HAPPY DIGITAL X
CITIES, SYSTEMS, PRODUCTS AND SERVICES
Kura Kura Bali

清华大学东南亚中心
Tsinghua Southeast Asia Center

清華大学深圳国际研究生院
Tsinghua Shenzhen International Graduate School

Introduction to GenAI

Happy Digital X

Happy Digital X | Tsinghua University

Today’s Agenda

1 Data Governance

Privacy regulations and data management

2 Product Development

Lifecycle, deployment, and ROI

Duration: 1 hour

Data Governance

The Data Imperative

“Organizations don’t have AI problems; they have data problems that AI exposes.”

Plan for 60–80% of GenAI project time
to be spent on data preparation.

Data Strategy Precedes AI Strategy

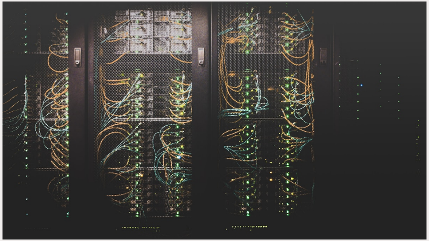
The Data Hierarchy of Needs:

1 **Data Collection** — Foundation

2 **Clean Data** — Must start here

3 **Analytics & Reporting**

4 **AI/ML** — Most start here (mistake)



Reality Check

Fortune 500 expected 4 months for GenAI. Actual: 15 months. Root cause: Data readiness.

Data Requirements for GenAI

- **Training Data**: Building/fine-tuning models
  Strategic value: Competitive moat
- **Context Data (RAG)**: Grounding model outputs
  Strategic value: Accuracy & relevance
- **Operational Data**: Real-time model inputs
  Strategic value: Timeliness

Quality Dimensions: Accuracy, Completeness, Consistency, Timeliness, Representativeness

Global Privacy Regulations

- **GDPR** (EU): Up to 4% global revenue
- **CCPA/CPRA** (California): Per-violation penalties
- **PIPL** (China): Up to 5% revenue
- **LGPD** (Brazil): Up to 2% revenue
- **POPIA** (South Africa): Up to 10M ZAR

Global Trend

Design AI systems with privacy by default.

GenAI-Specific Privacy Concerns

1. **Training Data Privacy**: Was personal data used with consent?
2. **Inference Privacy**: Can model be manipulated to reveal data?
3. **Output Privacy**: Do outputs contain personal information?
4. **Conversation Privacy**: Who accesses user interactions?
5. **Derived Data**: Are new personal insights generated?

The Consent Challenge

Traditional consent breaks down: capabilities hard to explain, data use unpredictable, untraining technically difficult.

Data Governance Framework

Key Components

- Data inventory & classification
- Access controls
- Consent management
- Retention policies
- Audit trails

Best Practices

- Minimize data collection
- Purpose limitation
- Regular compliance audits
- Incident response plans
- Cross-border controls

User Rights to Support

- **Right to Access**: Users request all data held about them
- **Right to Erasure**: Users request deletion
- **Right to Portability**: Data in machine-readable format
- **Right to Rectification**: Correct inaccurate data
- **Right to Object**: Object to certain processing
- **Automated Decision Rights**: Human review of AI decisions

China’s AI Regulatory Framework

The world’s most comprehensive AI regulations:

- **Algorithm Recommendations** (2022): Internet services
- **Deep Synthesis** (2023): Deepfakes, synthetic media
- **GenAI Service Measures** (2023): All public GenAI
- **AIGC Labeling** (Sept 2025): Mandatory AI content labels
- **National Standards** (Nov 2025): Security & governance

Scale: 350+ LLMs filed. 1.57M AI patents (38.6% of global total).

Product Development

The GenAI Development Reality

Key Statistics (2025)

Only **5%** of AI pilots achieve rapid revenue acceleration
67% success rate for purchasing/partnering
22% success rate for internal builds
46% have no structured ROI measurement

GenAI has entered the “Trough of Disillusionment”

Why Traditional Project Management Fails

Traditional

- Fixed requirements
- Binary success
- Predictable timeline
- Deterministic testing

GenAI

- Emergent requirements
- Probabilistic success
- Uncertain timeline
- Statistical testing

Implication

Waterfall always fails. Agile is better but insufficient

The AI Project Lifecycle

1 **Problem Framing** (Often Skipped): Should AI solve this?

2 **Data Assessment**: Inventory, gaps, quality

3 **Proof of Concept** (4–8 weeks): Time-boxed experimentation

4 **Pilot**: Limited production, controlled blast radius

5 **Production & Scale**: Infrastructure, monitoring

6 **Operations**: Performance monitoring, retraining

Rule of Thumb

Budget for 2–3 PoCs failing for every success.

