CSL Business Development

AWS Bioinformatics Agentic Workflow

Use Case Descriptions for Development Teams

March 2026 | Draft v1.0

Introduction

CSL's Business Development group operates in a fast-moving global landscape where identifying the right biotech and pharma partners – at the right moment – can define competitive advantage for years. The volume of scientific literature, clinical trial data, patent filings, and corporate disclosures now far exceeds the practical capacity of any human team to monitor and synthesise in real time.

This document describes four illustrative use cases for an AWS-hosted, multi-agent bioinformatics workflow designed to augment the BD team's capability. Each use case is built on a common architecture of Amazon Bedrock agents, AWS HealthLake, Amazon SageMaker, and purpose-built orchestration – working together in an 'agentic' pattern where specialised AI agents plan, reason, and act autonomously to complete complex, multi-step tasks with minimal human intervention.

The four use cases presented here are:

- Use Case 1: Real-Time Pipeline Intelligence & Partnership Opportunity Scanning
- Use Case 2: Accelerated Scientific Due Diligence
- Use Case 3: Genomic White Space Analysis for Co-Development Targeting
- Use Case 4: Dynamic Partnership Fit Scoring & Recommendation Engine

Use Case 1: Real-Time Pipeline Intelligence & Partnership Opportunity Scanning

An autonomous monitoring agent that continuously watches the global biotech and pharma landscape, surfacing relevant partnership signals before they become widely known.

AWS Services	Amazon Bedrock (multi-agent orchestration), AWS HealthLake, Amazon Comprehend Medical, Amazon EventBridge (scheduling), Amazon S3 (source data lake), AWS Lambda
BD Value	Reduces time-to-awareness of high-value partnership targets from weeks to hours; ensures BD team have first-mover advantage on emerging opportunities

The Challenge

CSL's BD team currently relies on periodic analyst reviews, conference attendance, and relationship networks to identify relevant partnership opportunities. This episodic approach means that promising biotech companies – particularly those raising Series A/B funding or advancing into Phase I/II trials – can go unnoticed until competitors have already initiated discussions. The sheer volume of sources to monitor (ClinicalTrials.gov, PubMed, SEC/EDGAR filings, EMA and FDA submissions, patent databases, press releases, and deal-tracking databases) makes comprehensive coverage manually impractical.

Agentic Solution

A scheduled Bedrock agent orchestrates a suite of specialised sub-agents that continuously ingest and parse structured and unstructured data across all relevant public sources. The agents apply CSL's strategic filter – including therapeutic area alignment (plasma-derived therapies, vaccines, nephrology/iron deficiency), geography, development stage, and modality – and rank emerging companies and assets by partnership relevance. A weekly digest is automatically generated and delivered to the BD team, while high-confidence signals trigger immediate notifications.

Key Workflow Steps

- EventBridge triggers the orchestrator agent on a defined schedule (daily/weekly).
- Sub-agents scrape and parse ClinicalTrials.gov, PubMed, SEC EDGAR, patent APIs (USPTO/EPO), and specialist databases (Evaluate Pharma, Citeline).
- Amazon Comprehend Medical extracts entities (drug names, targets, indications, trial phases) and links them to CSL's internal ontology stored in HealthLake.
- A scoring sub-agent ranks each signal against configurable BD criteria using a Bedrock-hosted LLM reasoning chain.
- Ranked outputs are written to a structured database; a report-generation agent produces a formatted partnership briefing document.
- High-scoring signals (≥ threshold) are pushed to the BD team via email or Slack with a one-paragraph rationale.

Example Output for the BD Team

A ranked weekly briefing showing 8–12 high-priority biotech companies with complementary pipeline assets, each accompanied by a plain-English rationale, the key data sources consulted, and a ‘next action’ recommendation (e.g. ‘Initiate outreach via mutual KOL contact’).

Use Case 2: Accelerated Scientific Due Diligence

An on-demand agentic workflow that compresses weeks of analyst due diligence work into hours, producing a structured scientific assessment of a target partner's pipeline.

