devices



Why IoT Matters for Analytics

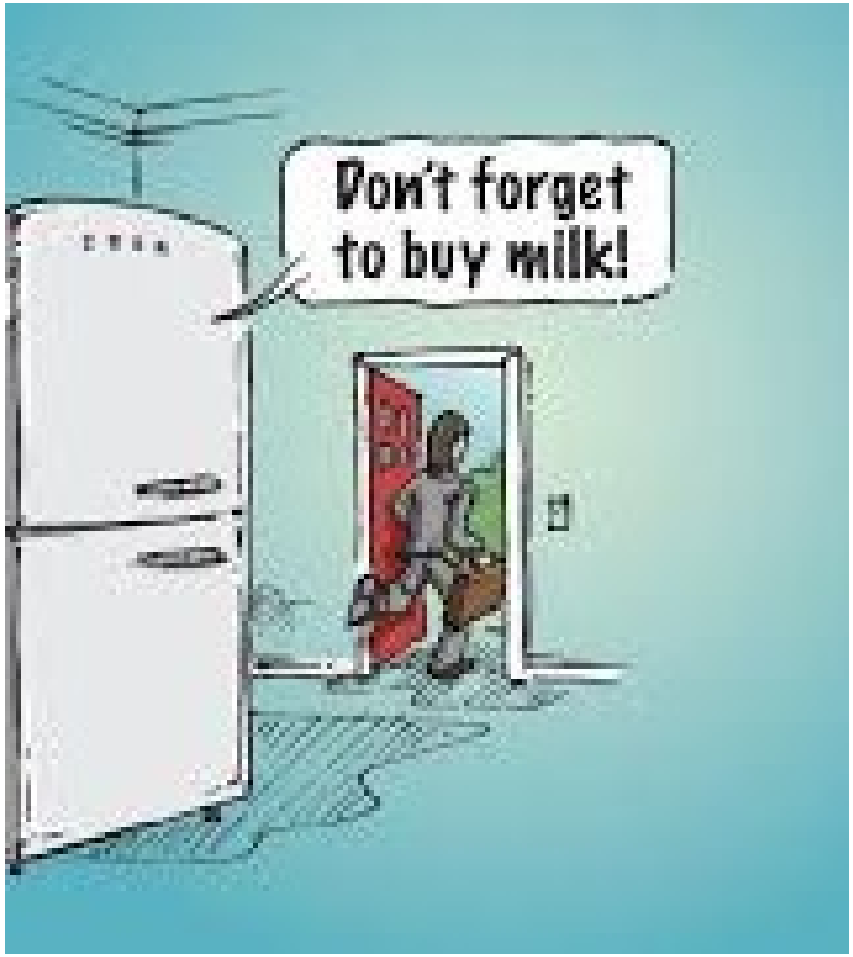
- Produces high-frequency, real-time data
- Expands observable phenomena

Without IoT:

- Many datasets would not exist

Example Use Cases

- Sensors reducing water waste in farming
- Smart meters monitoring electricity usage
- Fleet tracking reducing fuel consumption



Sustainability Applications

- Smart agriculture (water usage)
- Energy monitoring (smart grids)
- Logistics tracking (fuel efficiency)

"Bad news - the scale is threatening to cut off our access to the fridge..."

Benefits of IoT

- Real-time monitoring
- Improved efficiency
- Predictive maintenance
- Reduced waste

Challenges of IoT

- Security risks
- Data overload
- Integration complexity
- Infrastructure costs

Key Insight

IoT expands

what can be measured

AI: Data Exploitation

What is AI in This Context?

- Machine learning models
- Pattern recognition
- Prediction and optimization

AI and Big Data

- AI depends on:
  - Data quantity
  - Data quality

- No meaningful AI without data

Role in Business

- Forecast demand
- Optimize operations
- Automate decisions

Sustainability Applications

- Reduce waste (forecasting)
- Optimize energy use
- Improve logistics effic

Benefits of AI

- Efficiency improvements
- Cost reduction
- Better forecasting
- Reduced environmental impact

Challenges of AI

- Data dependency
- Bias and ethics
- High implementation cost
- Skills gap

Key Insight

AI transforms data into

actionable decisions

Blockchain: Data Trust

What is Blockchain?

