How Data Can Supercharge L&D

Lori Niles-Hofmann

Founder 8Levers

www.theeightlevers.com

+001.647.302.2227

Tip

Spend time with teenagers.

Three Levels of Analytics

1. Predictive
2. Triangulation
3. Closed-Loop Reporting

L&D runs on 20+ year-old technology

SCORM

Shareable: Content is designed to be reused across different systems and courses. Instead of creating new learning material for each platform, SCORM packages can be moved and reused easily

Content: Refers to learning material — such as lessons, quizzes, or entire courses — usually in digital format

Object: Each piece of content is treated as a building block (or object). This modular approach allows different pieces to be assembled into various courses

Reference Model: SCORM is a set of technical standards and guidelines that reference existing web technologies (like JavaScript, HTML, XML) to ensure interoperability between learning content and Learning Management Systems (LMSs)

SCORM

✅ Advantages

Interoperability — Works across different LMS platforms

Reusability — Content can be repurposed without major rework

Tracking — Monitors learner progress, completion, and scores

Standardization — Common technical framework simplifies content creation

⚠️ Limitations

Rigid Structure — Limited support for adaptive or personalised learning

Outdated Technology — Built on older web standards like JavaScript APIs

Limited Data — Only basic tracking like completion, score, and time

Complex Setup — Packaging and troubleshooting SCORM can be technical and time-consuming

LRS and xAPI

xAPI (Experience API) captures a wide range of learning experiences, both online and offline, in a consistent format

LRS (Learning Record Store) is a specialised data store that collects, stores, and retrieves learning activity data generated by xAPI

Together, xAPI and an LRS allow organisations to capture detailed learning behaviours such as reading an article, attending a seminar, completing a course, or interacting with a simulation beyond traditional LMS tracking

xAPI Data Examples

Element	Description
Actor	Who performed the activity (usually a learner, identified by name or email)
Verb	What action was taken (e.g., *completed*, *viewed*, *answered*, *failed*)
Object	What the action was taken on (e.g., a course, module, document, or video)
Result	Optional – score, success/failure, time spent, completion status
Context	Optional – additional metadata like platform, instructor, team, device
Timestamp	When the activity occurred
Authority	Who recorded the statement (e.g., LMS, platform, or device)
Attachments	Optional – supporting files like certificates, screenshots, PDFs

Challenges with Data in EdTech

Fragmented Systems	Organisational Gaps	Strategic Blind Spots
Data trapped in silos	Lack of analytics skills in teams	Difficulty linking data to outcomes
Limited interoperability	No clear ownership or governance	Overreliance on surface metrics
Inconsistent formats and poor data quality	Slow adoption of AI and predictive tools	Missing context around learner activities
Vendor lock-in and privacy concerns	Delays in turning data into insights	Inability to measure informal and social learning

Why was there a spike in searches for “my eyes hurt” in April 2024?


ITS A SOLAR ECLIPSE


2019
Train Your Dragon 3
take a screenshot
tie a tie
draw
change PSN name
Get Away With Murder
make slime
pronounce
lose weight
solve a Rubik's cube
delete Instagram
boil eggs
lose belly fat
write a check
scan a network for hidden cameras
2020
make a mask
make hand sanitizer
boil eggs
take a screenshot
file for unemployment
make French toast
cut your own hair
make whipped coffee
make buttermilk
change background on Zoom
draw
watch Trolls 2
get free Robux
delete Instagram
dye Easter eggs