Phase Gates for GenAI

- **Gate 0**: Business case, feasibility, ethics screening
- **Gate 1**: Requirements, data availability, build vs. buy
- **Gate 2**: Technical validation, benchmarks, user feedback
- **Gate 3**: Production-grade, security & ethics review
- **Gate 4**: Controlled deployment, monitoring setup
- **Gate 5**: Full deployment, continuous improvement

Kill Criteria: Define Before Starting

- **Technical**: Can't achieve accuracy threshold
- **Economic**: Cost exceeds value
- **Timeline**: 6-month delay, no path forward
- **Ethical**: Can't mitigate bias
- **Security**: Can't protect data
- **Regulatory**: Unacceptable compliance risk
- **Strategic**: Market opportunity gone

Imperative

Implementation Patterns

1 **Co-Pilot / Augmentation**

AI assists; humans decide. *Best for: High-stakes, building trust*

2 **Automation with Exceptions**

AI handles routine; humans handle exceptions. *Best for: High-volume*

3 **Full Automation**

AI autonomous with monitoring. *Best for: Low-stakes, speed critical*

4 **Internal Tool**

AI assists employees only. *Best for: Building capability, lower risk*

Build vs. Buy Decision

- **Build from Scratch**: $10M–$100M+; 12–24 months
  Only if: Massive data advantage
- **Fine-Tune**: $10K–$1M; weeks to months
  Best for: Domain-specific tasks
- **RAG**: $10K–$100K; weeks
  Best for: Current/proprietary information
- **Prompt Engineering**: $1K–$10K; days to weeks
  Best for: Quick wins
- **Buy SaaS**: Variable; days
  Best for: Non-differentiating capabilities

Success Metrics

Avoid Vanity Metrics:

✗ “We deployed an AI model”

✗ “95% accuracy” (on what?)

Focus on Business Outcomes:

✓ Customer satisfaction improved by X%

✓ Time to resolution decreased by Y hours

✓ Cost per transaction reduced by $Z

✓ Employee time redirected to higher-value work

Four-Layer Monitoring Framework

1 **Infrastructure**: Latency, error rates, throughput, cost

2 **Model Performance**: Accuracy, hallucination rate, drift

3 **Business**: Adoption, task completion, satisfaction, revenue

4 **Risk**: Incidents, near-misses, compliance, complaints

Principle

You can’t improve what you don’t measure. Monitor from day one.

ROI Reality (2025)

- Average ROI: **3.7x** per dollar (IDC/Microsoft)
- Top performers: **$10.3** return per dollar
- 74% meeting or exceeding expectations (Deloitte)
- **46% have no structured ROI measurement**

Timeline Expectations:

- Chatbots, RPA: 6–12 months
- Operational efficiency: 12–24 months
- Revenue generation: 18–36 months

Total Cost of Ownership

Initial Costs

- Infrastructure (GPUs)
- Software licenses
- Integration
- Data preparation
- Training

Ongoing Costs

- Compute resources
- API fees
- Model maintenance
- Monitoring
- Personnel

Hidden Costs: Compliance, legal/IP, incidents, technical debt, failed pilots

Minimum Viable AI Team

- **Executive Sponsor** (10–20%): Alignment, resources, blockers
- **Product Owner** (Full-time): Requirements, prioritization
- **Data Engineer** (Full-time): Pipelines, quality
- **ML Engineer** (Full-time): Model development
- **Domain Expert** (25–50%): Business logic, validation
- **MLOps Engineer**: Deployment, monitoring

Part 1 Key Takeaways

Summary

1. **Data First**: 60–80% of GenAI time is data preparation
2. **Privacy by Design**: Global regulations require it
3. **Expect Failure**: Budget for 2–3 PoCs failing per success
4. **Define Kill Criteria**: Before emotional investment
5. **Measure Everything**: Connect to business outcomes
6. **Build the Right Team**: Minimum viable AI team

Discussion Questions

1 What is the current state of data readiness in your organization?

2 Have you defined clear kill criteria for your AI projects?

3 How are you measuring ROI on AI investments today?

4 Do you have the right team composition for AI success?

Thank You

HAPPY DIGITAL X
CITIES, SYSTEMS, PRODUCTS AND SERVICES
Kura Kura Bali

www.hdx.edu

info@hdx.edu

@HappyDigitalX

Continue to Part 2: Ethics, Security & Imple