- Blockchain is a decentralized and distributed ledger technology that has gained prominence due to its application in cryptocurrencies like Bitcoin.

  https://andersbrownworth.com/blockchain/

What is Blockchain?

- Distributed, secure digital ledger
- Distributed ledger
- Immutable records
- Shared across participants
- Key features:
  - Transparency
  - Immutability
  - Decentralization

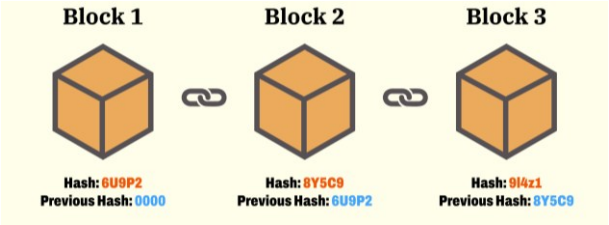
Consensus Mechanism

- Consensus mechanisms are protocols that enable nodes in the network to agree on the state of the ledger.
- Common consensus mechanisms include Proof of Work (used in Bitcoin) and Proof of Stake.
- These mechanisms prevent malicious actors from manipulating the ledger

	Proof-of-work	Proof-of-stake
Power consumption	Huge amounts of electricity required to secure the blockchain due to the processing needed.	Much lower amounts of electricity required to secure the blockchain.
Security	Required to have more than 50% of the processing power to hack.	Required to have more than 50% of the stake (coins) to hack. Can be more expensive to hack due penalties defined in the protocol such as loss of the stake.
Risk of centralisation	There is a risk of having mining pools, group of miners working together, controlling vast amounts of mining power. Currently, three different mining pools control more than 50% of the mining power [9].	Lower risk due to economies of scale being less of an issue. Not dependent on mining equipment.

Abreu, Aparicio & Costa (2019)

Blocks and Transactions



- Transactions are grouped together in blocks, and each block contains a reference to the previous block, forming a chain.
- This chain of blocks ensures the chronological order and integrity of transactions.