Autopsy vs. Diagnosis

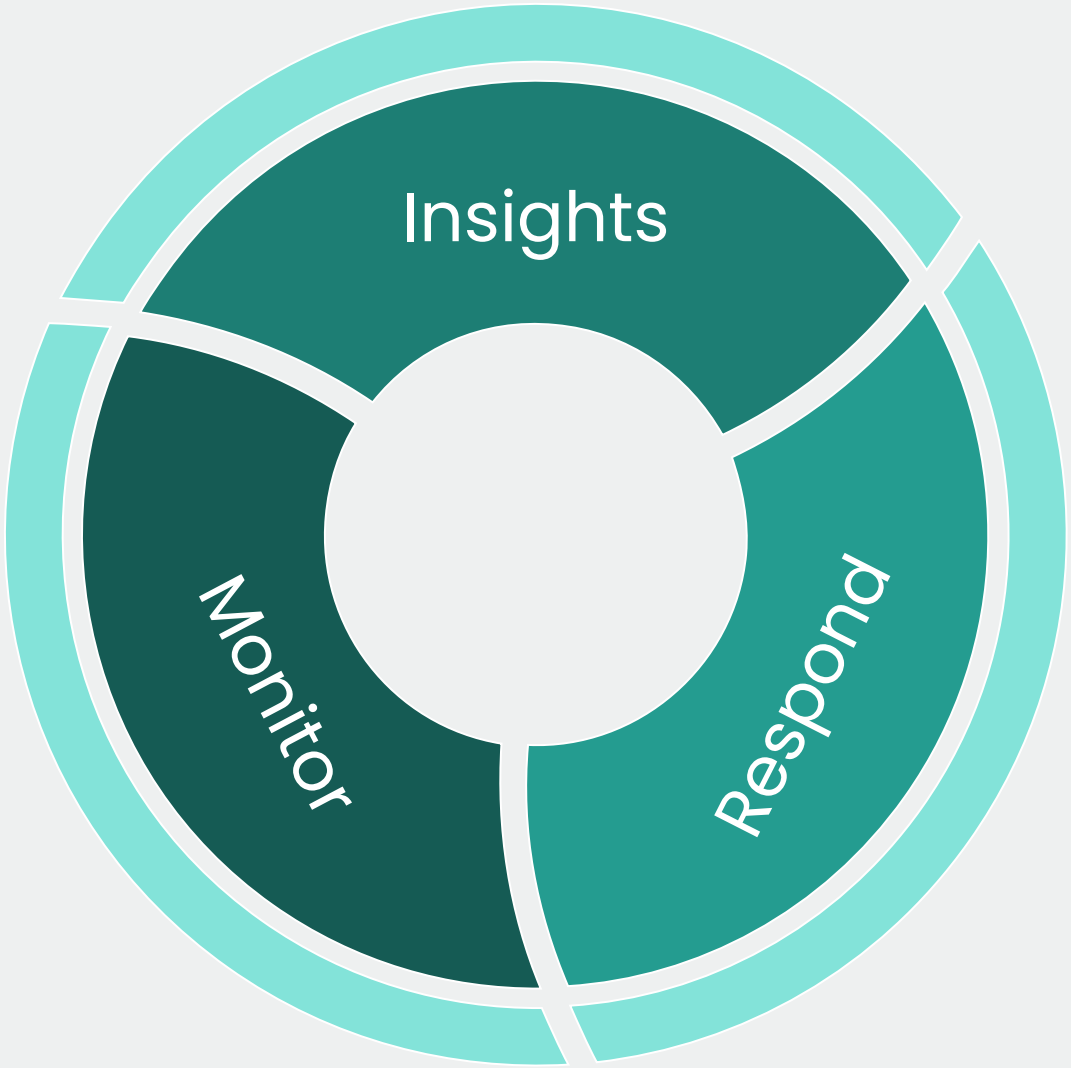
A simple data story

Kabbage

- Surge in searches for "Kabbage"
- Competitor offering lower rates
- Call times and churn rising
- Staff re-skilled to retain customers
- SME identified and job aid created
- Customer retention and reduce wait times

Data Source	Use Case
Internal search terms	Identify topics employees are struggling with or curious about
External search trends	Detect emerging skills and industry knowledge gaps
Internal platform hashtags	Spot trending interests and challenges
Performance goals (KPIs)	Align learning to business targets and skills gaps
Help desk tickets	Highlight technical and process issues needing training
Exit interviews	Identify gaps in onboarding and development support
Employee engagement surveys	Surface perceived skills gaps and growth areas

