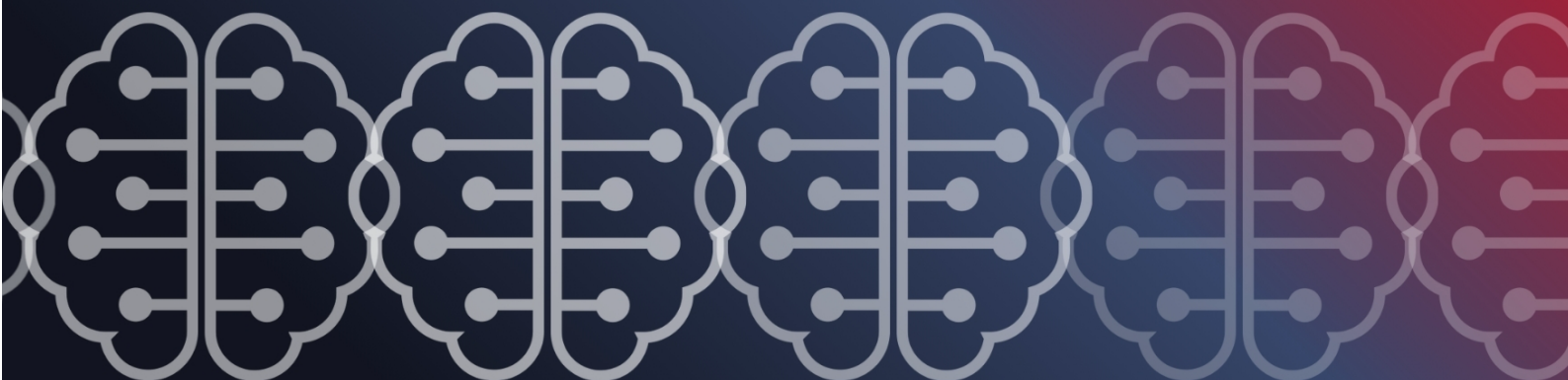




# ADVANCING AI AND ML IN LIFE SCIENCES R&D

From Model Innovation to  
Organizational Readiness



# Contents

Contents .....	1
Foreword: The Next Phase of AI in Life Sciences .....	3
FAIR by Design: From Principle to Practice .....	3
From Pilot Success to Decision Impact.....	4
Ownership and Structural Alignment.....	4
Technology Churn and Strategic Risk.....	4
Pre-Competitive Foundations.....	5
2026 and Beyond .....	5
Executive Summary.....	6
Survey Findings .....	6
Insights from Follow-Up Interviews .....	6
Organizational and Governance Constraints .....	6
Technology Dynamics and Adoption Capacity.....	7
Process Design and Workflow Readiness .....	7
Implications for Pre-Competitive Collaboration .....	7
1. Context and Purpose .....	8
1.1 Survey Demographics.....	9
1.2 Interviewee Demographics.....	10
1.3 Strategic Opportunity Gaps.....	10
1.4 Industry Readiness Landscape.....	11
1.5 Top Pre-Competitive Collaboration Opportunities.....	12
3. Key Themes.....	13
3.1 Validation and Governance Standards .....	13
3.2 Data Quality and Metadata Standards .....	14
3.3 Value Realization and ROI Measurement.....	14
3.4 Organizational Structure and Ownership Models.....	15
3.5 Technology Selection and Model Strategy.....	16
3.6 Culture, Literacy, and Enablement.....	17
3.7 Collaboration Friction and IP.....	17
3.8 Agentic AI: From Experimental Capability to Governance Challenge.....	18
High Collaboration Potential, Low Maturity .....	19
High Collaboration Potential, Medium Maturity .....	19
Medium Collaboration Potential, Low Maturity .....	19
Medium Collaboration Potential, Medium Maturity.....	19
5.1 Governance and Validation Sandbox.....	20
5.2 Data Quality and Metadata Consortium .....	20
5.3 AI Literacy and Enablement Network.....	20
5.4 Rapid Collaboration Blueprint .....	20

5.5 Agentic AI Governance Framework .....	20
5.6 Value and ROI Framework.....	20
5.7 Organizational Operating Models.....	20
Appendix A: Methods.....	21
Survey Phase.....	21
Survey Design .....	21
Survey Sample .....	21
Survey Outputs Used in This Report.....	21
Interview Phase.....	22
Data Collection .....	22
Interview Sampling and Topic Selection.....	22
Qualitative Coding Framework.....	22
Analytic Procedure and Integration .....	22
Appendix B: Glossary .....	23
Appendix C: Interview Analysis .....	26
Interview Methodology Summary.....	26

## Foreword: The Next Phase of AI in Life Sciences

The findings presented in this report reflect the state of thinking across the life sciences sector during the second half of 2025. In a field where technological cycles increasingly compress into months rather than years, it is important to acknowledge the temporal context. The pace of AI advancement continues to accelerate. Organizational capability evolves more gradually. The widening distance between these two curves is central to the conclusions of this report.

The global survey of 54 senior leaders across pharmaceutical, biotech, CRO, and technology organizations, followed by in-depth interviews with 9 executives, reveals a consistent pattern: ***internal demand for AI and machine learning is high, perceived strategic importance is high, yet organizational maturity remains comparatively low.***

The demand–maturity gap is particularly pronounced in multimodal AI, validation frameworks, governance practices, and data standards (see Figure 3). These are foundational capabilities required to embed AI into core R&D workflows.

This report integrates findings from both phases of the study. The survey provides the quantitative landscape: where demand is highest, where maturity lags, and where collaboration potential is strongest. The interviews add interpretive depth: *why* adoption stalls, *what* organizational factors shape outcomes, and *how* pre-competitive collaboration might accelerate progress. Where survey and interview findings align, confidence is high. Where they diverge, the contrasts are instructive and are flagged throughout.

External signals reinforce this picture. Industry analyses from McKinsey, BCG, and Deloitte continue to show strong executive prioritization of AI-enabled productivity in drug discovery and development. Regulatory authorities, including the FDA and EMA, have intensified focus on lifecycle management, transparency, and validation requirements for AI in regulated environments. The strategic intent is clear. The constraint is operational readiness.

Across interviews, limited impact is rarely attributed to algorithmic performance. Instead, executives describe uncertainty around value definition, weak linkage between pilots and decision processes, unclear ownership models, and insufficient data quality and metadata standards. As one interviewee noted, “The moat is not in the models, it is in the data.” This observation is consistent with broader empirical findings in biomedical AI: reproducibility, representativeness, and lineage transparency remain the primary determinants of trusted deployment.

### FAIR by Design: From Principle to Practice

Although the FAIR principles have been discussed across the industry for many years, implementation remains uneven. Many organizations still treat FAIRification as a downstream clean-up activity rather than an upstream design requirement. Ensuring that data are Findable, Accessible, Interoperable, and Reusable at the point of generation has not yet become standard practice.

