



PointCross Solutions
www.pointcross.com
pointcross.wordpress.com
get_started@pointcross.com

SEARCH & ORIENTEERING SOLUTIONS FOR PHARMACEUTICAL R&D

PointCross's Search & Orienteering offerings enable scientists to exploit institutional information from structured data and unstructured content, fostering new discoveries and finding safety signals sooner to drive innovation and better pipeline decisions.

THE SITUATION

Pharmaceutical companies have billions of dollars worth of structured and unstructured data within their own data warehouses, servers and drives, and across their partner and CRO network. Unfortunately, they are unable to exploit this information for pipeline decisions because their informatics infrastructure does not provide holistic, contextualized, and ready access to troll through these resources to find and discover patterns, correlations and insights.

Business users typically rely on data and application administrators to collect and analyze the data, which can often take weeks or months. The turnaround time is compounded if the information is spread across multiple applications, geographies, or partner organizations. Further, legacy data from retired applications is often lost forever and important legacy data may only exist in paper or PDF formats. Finally, unlocking knowledge from unstructured data contained within emails, documents, and meeting notes (which experts tell us now conservatively account for 80% of corporate information¹) is impossible with most information technologies being used today. In fact, data warehouses that house most of companies' raw data were designed primarily as regulatory repositories, not necessarily for scientists to formulate new concepts and look for patterns and correlations hidden within study datasets.

In such an environment, **meta-analysis** – an approach for cross study analysis that controls for measurement and sampling errors which may otherwise distort signals within individual studies – remains a distant dream. Meta-analysis relies largely on the ability to create and harvest metadata from raw data to enable the discovery of patterns and correlations that may otherwise be hidden. In most cases, vital patterns that meta-analysis can reveal is lost because Pharma organizations do not have access to the right informatics tools. As a result, scientists may be forced to conduct studies that may be unnecessary or that may have previously been completed in a different context. In fact, IDC tells us

¹ Knox, Rita, T. Eid, and A. White. "Management Update: Companies should align their structured and unstructured data," Gartner Research, Feb 2005.



that, “as high as 90% of the time creating information is actually recreating existing information.”² It is also well-documented that knowledge workers spend 30-40% of their time looking for information that already exists, translating into costs approaching \$26,000 for these roles in an enterprise.³

For Big Pharma, this means not only millions of dollars spent and time lost reinventing the wheel every year, but billions of dollars on failed drug candidates that whose toxic side effects could potentially have been identified earlier in development. Being able to decrease late-stage attrition may be the single greatest factor that can lower Pharma companies’ high clinical trial costs, while increasing pipeline throughput and profitability. Since scientists do not have relevant information from the past at their disposal to compare results with ongoing studies, early and accurate assessments of the viability of these substances are difficult to make and adverse indications are often not caught until Phase II or III when the risks and costs are significant. Then there is also the incalculable opportunity cost of new drugs that go undiscovered, or of discovering beneficial or unexpected secondary effects to open up new therapeutic opportunities, simply because scientists lack the ability to connect necessary knowledge and data patterns across the organization.

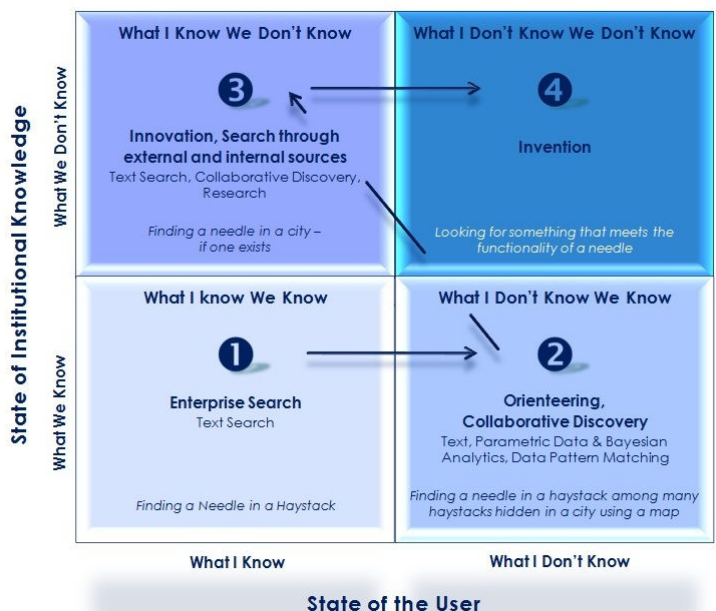
SOLUTION

PointCross’s **Search and Orienteering Solution** allows scientists to discover, “*What I don’t know that the organization knows,*” and to leverage institutional knowledge to make better pipeline decisions, gain valuable insights and provide opportunity for new scientific discoveries to continually drive innovation.