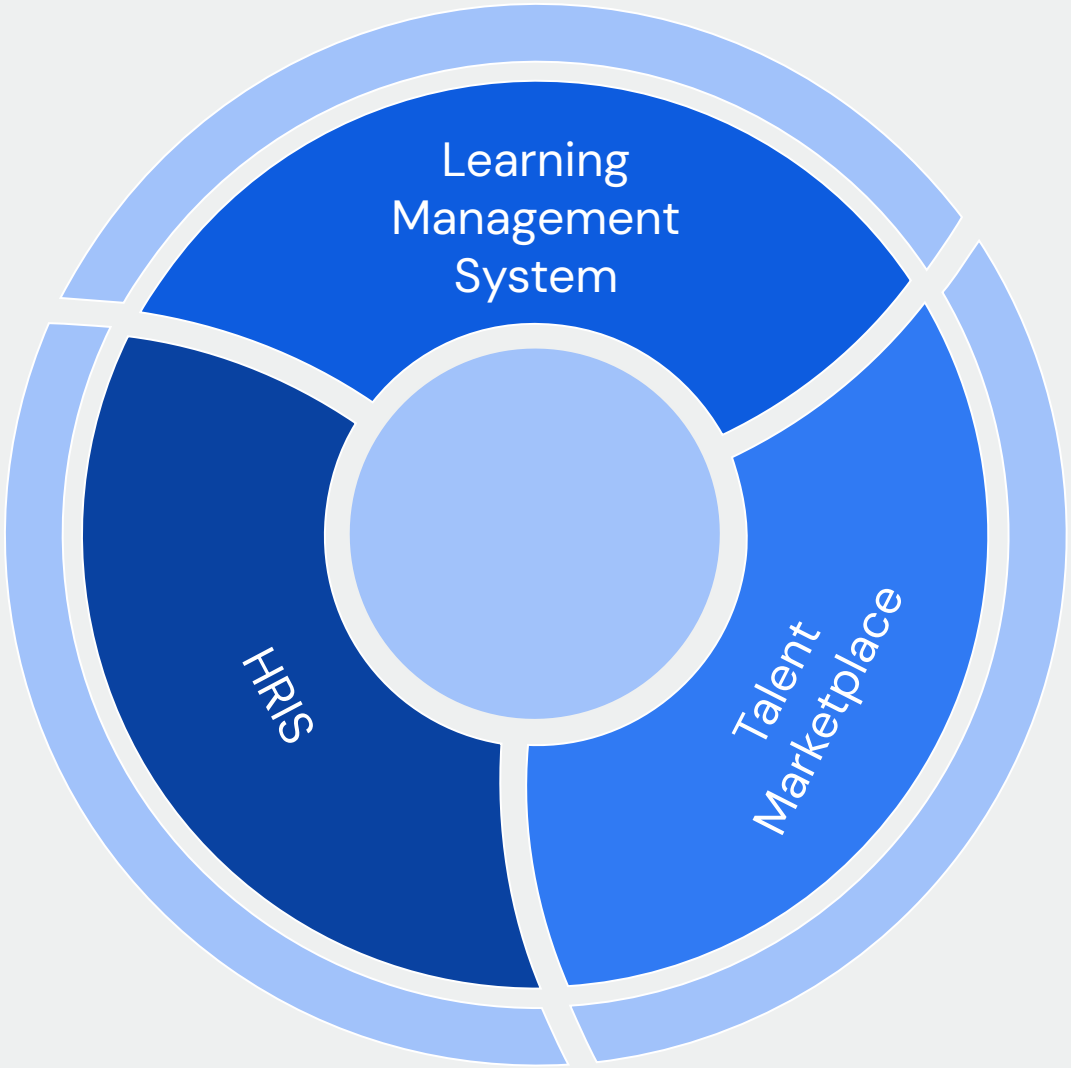
Data Source	Use Case
LMS usage data	Track course completion, drop-offs, and content popularity
Project management tools (e.g., JIRA)	Identify workflow bottlenecks and process training needs
Customer feedback (e.g., NPS scores)	Reveal service and product knowledge gaps
Performance review comments	Detect common development needs across teams
Industry certification trends	Predict future credentialing and upskilling needs
Meeting transcripts	Capture recurring misunderstandings or knowledge gaps
Internal chat transcripts	Surface real-time support needs and peer knowledge gaps
Recruitment trends	Signal evolving skills required for new roles


Insights
Respond
Monitor

Tip

Set up data feeds to track as many of these data sources as possible

Triangulation



Merging Data Sets

Use an **outer join** on `Employee_ID` to capture all records:

In Python (e.g., using pandas):

```
merged = pd.merge(hris_df, lms_df, on='Employee_ID',
how='outer', indicator=True)
```

- This adds an `_merge` column:
  - `both`: Found in HRIS and LMS
  - `left_only`: HRIS only
  - `right_only`: LMS only

Or, just ask ChatGPT/Copilot, etc.

