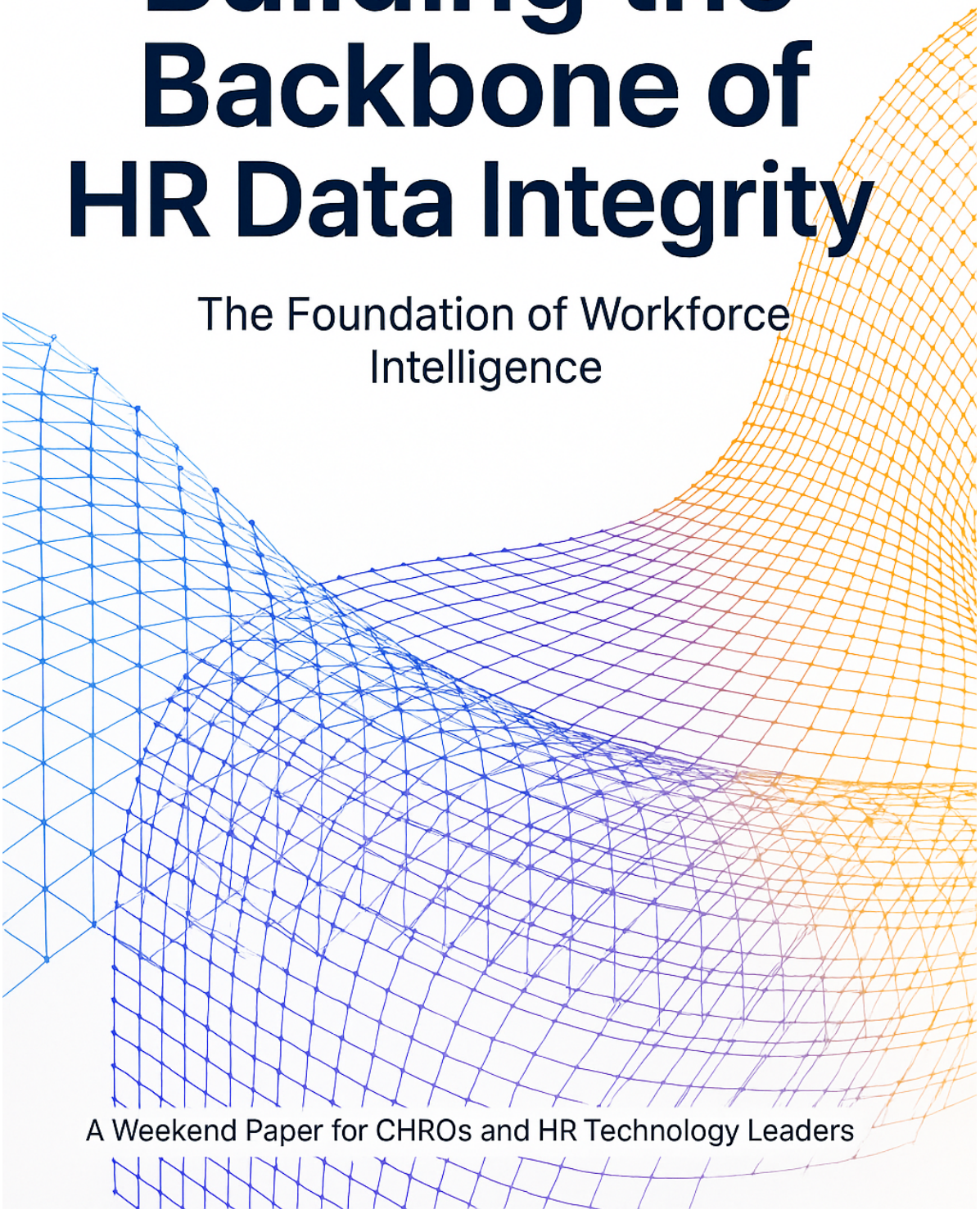
draup

Building the Backbone of HR Data Integrity

The Foundation of Workforce Intelligence



A Weekend Paper for CHROs and HR Technology Leaders

Executive Summary

HR technology is evolving at a remarkable pace. In just the past three years, we have made significant progress—with more HR vendors introducing AI-driven features and capabilities. Yet, building effective AI systems cannot be achieved solely by increasing budgets. The foundation for success lies in understanding and mapping the existing HR data infrastructure—and, just as importantly, the key assumptions and the operating model that shapes it.