The solution provides a rich set of capabilities that empower scientists to explore their institutional knowledge, allowing them to:

- ☑ Perform meta-analysis across semantically harmonized information and meta-data from multiple studies and sources that have been normalized *a priori*;
- ☑ Orienteer and navigate across qualitative and quantitative study data and meta-data, as well as within unstructured content (emails, meeting notes and documents);
- ☑ Find correlations, patterns, trends, and outliers in multiple dimensions, including study types, findings, compounds, and animals/patients;
- ☑ Combine metadata extractors, analytics, meta-

The Search for Knowledge and Innovation in an Enterprise



² Feldman & Sherman, 2003: The High Cost of Not Finding Information

³ Stewart, 2008: Building Enterprise Taxonomies

analysis, and the ability to sense weak pharmacovigilance & sentinel signals to support a “Principle of Precaution” and meaningful response via federated search facilities. It is possible to intertwine internal and external data sources, rich metadata extraction, meaning and concept matching, and automatic classification, with workflow triggers tied to concepts, and alerting to relevant roles/people within an organization.

- ☑ Allows scientists to interact with search results through tabular views, graphical visualizations and thematic data browsers;
- ☑ Create new concepts from search results and collaborate contextually; and
- ☑ Make better and faster strategic and pipeline decisions.

The search solution provides techniques to narrow down a vast set of data by selecting specific attributes, or facets, of the information that the user is seeking. By proactively offering a set of clarifiers that are salient to the user’s search parameters, user can ‘drill down’ with increasing specificity to the exact dataset s/he is seeking. This is a useful technique for navigating through a large amount of data in a multi-dimensional way to arrive at a focused expression of the user’s goal.

Orienteering is a related concept that allows users to chart multiple foraging pathways through institutional data, to understand the context of their search results, and to explore new and often unforeseen relationships across substances, people, applications, and geographies, breaking down these traditional silos. This exploration is the key to scientists connecting patterns across disparate, disconnected areas of knowledge, seeing new patterns that were previously hidden, and interacting with their peers for collaborative discovery. These discoveries can lead to innovation and the next successful drug – or the early realization that a substance in the pipeline is likely to fail.

The key technology components of PointCross’s Search & Orienteering solution include:

- ☑ **Semantic Data Exchanger (SDE):** an any-to-any data converter that semantically normalizes and unifies data from multiple sources (systems, formats and geographies) into a common Universal Operational Data Model (UODM). SDE ensures a common unified information base for search and orienteering by ensuring that:
 - Semantic differences in terminologies between multiple sources are harmonized to standard Controlled Terminologies by creating synonyms and parent-child linkages, while preserving relationships to local terms with which people are familiar;
 - Parametric data (e.g. measurements taken for Hematology, Clinical Pathology, or Clinical Observations) differences arising from using different units systems within and across the company firewalls are standardized.
- ☑ **Contextualization of Legacy Content with Free Text Search** for qualitative data like patient details, site visit observations or other findings data, as well as unstructured content (protocols, reports and related collateral). Facilities are provided to contextualize legacy content including electronic and paper versions from internal/external sources, and to harvest related data and metadata through federation. Also enabled is the ability to meta-tag all such information using a rich set of taxonomies including folksonomies, A.K.As, and controlled or industry standard terminologies.