DataRobot or Amazon SageMaker

What They Do:

- Ingest structured data (CSV, Excel, JSON)
- Automatically clean, normalise, and analyse datasets
- Detect patterns, trends, and predictive relationships without writing code
- Highlight the most influential factors behind learner behaviours, such as course dropouts or completion rates

How to Use:

1. Upload datasets from LMS, Talent Marketplace, and HRIS (CSV format works best).
2. Set Employee ID or Learner ID as the join key if needed across datasets.
3. Choose the target variable you want to predict or understand, such as:
    - Course Completion (Yes/No)
    - Time Spent on Course
    - Voluntary Access to Marketplace
4. Run "AutoML" to let the platform build models and rank what factors (features) are influencing outcomes.
5. Review Feature Importance charts to learn which inputs (device type, content length, job role) have the biggest effects.

Prompt:

You are a highly skilled data analyst.
I will upload three .csv (or Excel) files containing employee data from:

- LMS (learning management system)
- Talent Marketplace
- HRIS (human resources information system)

Each file has a field for employee number, but not all employees will appear in every file, and the rows are not aligned.
Instructions:

1. Merge the three datasets using **Employee Number** as the unique identifier
2. Create a unified dataset where each employee's information from all three sources is aligned side-by-side
3. Retain all records, even if an employee appears in only one or two of the sources (outer join)

Next:

- Analyse the merged dataset for **patterns, gaps, and anomalies**
- Automatically detect interesting insights — you decide what is statistically significant, unusual, or noteworthy
- Look for areas such as:
  - Skills gaps
  - Learning engagement vs mobility
  - Talent Marketplace participation vs HRIS performance or tenure
  - Employees missing from one or more systems
  - Possible data quality issues

Output:

- Provide a **summary of insights** in clear, bullet-point format
- Suggest possible **actionable recommendations** based on what you observe
- If relevant, suggest further **data visualisations** that would help explore these insights (e.g., bar charts, scatterplots)

Additional Notes:

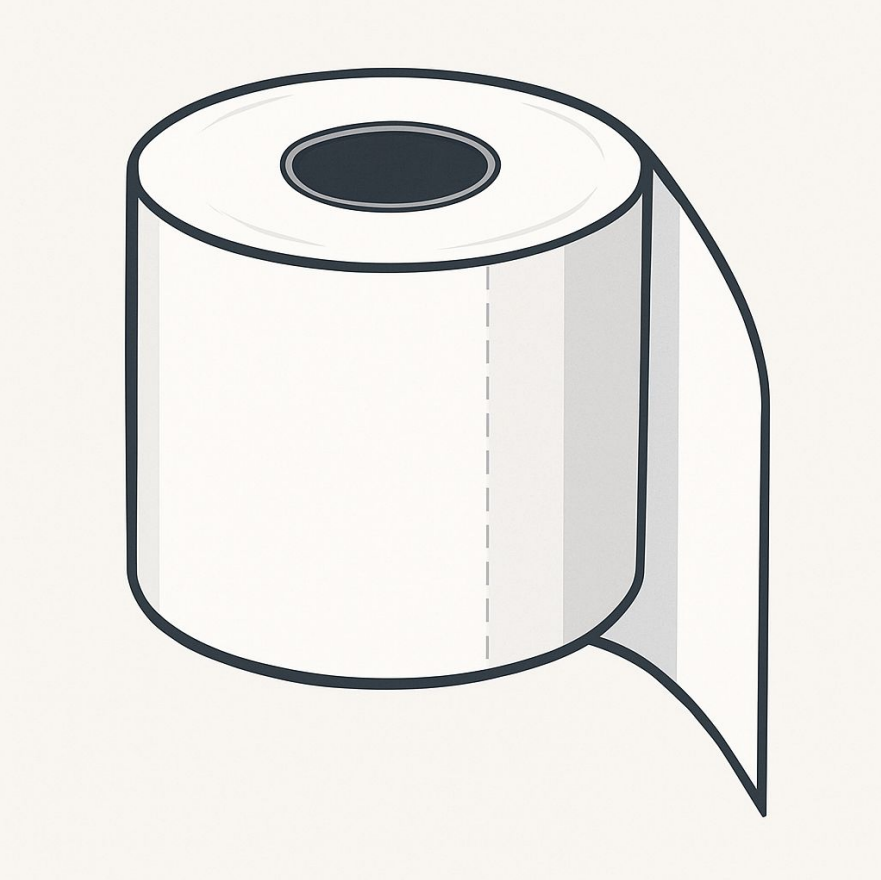
- You are allowed to create derived fields if it will improve the analysis (e.g., active vs inactive learner, frequent marketplace user
- Focus on surfacing anything that a human might miss on first glance
- Prioritise clarity and impact in your findings