This paper highlights the critical data-infrastructure requirements that Strategic Workforce Planners, Talent Acquisition leaders, and CHROs must consider to make AI both practical and valuable. It explores how mapping systems, clarifying data ownership, ensuring quality and interoperability, and establishing governance and privacy frameworks create the groundwork for sustainable AI adoption in HR.

High-Level Overview of HR Technology Systems

The continuous move to the cloud has transformed HR systems. It has led to the development of specialized tools for payroll, recruiting, learning, and analytics, all of which are connected through APIs. This shift enables agility and data-driven insights but also demands strong integration and governance.

1. Core HR Systems (System of Record)

Purpose: Maintain foundational employee data across the organization.
Key Data: Employee master data, job profiles, positions, compensation, benefits, organizational hierarchy, and employment history.
AI Relevance: Serves as the primary data source for workforce identity, structure, and longitudinal analytics.

2. Talent Acquisition Systems

Purpose: Manage sourcing, recruiting, and candidate pipelines.
Key Data: Job requisitions, candidate profiles, resumes, interviews, sourcing channels, hiring conversion metrics.
AI Relevance: Powers skills inference, candidate matching, and predictive hiring analytics.

3. Learning and Development Systems

Purpose: Enable training, learning pathways, and continuous skill development.
Key Data: Course enrollment, completions, skill attainment, certifications, and learning behavior.
AI Relevance: Supports skill taxonomies, personalized learning recommendations, and role-based upskilling insights.

4. Performance and Talent Management Systems

Purpose: Facilitate goal setting, reviews, succession planning, and career progression.
Key Data: Goals, review feedback, competency ratings, succession pipelines, and career aspirations.
AI Relevance: Connects behavioral and performance data to AI-driven potential, readiness, and mobility models.

5. Workforce Planning and Data Analytics Platforms

Purpose: Forecast workforce demand and supply, track costs, and analyse productivity.
Key Data: Headcount plans, FTE projections, attrition models, workforce segmentation, cost analysis.
AI Relevance: Central to AI-based forecasting, scenario modeling, and strategic workforce optimization.

6. Compensation and Rewards Systems

Purpose: Manage salary structures, incentives, and equity programs.
Key Data: Pay bands, bonuses, variable compensation, equity data, and geographic differentials.
AI Relevance: Enables pay equity insights, predictive retention analytics, and total rewards optimization.

7. Employee Experience and Engagement Platforms

Purpose: Capture and analyse employee sentiment and engagement levels.
Key Data: Survey responses, feedback themes, engagement scores, participation metrics.
AI Relevance: Provides human sentiment data to feed culture, well-being, and attrition prediction models.

8. Workforce Scheduling and Time Systems

Purpose: Manage attendance, scheduling, and utilization of workforce capacity.
Key Data: Timesheets, shifts, leave balances, absenteeism, and overtime data.
AI Relevance: Critical for modeling operational efficiency and optimizing frontline workforce deployment.

9. HR Service Delivery and Case Management Systems

Purpose: Handle HR requests, service delivery, and employee support workflows.
Key Data: Case categories, resolution times, request types, and employee interactions.
AI Relevance: Provides process-level data for automation, chatbots, and service improvement analytics.

10. External and Augmented Data Sources

Purpose: Integrate external labor market intelligence and skills benchmarks.
Key Data: Skill supply-demand trends, salary benchmarks, peer hiring data, and job taxonomies.
AI Relevance: Enables contextual calibration of internal data models and competitive benchmarking.

11. Generative Content Creation Layer – AI Models for HR Knowledge and Drafting

This emerging layer focuses on utilizing generative AI to generate initial drafts of HR documents, job descriptions, interview guides, and talent summaries. It leverages structured data across other layers and utilizes large language models to expedite the creation of HR content.

Examples of data assets:

- Job architecture repositories
- Competency libraries and skill dictionaries
- Past job descriptions, interview templates, and assessment rubrics

Primary function: Enable scalable, consistent, and data-informed HR content creation using generative AI while maintaining governance and review by HR experts.