Creating FAIR data by design requires structured metadata capture at the time of experiment execution, harmonized ontologies, consistent annotation standards, electronic lab notebook integration, imaging ingestion pipelines with traceable preprocessing steps,

and transparent lineage documentation. Without these elements embedded into laboratory and clinical workflows, model performance gains remain constrained by inconsistent and poorly contextualized inputs.

The evidence from this study reinforces a recurring conclusion: sustainable AI impact depends on operationalizing FAIR at the moment of data production. Retrofitting FAIR compliance after datasets are fragmented and undocumented introduces cost, delay, and avoidable bias. Embedding FAIR as a default design principle is not an aspirational objective; it is a prerequisite for scalable AI.

## From Pilot Success to Decision Impact

The AI Value Leakage Funnel (Figure 1) illustrates a recurring pattern described qualitatively across interviews: successful pilots frequently stall before influencing decisions or generating measurable ROI. Technical feasibility is demonstrated. Deployment remains partial. Decision impact is limited. Measured value is weak.

This pattern aligns with established research in digital transformation. The final stages of workflow integration and behavioural adoption account for the majority of unrealized value. In life sciences, the “last five percent problem,” as one interviewee described it, results in scientists re-checking outputs manually, negating efficiency gains.

AI adoption accelerates when value is framed in decision terms rather than performance metrics. Model accuracy improvements are insufficient unless linked to cycle time reduction, experimental throughput, risk mitigation, or confidence in hypothesis prioritization. Without a shared value framework, AI remains an interesting capability rather than a trusted instrument.

## Ownership and Structural Alignment

A second structural theme is ownership fragmentation. Figure 4 summarizes how responsibility for AI initiatives is distributed across R&D, digital, IT, and data science functions. End-to-end accountability for value realization is often unclear. Centralized digital teams may lack domain authority. Scientific groups may lack technical depth. Diffused responsibility slows adoption and complicates benchmarking.

Successful integration requires explicit operating models defining decision rights, platform ownership, domain accountability, and the emerging role of the AI Translator. The Translator function, highlighted in multiple interviews, bridges scientific expertise and machine learning practice, reducing misinterpretation and inappropriate application.

Organizational alignment is not an ancillary consideration. It is a determining factor of impact.

## Technology Churn and Strategic Risk

Interviewees consistently describe the pace of technological innovation as exceeding organizational evaluation capacity. Figure 5 contrasts rapid increases in available models, tools, and platforms with slower internal validation and deployment cycles. Technology churn is increasingly perceived as a strategic risk.

In regulated scientific environments, premature adoption carries measurable cost. Model lifecycle management, validation lag, and vendor dependency introduce long-term governance implications. Executives therefore seek shared evaluation benchmarks and structured model comparison frameworks rather than rapid tool proliferation.

The constraint is no longer access to models. It is the ability to assess fitness for purpose under scientific and regulatory standards.

## Pre-Competitive Foundations

The Priority Map (Figure 6) identifies areas with high collaboration potential and low organizational maturity. These include validation frameworks, data quality and metadata standards, ROI measurement, and ownership models. These domains represent structural enablers rather than competitive differentiators.

Pre-competitive collaboration in these areas reduces duplicated effort and shared risk while preserving differentiation at the level of proprietary datasets, therapeutic insight, and pipeline strategy. Shared validation sandboxes, interoperable metadata standards, standardized contracting templates, and common value measurement frameworks can accelerate adoption without compromising competitive advantage.

## 2026 and Beyond

Internal demand remains strong across nearly all surveyed domains. Foundation models, multimodal AI, and agent-based systems continue to attract attention. Enthusiasm is evident. Scepticism is disciplined. Senior R&D leaders require validation, transparency, and measurable impact before scaling deployment.

Whether 2026 becomes a pivotal year for AI integration into core R&D processes will depend less on model novelty and more on structural execution. Validation standards, FAIR-by-design data practices, ownership clarity, governance frameworks, and value measurement will determine the trajectory.

AI in life sciences has moved beyond theoretical promise. The next phase is operational discipline. Execution, grounded in scientific rigor and collaborative infrastructure, will define sustainable impact.

*This report is offered as a contribution to that execution.*

# Executive Summary

## Survey Findings

A global survey of 54 life sciences research and development leaders examined AI and ML adoption, internal demand, and collaboration readiness across 19 technology domains. Respondents rated each domain on three dimensions – organizational maturity, internal demand, and collaboration potential – on a 1–5 scale, enabling quantitative comparison across the sector. The survey revealed a consistent pattern:

- Internal demand and perceived strategic importance are high
- Organizational maturity remains comparatively low
- The demand–maturity gap is most pronounced in multimodal AI (0.92-point gap), real-world data integration (0.77), and reinforcement learning (0.74)

Foundation Models led all dimensions with the highest maturity (3.30), demand (3.89), and collaboration scores (3.51), while areas such as shared validation frameworks and cross-pharma governance showed high demand but significantly lower maturity. These patterns indicate persistent structural and capability constraints rather than isolated technical gaps.

## Insights from Follow-Up Interviews

To interpret the survey findings, a qualitative phase comprising nine semi-structured interviews with senior leaders was conducted. The interviews explored why areas of high demand frequently fail to translate into sustained impact.

Across interviews, a common set of explanations emerged:

- Limited AI impact is less frequently attributed to model performance
- Greater emphasis is placed on unclear value definitions and weak outcome measurement
- Many initiatives stall after pilots without influencing decisions or workflows

## Organizational and Governance Constraints

Interviewees consistently highlighted organizational factors shaping AI outcomes:

- Responsibility for AI initiatives is distributed across research, digital, IT, and data science functions
- End-to-end ownership of value realization is often unclear
- Fragmented ownership contributes to slower adoption and inconsistent benchmarking practices

These dynamics make it difficult to translate technical outputs into trusted scientific or operational decisions.

## Technology Dynamics and Adoption Capacity

Interviewees also described growing uncertainty driven by rapid advances in AI technologies:

- Strong demand exists for foundation models and multimodal AI
- Organizational capacity to evaluate, validate, and sustain adoption lags behind innovation pace
- Technology churn is increasingly perceived as a strategic risk rather than a technical issue

This misalignment complicates technology selection and long-term investment decisions.