AWS Services	Amazon Bedrock (agent orchestration + Titan/Claude models), Amazon SageMaker (custom bioinformatics models), AWS HealthLake (clinical data), Amazon Kendra (internal document retrieval), Amazon S3
BD Value	Compresses scientific due diligence from 2–4 weeks to 24–48 hours; standardises output quality; enables parallel assessment of multiple targets simultaneously

The Challenge

When a potential partnership target is identified, the BD team must perform rigorous scientific due diligence before committing to discussions. This involves reviewing clinical trial data, assessing the underlying science and competitive differentiation, mapping the intellectual property landscape, and evaluating regulatory history. This process currently takes two to four weeks per target and creates a bottleneck that slows the BD pipeline, particularly when evaluating multiple concurrent opportunities.

Agentic Solution

On request, a Bedrock orchestrator agent is given a target company name (or a specific asset). It autonomously dispatches specialist sub-agents to retrieve and synthesise information across scientific literature, clinical databases, patent filings, and regulatory submissions. The agents apply bioinformatics reasoning – for example assessing mechanism-of-action plausibility, biomarker data quality, and clinical endpoint appropriateness – before compiling a structured Due Diligence Brief in a standardised format the BD and scientific teams can act on immediately.

Key Workflow Steps

- BD analyst triggers the workflow via a simple web UI or API call, providing the target company/asset name.
- The orchestrator dispatches: a Literature Agent (PubMed, bioRxiv, patent databases), a Clinical Data Agent (ClinicalTrials.gov, FDA/EMA submissions), an IP Agent (USPTO, EPO, FreePatentsOnline), and a Competitive Positioning Agent.
- Amazon Kendra retrieves relevant CSL internal documents (prior assessments, strategic frameworks) to provide internal context.
- A SageMaker-hosted bioinformatics model evaluates the scientific validity of the target's mechanism of action against genomic and proteomic reference datasets in HealthLake.
- The orchestrator synthesises all sub-agent outputs using a Claude model to resolve conflicting signals and produce a coherent narrative.
- The final Due Diligence Brief is generated as a formatted document and stored in the BD team's SharePoint, with a summary pushed to the analyst.

Example Output for the BD Team

A structured 8–10 page Due Diligence Brief covering: scientific rationale, clinical data quality assessment, IP landscape map, regulatory risk flags, competitive differentiation, and a final ‘scientific confidence score’ with supporting reasoning – ready for BD team review in under 48 hours.

Use Case 3: Genomic White Space Analysis for Co-Development Targeting

A strategic intelligence agent that analyses genomic, proteomic, and clinical datasets to identify therapeutic white spaces where a co-development partnership would provide CSL with a first-mover advantage.

AWS Services	Amazon Bedrock, Amazon SageMaker (genomics models, variant calling pipelines), AWS HealthLake (clinical data), AWS Glue (ETL for multi-omics data), Amazon Athena (federated query), Amazon S3 (genomics data lake)
BD Value	Proactively surfaces co-development opportunities that are scientifically validated and competitively uncontested; reduces reliance on reactive, relationship-driven deal flow

The Challenge

CSL's portfolio strategy depends on identifying disease areas where the science is sufficiently mature to be de-risked, but where no dominant player has yet established a strong position. Identifying these 'white spaces' requires integrating multiple layers of evidence: genomic association data, existing pipeline competitor mapping, unmet clinical need metrics, and regulatory precedent. This kind of multi-dimensional analysis is rarely performed proactively because of the complexity and time required to assemble and query the relevant datasets.

Agentic Solution

A bioinformatics-specialised Bedrock agent orchestrates an analysis pipeline that ingests large-scale genomic association data (GWAS, multi-omics) from public repositories alongside CSL's internal R&D datasets, then cross-references these against competitor pipeline data and published clinical outcome benchmarks. The agent identifies disease indications where strong genetic validation exists, unmet clinical need is high, and CSL's existing capabilities (e.g. plasma-derived biologics, recombinant proteins) could be applied via a co-development partnership. The output is a prioritised matrix of white space opportunities with suggested partner profiles.

Key Workflow Steps

- The orchestrator agent queries public genomic databases (GWAS Catalogue, Open Targets, gnomAD) via API, ingesting genetic association data for disease areas relevant to CSL's portfolio.
- AWS Glue ETL pipelines process and harmonise multi-omics datasets from disparate sources into a unified S3 data lake.
- A SageMaker model scores disease-gene associations by evidence strength, druggability, and translational maturity.
- A Competitor Mapping sub-agent cross-references scored targets against current pipeline databases to identify areas with limited competitive activity.
- An Unmet Need sub-agent analyses published epidemiology, patient advocacy data, and clinical outcome benchmarks to score the burden of disease.