Use cases that HR Tech Systems Power

These technologies power several use cases. Reporting and Analytics, Employee Experience, SWP, Skills Management, Succession Planning, HR Service Delivery, and so on

Data Operating Model

The **data operating model** for each HR system is critical. Many organizations purchased a system years ago and gradually built custom processes and integrations around it. Over time, this orchestration defines how data flows, who owns it, and how insights are generated. **However, just owning a system allows additional data enhancements to flow in easily**. If the data structure, workflows, and governance aren't aligned with newer enhancements or API/MCP extensions, you may be unable to take advantage of updates or emerging AI features. Ensuring each HR system has a clearly defined data model—covering inputs, outputs, and inter-system relationships—helps maintain scalability, seamless integration, and readiness for future innovations.

Example: Data Operating Model Focused on Core HRIS System

For example, you might have an HRIS system that can store and manage multiple categories of employee data — such as personal details, job information, compensation, and performance records.

Categories	Meaning	Example
Job Family	A logical grouping of related job profiles that share a similar nature of work or discipline. Used for career paths, benchmarking, and reporting.	*Finance*, *Human Resources*, *Engineering*, *Sales*
Job Profile	The **blueprint** of a role defines responsibilities, qualifications, and skills. Multiple positions can share one job profile.	*Financial Analyst II* under the *Finance* family
Position	The **specific seat** in the organization that uses a job profile and belongs to a department or manager. It can be vacant or filled.	*Position ID 000245 – Financial Analyst II, Dallas, reporting to the CFO*
Worker	The **person** assigned to a position can change over time as people are hired, transferred, or leave.	*Jane Doe is currently occupying Position 000245*

There could be more categories, but for our example, we have restricted ourselves to these sets. The issue is that you may not be using these categories of data as intended by the system. You may use Job Profile as a catch-all, and you may not be diligently updating position details. Now, when you have obtained additional data to enhance the Job Profile (for example), it will fall short. If, for example, you are using the Finance Job profile as a catch-all job profile for all Finance jobs, getting granular skills across Finance Analyst, Treasurer, Controller, etc., will not be valuable for you, as you cannot store the same.

Core Data Operating decisions that HR must lock down

1. **Job architecture (the catalog)**
   - Job family groups → job families → **Job Profiles** (with levels/grades).
   - Standardized naming and **unique codes** (e.g., ENG-SWE-III), separate from "business titles."
   - On each profile: FLSA/exemption, EEO category, union code (if applicable), grade profile, typical schedule, remote/on-site flag, required certifications, **skills taxonomy** mapping.
   - Keep it global first with explicit local variations via profile attributes or local profiles if truly needed.

2. **Staffing model & position management**
   - Decide where to use **Position Management** vs **Job Management**. For most enterprises, use Position Management for permanent/headcount-controlled populations; Job Management only where seats are fluid and controls are looser.

- Define **position restrictions**: job profile, time type, worker type, location, default comp plan, default costing/worktags, minimum qualifications.

3. **Organizational model**
   - Clean **supervisory organizations**; rules for manager changes and reorgs.
   - Align **Company**, **Cost Center**, **Location**, **Worktags** (Program/Project/Fund) to Finance.
   - Define rules for **matrix** structures (e.g., additional managers vs. role-based relationships).

4. **Compensation architecture**
   - Grade structures and **grade profiles** with currency/locale variations.
   - Default compensation plans at the job profile or position as needed.
   - Clear policy for **business title** vs **job profile title** (analytics use the profile, not the business title).

5. **Worker types & employment models**
   - Employee vs. contingent worker design; how vendors are represented; concurrent jobs; global employment.
   - Standard time types (full-time, part-time) and **work schedules** for accruals and time rules.

6. **Effective-dating discipline**
   - Always use **effective dates**; avoid corrections where a retro change will do.
   - Standard rules for back-dating (window, approvers, and downstream notifications).

Closing Thoughts

Create a cross-functional team comprising SWP, Analytics, and Talent Acquisition to audit and document the current operating model across systems, ensuring the organization can fully leverage data integrations and emerging technologies.