## Process Design and Workflow Readiness

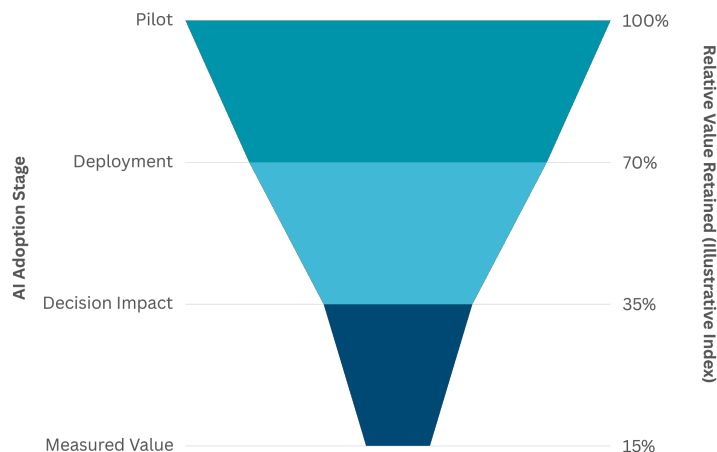
Finally, interviews emphasized the importance of workflow design:

- Applying AI to existing processes without redesign limits effectiveness
- Where workflows lack standardization, reproducibility, or decision clarity, AI amplifies inefficiency
- Sustainable impact requires aligning AI deployment with process maturity

## Implications for Pre-Competitive Collaboration

Taken together, the findings point toward a pre-competitive agenda that extends beyond technical enablement. Progress will depend on shared approaches to:

- Validation and reliability frameworks
- Data quality and metadata standards
- Value and ROI measurement
- Organizational operating models for AI
- Technology evaluation and selection guidance
- Process redesign and workflow standardization
- Responsible integration of emerging agent-based systems



**Figure 1: AI Value Leakage Funnel.** The funnel illustrates the relative drop-off observed as AI initiatives progress from pilots to measurable value. Stages and proportions are derived from recurring interview patterns describing successful pilots, partial deployment, limited decision impact, and weak ROI measurement. Values are illustrative and represent direction and magnitude, not absolute percentages.

## 1. Context and Purpose

This white paper builds on a global survey and follow-up interviews examining AI and ML adoption, internal demand, and collaboration readiness in life sciences research and development. The survey identified multiple domains in which internal demand and collaboration potential are high, while organizational maturity remains low, indicating persistent structural and capability gaps across the sector.

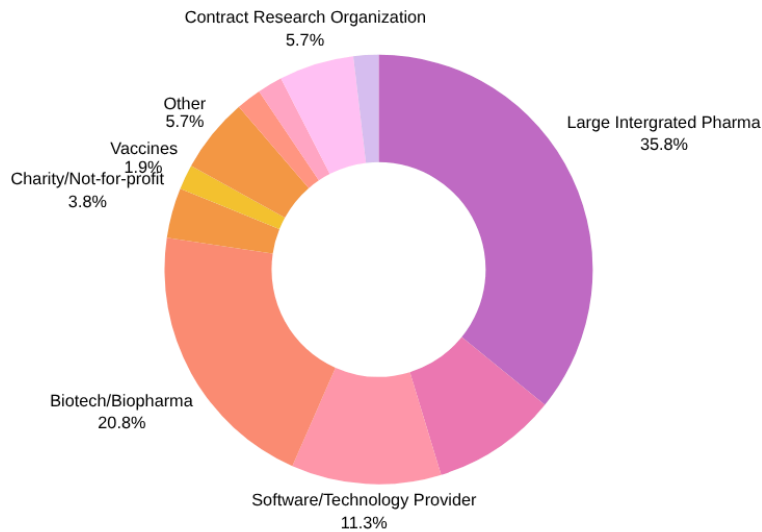
To contextualize and interpret these findings, a qualitative phase comprising nine semi-structured interviews was conducted. The objectives of this phase were to:

- Explain the underlying rationale behind survey responses and scores
- Identify recurring blockers and enablers influencing AI adoption
- Distinguish systemic, cross-industry constraints from organization-specific challenges
- Surface organizational, operational, and economic tensions not fully captured through survey data
- Map areas where pre-competitive collaboration could reduce duplicated effort and shared risk

Interview participants represented a broad cross-section of life sciences research and development functions, including discovery science, clinical development, digital and data science, laboratory and automation functions, enterprise governance, and organizational learning. This diversity enabled triangulation of perspectives across scientific, technical, and operational domains.

## 1.1 Survey Demographics

The survey was conducted in August 2025 via the Pistoia Alliance network and received 54 completed responses from 1,724 views. Respondents spanned 14 countries, with strong transatlantic representation providing an excellent foundation for global collaboration initiatives.



**Figure 2: Survey Demographics.** Insights from 54 pharmaceutical industry leaders across Europe (52%), North America (33%), and Asia Pacific (11%). Respondents represent senior leadership from R&D IT/Digital teams, AI/ML specialists, and C-suite executives across large integrated pharma (35%), biotech (21%), IT/digital service providers (21%), plus CROs and technology firms.

The balanced representation between R&D scientists and IT/digital professionals indicates strong cross-functional interest in AI applications:

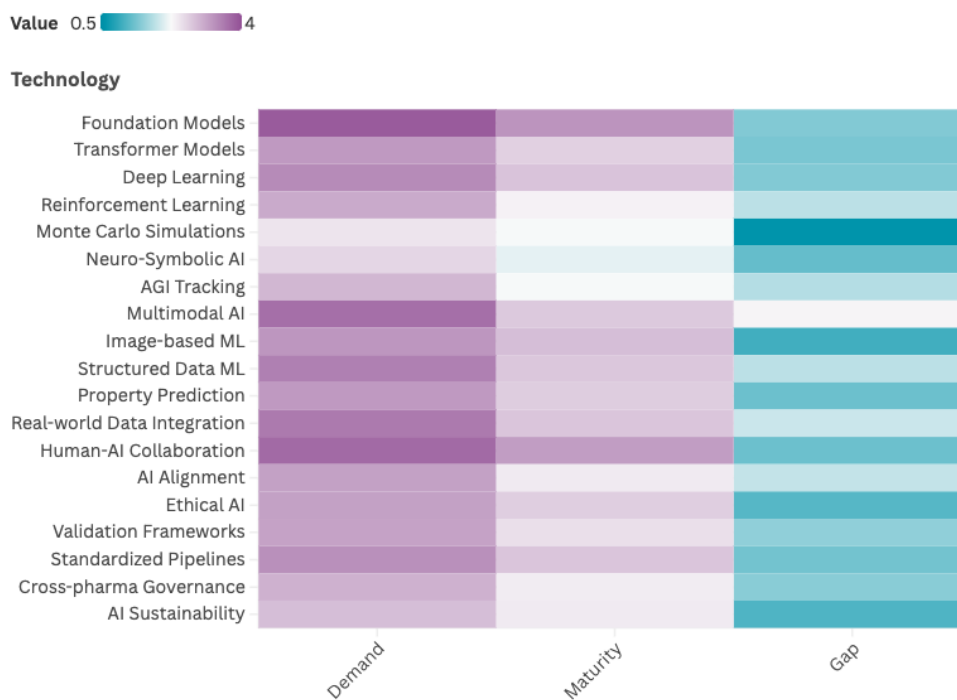
Job Role	Count	Percentage
Research/Discovery	12	22.2%
R&D IT/Digital	12	22.2%
Other Biopharma IT/Digital	6	11.1%
Pharmaceutical Development/CMC	5	9.3%
R&D Data	4	7.4%
Other	15	27.8%

Notably, 38.9% of respondents were individual contributors and 25.9% led teams of 6–20 people, suggesting grassroots interest in AI collaboration alongside institutional support from leadership.