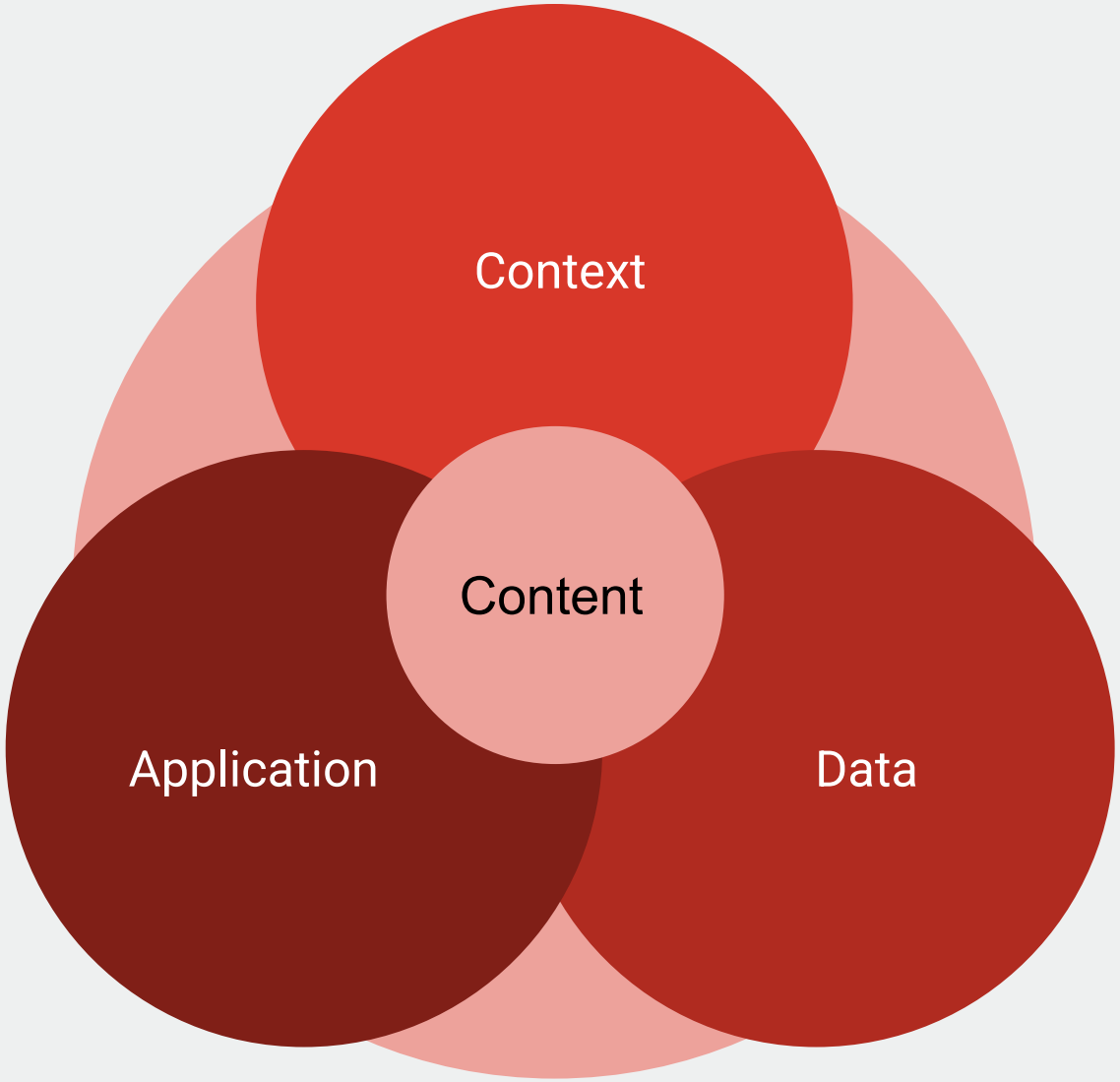
Skills Gaps and Learning Needs	Readiness and Succession
Skills identified in the Talent Marketplace but unsupported by learning in the LMS	Employees with strong performance, active learning, and Marketplace interest
Courses completed in the LMS but not linked to roles or gigs in the Talent Marketplace	Critical roles in HRIS with no internal successors in the Talent Marketplace
Skills learned historically but not applied or endorsed recently	Internal applicants consistently missing required certifications or badges
Employees applying for roles they are under-skilled for based on LMS data	Departments with high learning engagement but low internal mobility
High-demand skills with no internal supply	High-potential employees not being offered developmental opportunities