- ☑ **Metadata character search** for rapid queries that only scan metadata within contexts (business topics), email, document, meetings and selected processes for matching character string patterns. Results are automatically contextually organized along multiple facets.
- ☑ **Full Content search** through embedded Lucene/SOLR search capabilities that harvest information across multiple content types; other search engines including Microsoft FAST, Google and others are readily supported.
- ☑ **Parametric Search** for quantitative data such as clinical pathology, hematology, body weight and others, by indexing not only the raw data but also extracting and indexing statistical metadata and associated patterns within and across studies. These statistical indexes are created using a wide variety of statistical analytics including linear and nonlinear modeling, classical statistical tests, time-series analysis, classification, and clustering.
- ☑ **Security and Authorization** for managing people and role access on a granular, need-to-know basis. Rather than relying solely on Access Control Lists for individual applications, the solution uses a context-based approach to harmonize authorization in line with business needs. **Blind Search** facilities trigger workflows to authorized users who can choose to grant information access to others who may not otherwise be able to tap into it.
- ☑ **Collaborative** controls to interact with peers in context with live and saved search results. All collaborative facilities directly integrate with common enterprise applications such as Microsoft SharePoint and Office.
- ☑ **Governance and Compliance Services** for importing and normalizing legacy data and documents including dashboards and controls for reconciling data discrepancies, applying transformation rules, and capturing annotations to ensure full audit traceability back to the original source.

BENEFITS

Pharmaceutical companies have spent years and millions of dollars in search of the holy grail of data: holistic, secure, and contextualized access to institutional knowledge across the organization and corporate network. Within a few months, and at a fraction of the cost, PointCross's Search and Orienteering solution delivers tangible results by affordably empowering scientists to explore their institutional knowledge in meaningful ways. The solution also:

- ☑ Empowers scientists and pharmaceutical companies to identify potential drug candidates, and failures, in a fraction of time and drive better decision making and reduced cost by lowering the attrition rate in later drug developmental phases;
- ☑ Identifies and establishes common objects and concepts, and resolves conflicts between different sources by meta-tagging based on Controlled Terminologies;
- ☑ Leverages contextualized meta-data concepts while integrating free text search to provide visual cues of context for organization of results as well as for subsequent queries. This helps suggest logical steps in search and perhaps unexpected alternatives, which is vital for drug discovery, and which can act as an important scaffold for rapid incremental and reversible exploration for new or hidden trends, gaps, and outliers. The solution helps pharmaceutical researchers to reduce the burden of information overload and aids in discovering novel information from vast diverse data;
- ☑ Allows researchers, via simple yet comprehensive dashboards, to understand concepts/entities associated with search terms at a quick glance before navigating into thematic data views, thereby reducing the time spent in searching for information and allowing scientists to instead better understand patterns and correlations;

- ☑ Combines search with statistical analytics to improve research productivity and serendipitous discovery. Unlike traditional search engines which efficiently retrieve the data but not any hidden patterns, PointCross's search solution retrieves both raw data as well as automatically extracted meta-data associated with it;
- ☑ Expedites therapeutic efficacy of a drug molecule from among a heap of studies by enabling scientists to perform meta-analysis and draw more definitive conclusions sooner.

ABOUT POINTCROSS

PointCross is a global provider of advanced strategic business solutions to knowledge-rich markets, including the pharmaceutical industry. Our Integrated Drug Development Suite (IDDS™) specifically addresses the pharmaceutical industry's key R&D business needs. The search and orienteering solution built on our Codetta™ search environment is integral to IDDS. At the heart of IDDS is Orchestra+Solo™, an adaptive, contextual knowledge environment and personalized client that orchestrates core business processes. This robust solution set delivers the following capabilities:

- ☑ Single point of access to contextualized tacit and structured knowledge across the enterprise, with search and guided navigation within and across contexts;
- ☑ Robust search and orienteering capabilities across studies, emails, documents, meta-data and more across the entire organization, CROs and partners
- ☑ Flexible, fool-proof IP security based on contexts and roles, determined by business rules;
- ☑ Predictive analytics for clinical and non-clinical data;
- ☑ Secure multi-party workflows for knowledge sharing and business-social networks within and across companies;
- ☑ Semantic Data Exchange (SDE) for vendor-neutral data exchange, normalization and unit harmonization;
- ☑ Business development, in/out-licensing, e-discovery, audit, regulatory submissions, compliance, and more;
- ☑ Scalable architecture and development toolkits for additional capabilities.

PointCross represents a new way of doing business. We deliver business ready solutions in 1/10th the time and a fraction of the costs compared to standard technologies, while offering strategic advice from people who know the pharmaceutical industry.

We are headquartered in the California Bay Area. We are a CDISC solution provider to the Pharmaceutical Industry, and a Microsoft Gold Certified partner. We also have a global network of service, consulting and infrastructure partners.



For more information, visit us at www.pointcross.com and call us at (650) 350-1900. Also, check out our blog at <http://pointcross.wordpress.com>.