## 1.2 Interviewee Demographics

Dimension	Description
Number of interviews	9
Seniority	Senior leadership, directors, and heads of function
Primary roles	AI and digital leadership, R&D leadership, scientific operations, enterprise governance
Decision scope	Strategy definition, portfolio prioritization, oversight of AI deployment
Functional representation	Discovery science, clinical development, digital and data science, laboratory automation, enterprise architecture and governance, organizational learning
Geographic scope	Multiple regions represented

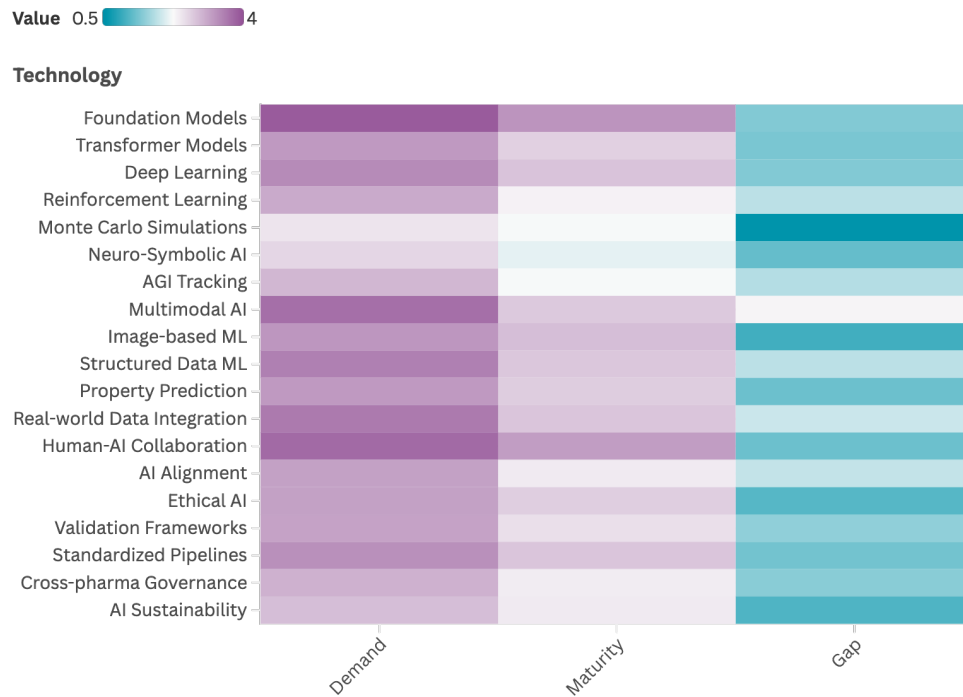
## 1.3 Strategic Opportunity Gaps



**Figure 3: Strategic Opportunity Gaps.** Identified from the survey. Multimodal AI shows the largest gap (0.92 points) – high demand but organizations aren’t ready. Real-world data integration (0.77 gap) indicates urgent need for infrastructure investment. Governance concerns outpace implementation readiness. Reinforcement Learning (0.74 gap) suggests experimental optimization remains aspirational.

## 1.4 Industry Readiness Landscape

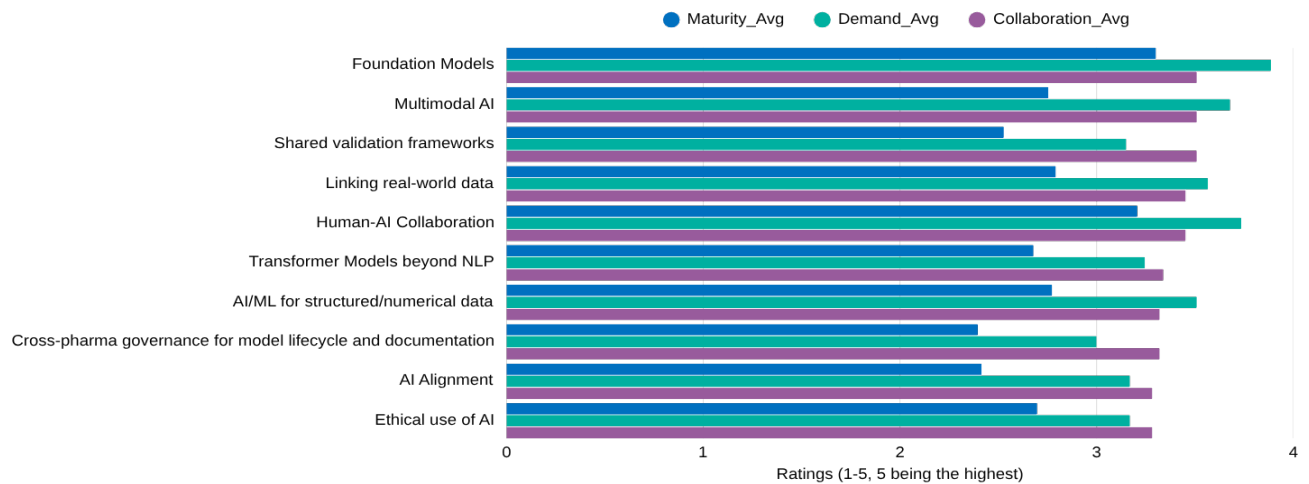
The survey assessed all 19 AI/ML technology domains across demand, maturity, and the resulting strategic gap. The heatmap below provides a visual overview of the full landscape, highlighting where demand significantly outpaces organizational readiness.



**Figure 3a: Industry Readiness Heatmap.** Demand, maturity, and strategic gap across all 19 surveyed AI/ML technology domains. Darker purple indicates higher scores for demand and maturity; darker teal indicates larger gaps. Foundation Models show the highest demand and maturity, while emerging technologies such as Neuro-Symbolic AI and Monte Carlo Simulations remain exploratory with lower scores across all dimensions.

## 1.5 Top Pre-Competitive Collaboration Opportunities

Ranking the 19 domains by combined scores across all three dimensions identifies the areas with the greatest potential for shared industry action. Foundation Models led all dimensions, while Multimodal AI showed strong demand and collaboration potential despite notably lower maturity.



**Figure 3b: Top 10 Pre-Competitive Collaboration Opportunities.** Foundation Models lead across all dimensions (3.30 maturity, 3.89 demand, 3.51 collaboration). Human-AI Collaboration scores highest for practical adoption (3.21 maturity, 3.74 demand). Multimodal AI shows significant potential (3.68 demand, 3.51 collaboration) despite lower maturity (2.75). Real-world data integration is a critical priority (3.57 demand, 3.45 collaboration).

## 2. Evidence Clusters

The following table summarizes the key evidence clusters identified through the combined survey and interview analysis. Each cluster represents a thematic area where consistent patterns emerged across multiple organizations and interviewees.

Cluster	Subtheme	Representative Quote
Trust and Governance	Validation standards and safety	“The moat is not in the models, it is in the data.”
Trust and Governance	Enterprise guardrails and review	“We want innovation, but it must be regulated and safe.”
Data Quality and Standards	Selection bias and modality gaps	“Models built on oncology do not work well for other diseases.”
Data Quality and Standards	Metadata and reproducibility	“Use electronic lab books and metadata standards so others can reproduce.”
Value Realization	ROI and decision impact	“The last five percent problem means we still check everything.”