Engagement and Utilisation	Equity and Inclusion
Employees engaging with the Talent Marketplace but not the LMS	Mobility opportunities accessed unevenly across locations or demographics
Skilled employees not expressing interest in internal opportunities	Learning engagement disparities by team or region
Learning-heavy employees showing no progression or application	High-performing underrepresented groups lacking mobility visibility
Low LMS activity in business-critical units	Learning completions not translating to promotion opportunities
Career stalling indicators based on combined inactivity across systems	Barriers to opportunity hidden in Marketplace search or match algorithms

Organisational Agility	Data Integrity and Process Gaps
Employees applying new skills quickly after learning	Employees present in one or two systems but missing from others
High correlation between learning and successful gig assignments	Talent profiles outdated despite active training
Speed from training completion to internal placement	Courses marked mandatory but not completed by Marketplace candidates
Teams most active across all three platforms showing higher retention	No post-onboarding development activity
Reskilling efforts translating to lateral moves or stretch assignments	Roles posted with outdated skill requirements

A Story About Loo Roll




Context
Content
Application
Data

Practical Example

- An employee completes Level 4 Spanish on the LMS.
- The next time they log into Excel, they're prompted to change their language settings to Spanish, providing an opportunity for continued practice.
- This information is then sent to the talent marketplace, where the employee is matched with an opportunity that offers more chances to speak Spanish, such as a project involving South American clients.