- The orchestrator synthesises scores into a prioritised White Space Matrix and generates a strategic narrative for each opportunity, including suggested biotech/pharma partner archetypes CSL should be targeting.

Example Output for the BD Team

A prioritised White Space Matrix presenting the top 10–15 disease-area opportunities ranked by a composite score (genetic evidence strength, CSL capability fit, competitive openness, unmet need). Each opportunity includes a one-page strategic summary and a description of the ideal partner profile CSL should be seeking.

Use Case 4: Dynamic Partnership Fit Scoring & Recommendation Engine

A configurable multi-agent system that continuously scores and ranks potential biotech and pharma partners against CSL's current strategic priorities, generating explainable recommendations for the BD team.

AWS Services	Amazon Bedrock (multi-agent orchestration, reasoning chains), Amazon DynamoDB (Partner Intelligence Database), Amazon SageMaker (scoring models), Amazon QuickSight (BD dashboard), AWS Lambda, Amazon API Gateway
BD Value	Transforms subjective, episodic partner prioritisation into an objective, real-time, explainable process; enables BD leadership to rapidly re-prioritise the pipeline as strategy evolves

The Challenge

BD teams often maintain large watchlists of potential partners – hundreds of companies at any given time – but lack a systematic, up-to-date, and consistent method for prioritising which relationships to invest in. Scoring is typically subjective, influenced by recency bias, and rarely reflects the full picture of a company's evolving scientific and commercial position. This leads to missed opportunities and inefficient use of the BD team's time.

Agentic Solution

A persistent Bedrock agent maintains a continuously updated Partner Intelligence Database, refreshed by the monitoring agents from Use Case 1. BD leadership can configure the scoring framework via a simple interface – weighting criteria such as therapeutic area alignment, technology platform complementarity, pipeline stage, IP position, geographic fit, and financial stability – and the scoring agent automatically re-ranks the entire watchlist whenever weights are adjusted or new data arrives. All recommendations are fully explainable: the agent provides a structured rationale for each score, citing the specific data points that drove it.

Key Workflow Steps

- The Partner Intelligence Database in DynamoDB stores structured profiles for all companies on the BD watchlist, continuously refreshed by the pipeline scanning agents.
- BD leadership configures scoring weights via a lightweight web interface (API Gateway + Lambda); weights are versioned so historical scoring runs can be compared.
- The Scoring Agent retrieves each company's profile and applies the weighted scoring model, using a Bedrock-hosted LLM to reason over qualitative factors (e.g. management team track record, partnership history) that resist pure quantification.
- A Recommendation Agent identifies the top 20 partners for immediate BD outreach and generates a structured rationale for each, citing the specific data points that drove the recommendation.
- An Explainability sub-agent produces a plain-English justification for every score, ensuring BD leaders can interrogate and challenge the model's reasoning.

- Results are surfaced via an Amazon QuickSight dashboard giving the BD team a real-time ranked view of their entire partner landscape, with drill-down capability to the underlying evidence.

Example Output for the BD Team

A live BD Partner Radar dashboard showing the full watchlist ranked by composite partnership fit score, with colour-coded priority tiers and one-click access to each company's detailed rationale. Weekly email digests highlight significant rank changes (e.g. a company that has moved up sharply following a positive Phase II readout), ensuring the BD team always act on the freshest intelligence.

Common Architecture Principles

All four use cases share a consistent set of architectural principles that simplify development, reduce cost, and ensure the platform can be extended over time:

- Modular agent design – each sub-agent is independently deployable and testable, enabling incremental development and easy substitution of components as AWS services evolve.
- Human-in-the-loop checkpoints – the orchestrator is designed to flag low-confidence outputs for human review rather than suppressing uncertainty, preserving analyst trust in the system.
- Auditability – every agent action, data source consulted, and reasoning step is logged to CloudWatch and S3, providing a full audit trail for compliance and model governance purposes.
- Security & data governance – all data flows remain within the CSL AWS environment; no proprietary or clinical data leaves the account boundary; IAM policies enforce least-privilege access at the agent level.
- Configurable thresholds – scoring weights, alert thresholds, and data source priorities are all managed as configuration, not code, enabling the BD team to tune the system without developer involvement.

Development teams should treat Use Cases 1 and 2 as the highest-priority starting points, as they deliver the most immediate and tangible value to the BD group and establish the core agent orchestration and data ingestion infrastructure that Use Cases 3 and 4 build upon.