Organizational Alignment	Ownership and silos	<i>“Everyone supports AI, but no one owns the outcome.”</i>
Technology Strategy	Tool sprawl and model churn	<i>“The landscape changes faster than we can evaluate tools.”</i>
Culture and Enablement	Agentic experimentation	<i>“Duplication is fine because it helps people learn.”</i>
Collaboration Friction	Legal and procurement delays	<i>“Months to finalize contracts and momentum is gone.”</i>
Tech Readiness	Biomedical limitations of LLMs	<i>“Foundational models do not work for many modalities in pharma.”</i>

### 3. Key Themes

The following themes are ordered by the frequency of their occurrence across interviews. Each theme includes a priority assessment, collaboration potential rating, maturity level, scientific observation, and recommended opportunity for pre-competitive action. Where relevant, quantitative findings from the survey are presented alongside the qualitative interview evidence to show alignment or instructive contrasts between the two data sources.

#### 3.1 Validation and Governance Standards

**Priority:** Very High

**Collaboration Potential:** Very High

**Maturity:** Moderate to High internally; Low cross-industry alignment

Validation frameworks vary widely between organizations. Participants emphasize the need for reproducibility, explainability, traceability, and bias evaluation. Governance practices are most effective when they enable experimentation rather than restrict it.

**Survey evidence:** Shared validation frameworks ranked among the top 10 pre-competitive collaboration opportunities in the survey (3.30 demand, 3.15 collaboration), and “Validation Frameworks & Bias Mitigation” was the third-highest collaboration priority identified by survey respondents. Cross-pharma governance for model lifecycle and documentation also scored highly, confirming the interview emphasis on this theme.

**Survey–interview contrast:** Survey respondents elevated “Ethical AI in Patient-Facing Applications” as the single highest collaboration priority, but this specific framing did not surface as prominently in interviews. Interviewees focused more on practical validation and trust-building than on patient-facing ethics, suggesting that the broader population places greater emphasis on external-facing responsibility, while senior operational leaders prioritize internal scientific trust.

**Scientific Observation:** Trust rather than benchmark performance is now the limiting factor for AI adoption. Regulatory and scientific stakeholders require transparent reasoning, defined failure modes, and consistent documentation.

**Opportunity:** A shared validation sandbox focusing on specific scientific domains such as imaging or mechanistic modelling. Outputs may include validation protocols, bias testing methods, audit checklists, and model documentation templates.

## 3.2 Data Quality and Metadata Standards

**Priority:** Very High

**Collaboration Potential:** Very High

**Maturity:** Low

Data variability, missing metadata, inconsistent lineage, and selection bias are the primary reasons for poor model generalizability. Imaging data suffers from inconsistent ingestion pipelines and annotation methods.

**Survey evidence:** “Data Standardization & Ontologies” was the second-highest collaboration priority identified by survey respondents, with strong interest in common data models, interoperability frameworks, and semantic data integration. The survey also flagged FAIR principles implementation and multi-modal data quality assessment as key quality and governance focus areas. This directly corroborates the interview finding that data, not models, is the binding constraint.

**Scientific Observation:** Most performance failures originate in the data, not the model architecture. Reliable modelling requires minimal metadata standards, consistent experimental annotation, and unified data engineering pipelines.

**Opportunity:** A Data Quality and Metadata Consortium that delivers interoperable schemas, experimental documentation standards, imaging ingestion patterns, and dataset data cards.

## 3.3 Value Realization and ROI Measurement

**Priority:** Very High

**Collaboration Potential:** High

**Maturity:** Low

Despite widespread experimentation, organizations struggle to connect AI outputs to measurable scientific or business value. Pilots often succeed technically but fail to influence decisions, reduce cycle times, or justify continued investment.

**Survey–interview contrast:** ROI measurement and value realization did not emerge as a distinct collaboration priority in the survey’s structured responses, which focused more on technology domains and ethics. This theme was surfaced almost exclusively through the interview phase, where senior leaders described it as a critical bottleneck. The contrast suggests that value measurement is more visible as a problem to those responsible for strategic deployment decisions than to the broader practitioner population.

**Scientific Observation:** AI adoption stalls when value is not framed in decision terms. Performance metrics alone are insufficient without linkage to experimental throughput, error reduction, or time-to-insight.

**Opportunity:** A shared value framework focused on decision quality, cycle time, and scientific confidence, including templates for value hypotheses, success criteria, and reporting of negative results.

### 3.4 Organizational Structure and Ownership Models

**Priority:** High

**Collaboration Potential:** High

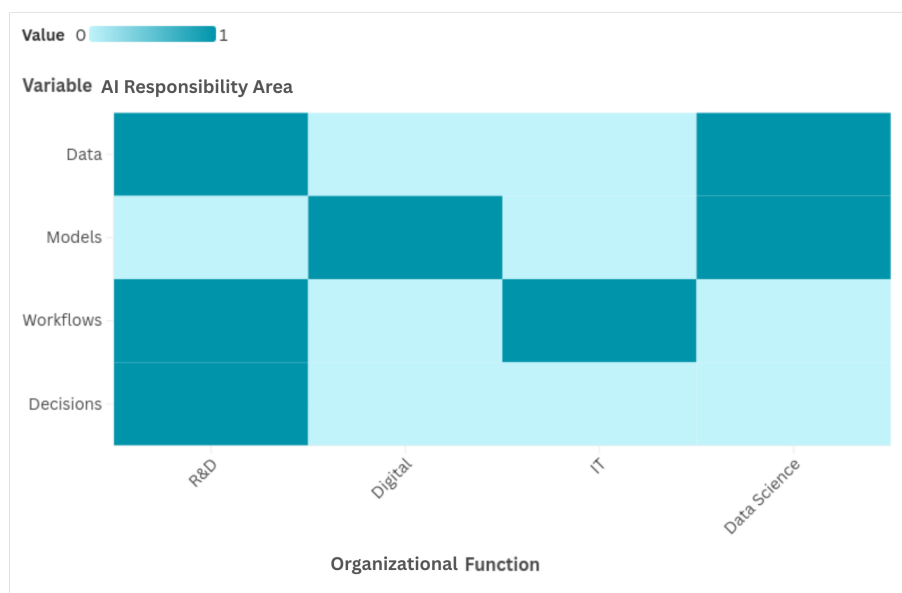
**Maturity:** Low to Medium

AI initiatives frequently span R&D, digital, IT, and data science, but ownership of outcomes is unclear. Centralized platforms often lack domain authority, while local teams lack technical support.

**Survey evidence:** The survey’s respondent profile provides indirect evidence for this theme: 38.9% were individual contributors and the remainder were distributed across team sizes from 5 to 500+, reflecting the fragmented organizational landscape that interviewees described. The balanced representation of R&D scientists (22%) and IT/digital professionals (22%) underscores the cross-functional nature of AI ownership challenges.

**Scientific Observation:** AI fails when accountability for outcomes is diffused. Successful adoption requires explicit decision rights and operating models.

**Opportunity:** Pre-competitive sharing of operating models that define ownership, decision rights, and roles such as platform owner, domain owner, and AI translator.