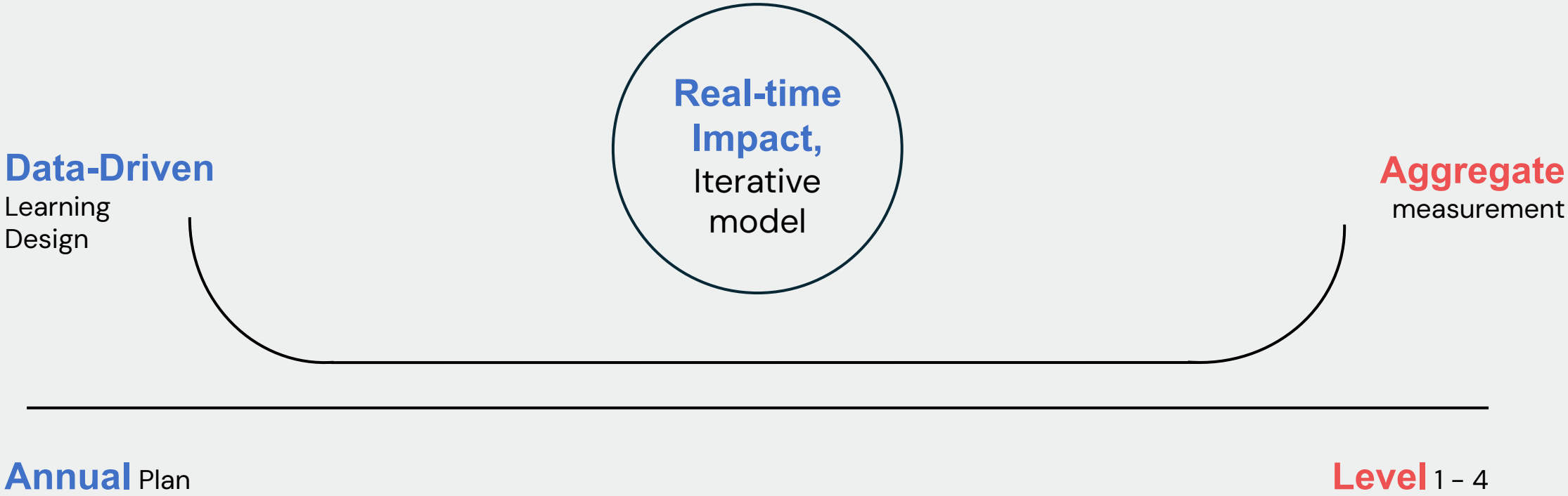
The Lion of Gripsholm

Closed-Loop Reporting

Annual Plan

Level 1 – 4

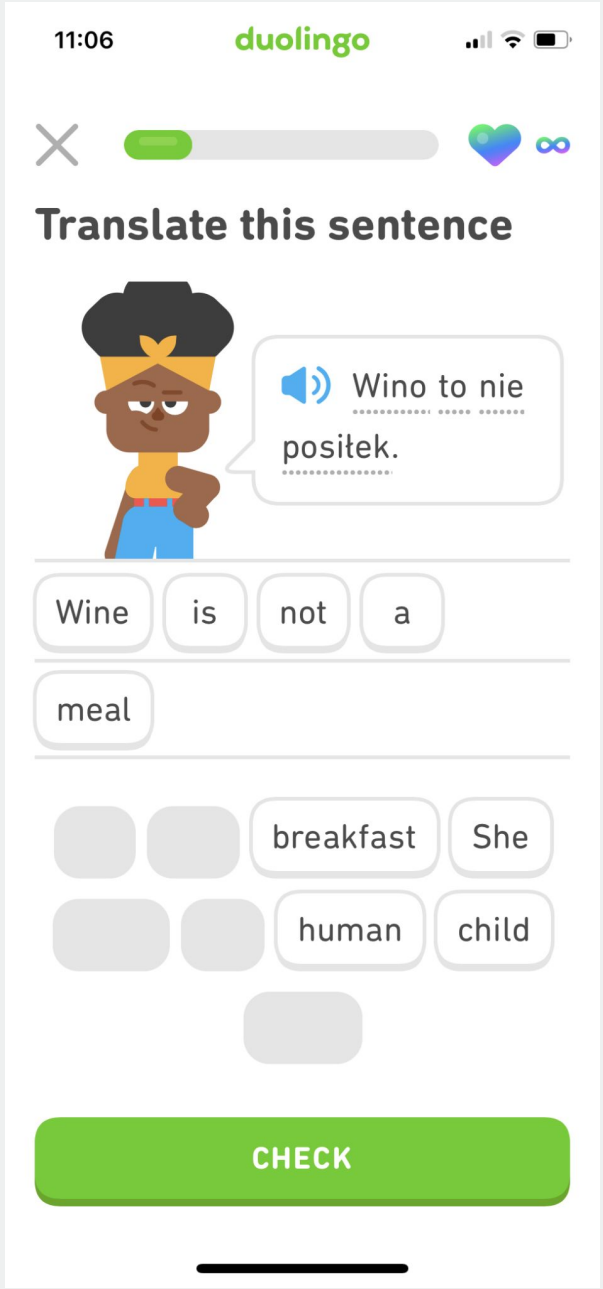
Closed-Loop Reporting

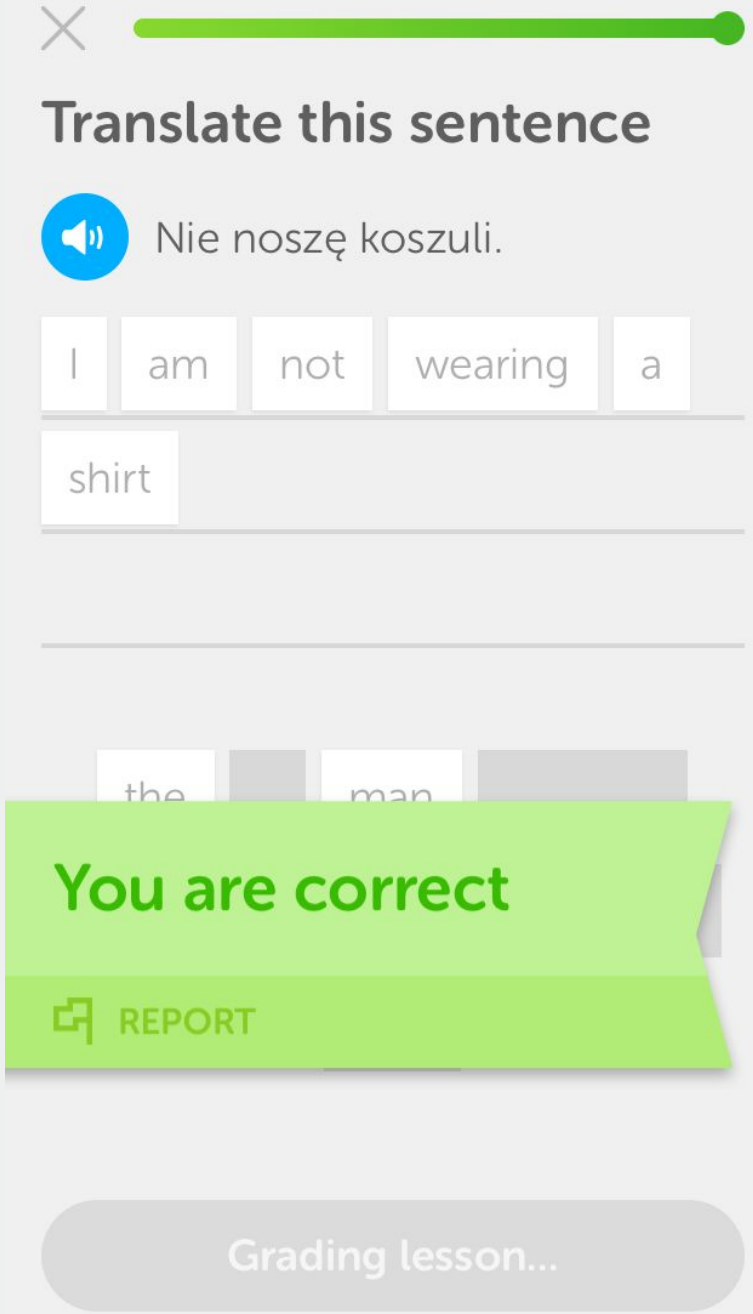

Real-time
Impact,
Iterative
model
Data-Driven
Learning
Design
Aggregate
measurement
Annual Plan
Level 1 - 4

Agent Architecture Overview

Agent Chain:

1. Data Agent (automatically connects to xAPI + CRM)
2. Analysis Agent (identifies misalignment between learning and KPIs)
3. Strategy Agent (defines content improvement approach)
4. Content Generator Agent (rewrites the module, quiz, or copy)
5. Deployment Agent (updates authoring system or flags for review)





Can we Reverse Engineer the learning process?

The latest news from Microsoft:

- Impressively accurate user skill profile inferences based on Microsoft Graph activity data (meetings, documents, chats, emails, etc)
- Has a frequent refresh cadence so skill profiles are always up-to-date and relevant

https://techcommunity.microsoft.com/blog/microsoft365copilotblog/announcing-people-skills-general-availability-and-new-skills-agent/4406364

Tip

If you want to play with data but are concerned with **data privacy**, try generating mock data sets: www.mockaroo.com

Thank you.

Lori Niles-Hofmann

Co-Founder NilesNolen

+001.647.302.2227