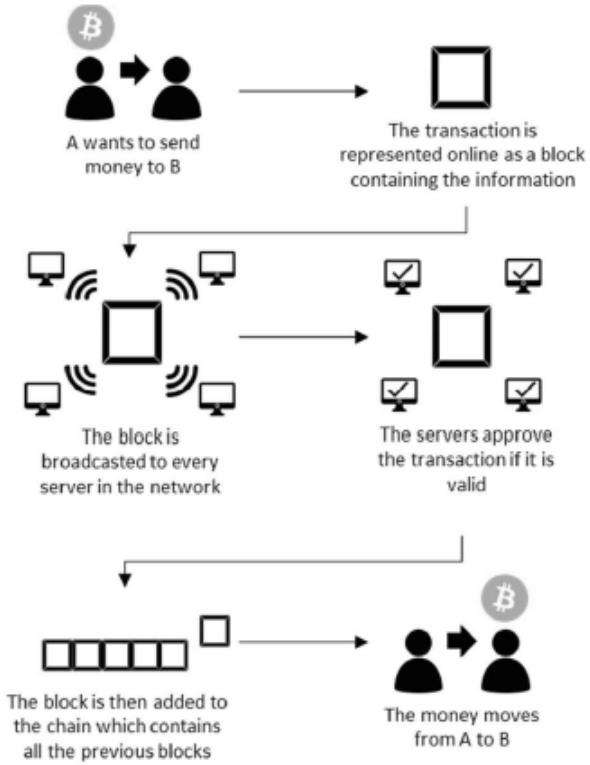
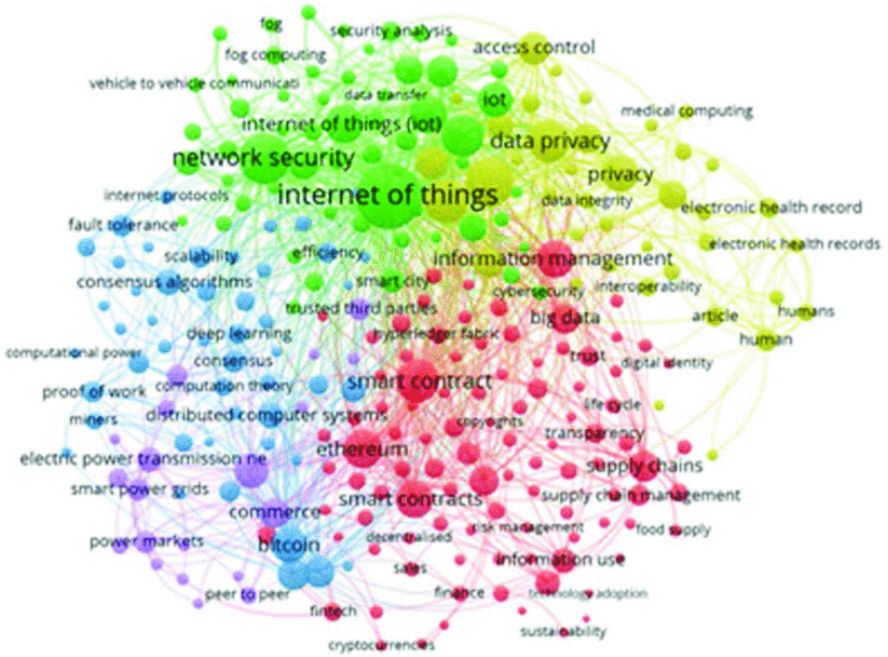
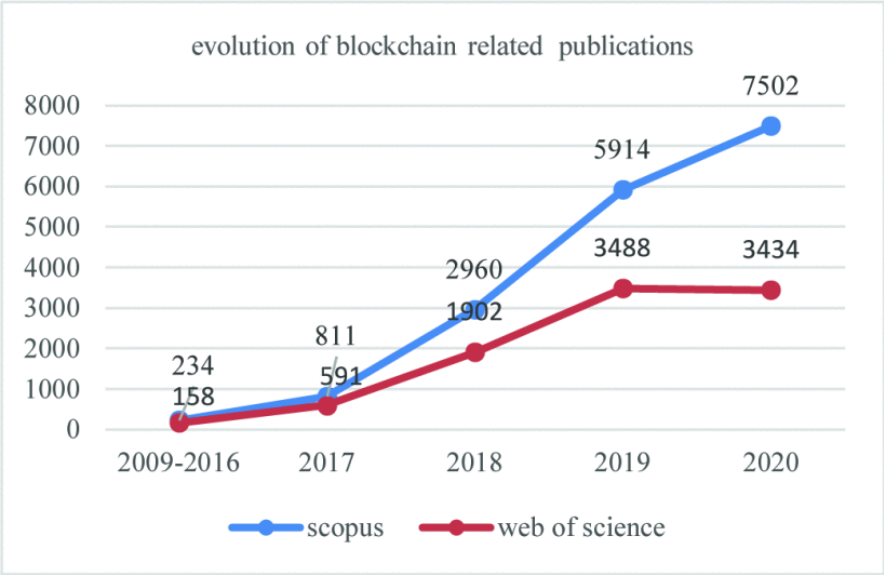


Figure 2: How does a transaction in the blockchain work?

Abreu, Aparicio & Costa (2019)





Bernardino, Costa, & Aparício, (2022)

Research

Decentralized Finance (DeFi)

- is a financial system built on blockchain technology that aims to recreate and improve upon traditional financial services in a decentralized manner.

Decentralized Finance (DeFi)

- is a financial system built on blockchain technology that aims to recreate and improve upon traditional financial services in a decentralized manner.