**Figure 4: Organizational Ownership Fragmentation.** This figure summarizes how responsibility for AI initiatives is distributed across organizational functions at different lifecycle stages. Positioning reflects recurring qualitative descriptions of ownership and hand-offs from interviews, highlighting areas where accountability is fragmented or unclear.

### 3.5 Technology Selection and Model Strategy

**Priority:** High

**Collaboration Potential:** Medium

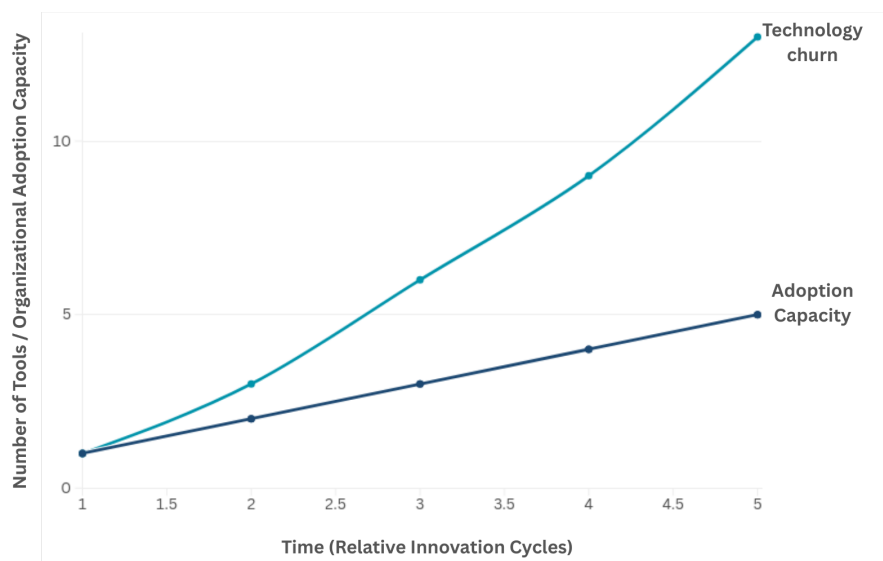
**Maturity:** Low

Rapid innovation in models and tools has created uncertainty. Organizations struggle with build versus buy decisions, foundation versus domain-specific models, and vendor evaluation.

**Survey evidence:** The survey quantifies this challenge precisely. Foundation Models led all 19 domains with the highest maturity (3.30), demand (3.89), and collaboration scores (3.51). However, Multimodal AI showed the largest demand–maturity gap of any domain (0.92 points: 3.68 demand versus 2.75 maturity), confirming interview descriptions of a technology landscape evolving faster than organizations can evaluate. The heatmap (Figure 3a) further illustrates how emerging technologies such as Neuro-Symbolic AI and AGI Tracking remain exploratory, with low scores across all three dimensions.

**Scientific Observation:** The pace of technological change exceeds organizations’ ability to evaluate fitness for purpose.

**Opportunity:** Shared evaluation benchmarks, model comparison frameworks, and guidance on technology lifecycle management.



**Figure 5: Technology Advancement Versus Organizational Adoption Capacity.** This figure contrasts the pace of AI technology advancement with organizational adoption capacity. Positioning is derived from interview evidence describing rapid increases in available models, tools, and platforms relative to slower internal evaluation, validation, and deployment cycles. Values are illustrative and represent relative divergence rather than measured counts or timelines.

### 3.6 Culture, Literacy, and Enablement

**Priority:** High

**Collaboration Potential:** High

**Maturity:** Medium

AI literacy gaps exist across scientists, clinicians, digital teams, and leadership. Adoption is often driven bottom-up by enthusiastic individuals. The missing capability across organizations is the AI Translator who can bridge domain science and ML practice.

**Survey evidence:** *Human-AI Collaboration scored highest of all 19 domains for practical adoption readiness (3.21 maturity) while maintaining strong demand (3.74) and collaboration scores (3.51). The survey also identified “Human-AI Collaboration Models” as the fifth-highest collaboration priority, with specific interest in AI tools that augment rather than replace human expertise, user experience standards, and training frameworks. This aligns closely with the interview emphasis on literacy, but the interviews added depth by identifying the specific AI Translator role as a critical missing capability not captured in the survey.*

**Scientific Observation:** Democratization is effective only when combined with governance and safety boundaries. Scientists need clarity on what is safe to attempt, how to interpret model outputs, and when to escalate to specialists.

**Opportunity:** An AI Literacy and Enablement Network that develops role-specific curricula, shared vocabulary, best practices for scientific prompt design, and patterns for safe experimentation.

### 3.7 Collaboration Friction and IP

**Priority:** Medium to High

**Collaboration Potential:** High

**Maturity:** Medium

Contracting cycles often take several months. Organizations over-protect low-value assets and struggle with publisher restrictions that prevent the use of generative models for literature analysis.

**Survey–interview contrast:** *Contracting friction and IP barriers were surfaced almost exclusively through the interview phase. The survey’s collaboration priorities focused on technical domains (validation, data standards, multimodal AI) and ethical frameworks, but did not directly capture the procedural and legal blockers that interviewees described as significant impediments to collaboration. This is a notable gap: the operational friction of “months to finalize contracts” appears invisible in structured survey data but was described consistently by senior leaders as a critical bottleneck to pre-competitive action.*

**Scientific Observation:** The barriers to collaboration are procedural and institutional rather than scientific. Most organizations agree on the value of collaboration but lack efficient ways to operationalize it.

**Opportunity:** A Rapid Collaboration Blueprint with standard IP provisions, template data-sharing agreements, streamlined security assessments, and a four to six week target timeline for initiating collaboration.

### 3.8 Agentic AI: From Experimental Capability to Governance Challenge

**Priority:** Emerging

**Collaboration Potential:** Medium

**Maturity:** Low to Medium

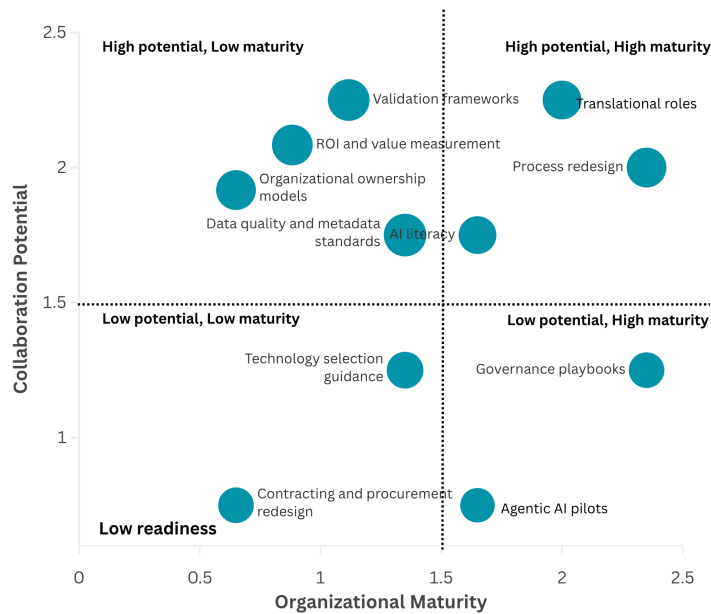
Agent-based systems are increasingly used for protocol drafting, laboratory automation, multimodal analysis, and scientific reasoning. Reliability and control remain major concerns.

**Survey evidence:** *While agentic AI was not one of the 19 individually scored technology domains, the survey captured it as an emerging focus area under infrastructure and integration, with respondents noting interest in “agentic ecosystem frameworks,” “API usage orchestration,” and “lab operations orchestration.” The presentation of survey results at the Pistoia Alliance conference in Boston positioned an “Agentic AI Framework Alliance” as one of the recommended technology development consortiums, suggesting growing consensus around the need for shared governance in this space.*

**Scientific Observation:** Agentic systems can reshape computational workflows but require strong human oversight, logging, and rollback capabilities.

**Opportunity:** A reference architecture for agentic systems that defines allowed actions, observability requirements, integration patterns with laboratory devices, and human-in-the-loop checkpoints.

## 4. Priority Map



**Figure 6: Priority Map for Pre-Competitive AI Collaboration.** Themes are positioned based on organizational maturity (x-axis) and collaboration potential (y-axis), as classified from survey responses and follow-up interviews. Quadrants indicate categorical positioning rather than continuous measurement. Bubble size reflects relative demand intensity based on interview frequency and emphasis.

### High Collaboration Potential, Low Maturity

- Validation frameworks
- Data quality and metadata standards
- ROI and value measurement
- Organizational ownership models

### High Collaboration Potential, Medium Maturity

- AI literacy
- Translation roles
- Process redesign

### Medium Collaboration Potential, Low Maturity

- Technology selection guidance
- Contracting and procurement redesign

### Medium Collaboration Potential, Medium Maturity

- Agentic AI pilots
- Governance playbooks

## 5. Strategic Recommendations

### 5.1 Governance and Validation Sandbox

Develop shared validation protocols, reproducibility checklists, bias testing methods, and domain-specific documentation standards.

### 5.2 Data Quality and Metadata Consortium

Deliver minimal metadata requirements, experimental documentation standards, imaging data QC templates, and data cards.

### 5.3 AI Literacy and Enablement Network

Create role-based learning materials, shared vocabulary, safe experimentation guidelines, and pilots for embedded AI translators.

### 5.4 Rapid Collaboration Blueprint

Provide standard contracting templates, IP frameworks, security and privacy models, and predictable timelines for collaboration onboarding.

### 5.5 Agentic AI Governance Framework

Define agent action boundaries, audit and rollback mechanisms, laboratory system integration patterns, and evaluation criteria for production deployment.

### 5.6 Value and ROI Framework

Develop shared metrics linking AI to decision quality, cycle time, and scientific outcomes.

### 5.7 Organizational Operating Models

Publish reference models for ownership, accountability, and governance across R&D, digital, and data science.

## 6. Closing Statement

Progress in AI for life sciences depends less on model novelty and more on validation, data quality, organizational alignment, and value realization. Pre-competitive collaboration across these foundations offers a practical path to scale AI while preserving scientific rigor, patient safety, and long-term impact.

## Appendix A: Methods

This study used a sequential mixed-methods design consisting of an industry survey followed by semi-structured interviews. The survey established the scale and prioritization of AI and ML challenges across the sector. The interviews were conducted as follow-up to explain the reasoning behind survey responses and to surface implementation constraints not captured in structured survey data.

### Survey Phase

#### Survey Design

A global survey was conducted in August 2025 to assess AI and ML adoption and collaboration readiness in life sciences research and development. The survey received 54 completed responses from 1,724 views (3.1% completion rate), distributed via the Pistoia Alliance network. Respondents evaluated 19 AI and ML technology domains grouped into four categories – foundational technologies and approaches, data modalities and multi-source integration, human-centric and responsible AI, and infrastructure and governance – across three dimensions:

- **Organizational Maturity**, ranging from not started to widely adopted (1–5)
- **Internal Demand**, ranging from no interest to high demand (1–5)
- **Collaboration Potential**, ranging from not suitable to very high potential (1–5)

A strategic gap score was computed as internal demand minus organizational maturity to highlight domains where demand outpaces readiness.

#### Survey Sample

The survey included 54 respondents across 14 countries spanning multiple sectors, including large pharmaceutical organizations (35.2%), biotech and biopharma (20.4%), software and technology providers (11.1%), IT/digital services (9.3%), contract research organizations (5.6%), and others (18.5%). Geographic distribution was concentrated in Western Europe (51.9%) and North America (50.0%), with additional representation from Northern Europe (38.9%), Southern Europe (18.5%), and other regions (27.8%).

#### Survey Outputs Used in This Report

Survey results were used to:

- Identify domains with high demand and high collaboration potential but low maturity
- Quantify demand–maturity gaps to support prioritization
- Define the set of topics selected for qualitative follow-up interviews
- Provide quantitative context for interview themes, including specific scores and rankings where relevant

## Interview Phase

### Data Collection

- Informed consent was obtained from all participants
- Interviews were approximately 30 to 40 minutes in duration
- Conversations were audio recorded and transcribed for analysis
- Interviewees were anonymized in the main report to enable candid discussion

### Interview Sampling and Topic Selection

Nine semi-structured interviews were conducted with senior leaders across life sciences R&D functions. Interview topics and probes were derived directly from survey domains showing high demand, high collaboration potential, and low or uneven maturity.

Participants represented discovery science, clinical development, digital and data science, laboratory and automation functions, enterprise governance, and organizational learning.

### Qualitative Coding Framework

Transcripts were coded using a structured framework encompassing five analytic dimensions:

- **Trust**, including validation practices, explainability, and risk perception
- **Friction**, including technical, organizational, legal, and process-related barriers
- **Delight**, capturing enabling conditions and successful patterns
- **Collaboration**, assessing feasibility, appetite, and perceived obstacles
- **Readiness**, reflecting maturity, resourcing, and organizational alignment

### Analytic Procedure and Integration

The analysis followed a multi-step synthesis process:

- Transcription review and tagging of timestamped quotes, emphasis, and observed behaviours
- Affinity mapping to cluster recurring themes across interviews
- Matrix mapping to align qualitative themes with survey dimensions of demand, maturity, and collaboration potential
- Cross-interview triangulation to test consistency and identify convergent findings

Survey results established the priority landscape at scale. Interview findings were used to interpret survey patterns, explain gaps between demand and maturity, and identify mechanisms limiting adoption and value realization. Where survey and interview findings converged, confidence in the finding was strengthened. Where they diverged, the contrasts were examined and reported as instructive tensions between the quantitative and qualitative evidence.

## Appendix B: Glossary

### **Adoption Capacity**

An organization's ability to evaluate, validate, deploy, and sustain AI systems beyond isolated pilots.

### **AI Alignment**

Practices ensuring that AI system behaviour remains consistent with human intent, domain constraints, ethical principles, and organizational objectives.