Why Trust Matters in Data Systems

Problems:

- Data manipulation
- Lack of transparency
- Uncertain data provenance

Role of Blockchain

- Ensures integrity of records
- Tracks origin of data
- Enables auditability

Example Use Cases

- Tracking origin of raw materials
- Verifying fair trade ptoducts
- Monitoring emissions

Sustainability Applications

- Supply chain traceability
- Ethical sourcing verification
- Carbon emission tracking

Benefits of Blockchain

- Trust and transparency
- Reduced fraud
- Improved traceability

Challenges of Blockchain

- Energy consumption
- Scalability issues
- Regulatory uncertainty
- Adoption barriers

Key Insight

Blockchain addresses: “Can we trust the data and its history?”

Integrated Systems

A Data-Centric System

A realistic system includes:

- IoT → generates data
- Big Data → processes data
- AI → analyzes data
- Blockchain → validates data

Example: Sustainable Logistics

- IoT sensors - track vehicles
- Big Data - aggregate data
- AI - optimize routes
- Blockchain - verify emissions data

Role of Technology

Technologies enable:

- Resource efficiency
- Transparency
- Automation
- Data-driven decision-making

System-Level Insight

Value comes from **integration**, not individual technologies

Wrap-Up

Final Takeaways

- Big Data enables large-scale analysis
- IoT expands data availability
- AI extracts value
- Blockchain ensures trust
- Real impact comes from **integrated systems**

References

Abreu, P. W., Aparicio, M., & Costa, C. J. (2018). Blockchain technology in the auditing environment. In 2018 13th Iberian Conference on Information Systems and Technologies (CISTI) (pp. 1-6). IEEE.

Aparicio, J. T., Romao, M. & Costa, C. J. (2022) "Predicting Bitcoin prices : The effect of interest rate, search on the internet, and energy prices," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1-5, IEEE

Bernardino, C, Costa, J. & Aparício, M. (2022) "Digital Evolution: blockchain field research," 2022 17th Iberian Conference on Information Systems and Technologies (CISTI), pp. 1-6, IEEE

Boritz, J. E. (2005). IS practitioners' views on core concepts of information integrity. International Journal of Accounting Information Systems, 6(4), 260-279.

Cesario, F., J. Costa, C., Aparicio, M., & Aparicio, J. (2023). Blockchain Technology Adoption: Factors Influencing Intention and Usage. In A. R. da Silva, M. M. da Silva, J. Estima, C. Barry, M. Lang, H. Linger, & C. Schneider (Eds.), Information Systems Development, Organizational Aspects and Societal Trends (ISD2023 Proceedings). Lisbon, Portugal: Instituto Superior Técnico. ISBN: 978-989-33-5509-1. https://doi.org/10.62036/ISD.2023.9

Christensen, C. M., Bohmer, R., & Kenagy, J. (2000). Will disruptive innovations cure health care?. *Harvard business review*, *78*(5), 102-112.

Christensen, C. M., Raynor, M. E., & McDonald, R. (2015). Disruptive innovation. *Harvard Business Review*, *93*(12), 44-53

Costa, C. J. (2024). DeFi: Concepts and Ecosystem. arXiv preprint arXiv:2412.01357. https://doi.org/10.48550/arXiv.2412.01357

González-Mendes, S, González-Sánchez, R., Costa, C. & García-Muiña, F. (2023) "Analysing the state of the art of Blockchain application in Smart Cities: A bibliometric study," 2023 18th Iberian Conference on Information Systems and Technologies (CISTI), Aveiro, Portugal, pp. 1-6, doi: 10.23919/CISTI58278.2023.10211371.

Nagy, D., Schuessler, J., & Dubinsky, A. (2016). Defining and identifying disruptive innovations. *Industrial Marketing Management*, *57*, 119-126.

Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system.

Raymond, E. (1999). The cathedral and the bazaar. *Philosophy & Technology*, *12*(3), 23.