### **AI and ML Technologies**

Nineteen surveyed technology domains spanning foundational methods, data modalities, human-centric and responsible AI, and infrastructure and governance, representing the pharma AI innovation landscape.

### **AI Translator**

A bridging capability between domain science, data science, and operations that enables interpretation, appropriate use, and adoption of AI outputs.

### **Artificial General Intelligence**

A hypothetical class of AI systems with broad, human-level cognitive capabilities across domains, included in the survey for horizon scanning and perception tracking rather than near-term implementation.

### **Build vs Buy Tension**

The recurring decision challenge between internal development and external procurement of AI capabilities under rapid technology change.

### **Collaboration Potential**

Industry rating on a 1 to 5 scale from Not suitable to Very high potential, measuring suitability for cross-company collaboration on shared AI and ML standards, validation frameworks, and capabilities.

### **Cross-Organization Governance**

Policies, processes, and oversight mechanisms that manage AI systems across their lifecycle, including documentation, accountability, risk management, and compliance across organizational boundaries.

### **Data Quality**

The degree to which data are complete, consistent, well-annotated, and fit for use in modelling and decision making, including considerations of bias and representativeness.

### **Decision Impact**

The degree to which AI outputs directly influence scientific, clinical, or operational decisions.

### **Deep Learning**

Neural network-based methods used for pattern recognition and representation learning in complex, high-dimensional datasets.

**Ethical Use of AI**

Application of AI in ways that respect patient safety, fairness, privacy, and regulatory expectations, particularly in patient-facing or clinical contexts.

**Foundation Models**

Large pre-trained models, including large language models and other architectures, adapted to downstream tasks through fine-tuning or prompting.

**Human–AI Collaboration**

AI systems designed to support or augment human users rather than fully automate decisions or replace expert judgment.

**Image-Based Machine Learning**

Machine learning methods applied to biomedical images, including histopathology, microscopy, and radiological data.

**Internal Demand**

Stakeholder priority rating on a 1 to 5 scale from No interest to High demand, indicating organizational need and appetite for adoption of specific AI and ML technologies.

**Metadata**

Structured information describing data origin, experimental context, processing steps, and lineage required to support reproducibility and reuse.

**Monte Carlo Simulations**

Computational methods using repeated random sampling to explore uncertainty, variability, and probabilistic outcomes in complex systems.

**Multimodal AI**

AI systems that integrate and reason across multiple data modalities such as text, images, molecular structures, omics data, and tabular datasets.

**Neuro-Symbolic AI**

Approaches combining symbolic reasoning with data-driven learning methods to improve interpretability and reasoning capability.

**Organizational Maturity**

Current capability assessment on a 1 to 5 scale from Not started to Widely adopted, reflecting infrastructure, skills, processes, and governance supporting AI and ML deployment.

**Ownership Fragmentation**

Distribution of responsibility for AI initiatives across multiple functions without a single accountable owner for end-to-end outcomes.

**Process Debt**

Accumulated workflow inefficiencies and lack of standardization that limit the effectiveness of AI.

### **Reinforcement Learning**

Machine learning approaches in which agents learn optimal actions through interaction with an environment, commonly used for optimization problems.

### **Strategic Gap**

Calculated difference between Internal Demand and Organizational Maturity, scaled by 2.5 for visualization, identifying priority areas where industry needs exceed current capabilities.

### **Sustainability Practices for AI**

Approaches to managing the computational, financial, and environmental cost of AI systems, including training efficiency and carbon footprint.

### **Technology Churn**

Rapid turnover of AI models, tools, and platforms relative to organizational evaluation and adoption cycles.

### **Transformer Models Beyond NLP**

Transformer-based architectures applied to non-text data such as molecular representations, numerical scientific data, or time series.

### **Validation Frameworks**

Structured approaches for assessing model performance, robustness, bias, explainability, and reliability in scientific and regulated environments.

### **Validation Lag**

Delay between model development and trusted deployment due to unclear acceptance criteria, governance gaps, or regulatory uncertainty.

### **Value Leakage**

Progressive loss of potential value as AI initiatives move from pilot to deployment, decision adoption, and measurement.

## Appendix C: Interview Analysis

The following table provides a detailed mapping from themes to subthemes and representative quotes, offering a full audit trail for the findings presented in this report.

Cluster	Subtheme	Representative Quote
Trust & Governance	Validation standards; patient safety first	<i>"The moat is not in the models, it's in the data."</i>
Trust & Governance	Enterprise guardrails, AI advisory, risk-based approvals	<i>"We love innovation, but we want to make sure it's regulated and safe."</i>
Data Quality & Standards	Selection bias; poor cross-modality transfer	<i>"Models built on oncology don't really work very well with other diseases."</i>
Data Quality & Standards	Metadata, ELNs, reproducibility	<i>"Use electronic lab books and metadata standards so others can reproduce."</i>
Culture & Enablement	End-user democratization; agentic exploration	<i>"We're seeing hundreds of agents... duplication is fine for now; we want people to learn."</i>
Collaboration Friction	Legal, procurement, and process delays	<i>"Months to finalize contracts... momentum is gone."</i>
Tech Readiness	LLM limitations in biomedical modalities	<i>"Most foundational models... do not work for many modalities in pharma."</i>
Tech Readiness	Imaging ML interpretability limitations	<i>"Cellular imaging ML struggles – we can't define the changed feature."</i>
Tech Readiness	Reliability barrier (last 5%)	<i>"The last 5% of errors forces users to re-check 100%."</i>
Collaboration Friction	Publisher & vendor ecosystem constraints	<i>"Half our PDFs cannot be processed by GenAI due to licensing."</i>
Culture & Enablement	Literacy gap and vocabulary mismatch	<i>"End users don't understand the language... we need shared vocabulary."</i>
Value & Scientific Insight	Blind-spot discovery	<i>"Ask AI what matters most – you might learn it's not what you thought."</i>

### Interview Methodology Summary

The following steps describe the structured approach used to conduct and analyse the interview phase of this study:

- 1. Standardized questions:** All interviewees were asked the same core questions, making it possible to compare answers across people and companies.
- 2. Verbatim transcription:** All interviews were recorded and transcribed word for word to ensure no details were missed or reliant on memory.
- 3. Key moment tagging:** Important, surprising, or emotionally charged statements (such as frustration or excitement) were tagged to identify what mattered most to participants.
- 4. Thematic grouping:** Similar comments about validation, data quality, culture, and other topics were grouped into thematic buckets.

**5. Cross-interview pattern analysis:** Patterns were examined for recurring topics, points of agreement and disagreement, urgency, and enthusiasm to build a cross-company picture.

**6. Representative quote selection:** The strongest quotes were chosen to represent each theme, ensuring the real voice of the industry is reflected rather than the authors' interpretation.

**7. Theme rating:** Each theme was rated by demand (how many people care about it), maturity (how far companies have progressed), and collaboration potential (how well it can be solved together) to identify the best pre-competitive opportunities.

**8. Full audit trail:** Every theme can be traced back to the exact interview and quote. Nothing was assumed; everything is evidence-based.