6-Step Guide to Turn Clinical Data into Gold

For Better Patient Care and New Revenue

melissa
INFORMATICS

www.melissainformatics.io

Tap Your Clinical Data Goldmine

Hospitals, Clinical Care and Clinical research organizations that have collected data for several years are sitting on a data goldmine. Now, you can realize new value from your clinical data to:

- Improve patient care and support precision medicine research
- Engage high paying pharmaceutical research partnerships
- Realize new intellectual property from clinical practice data
- Capture and report research quality clinical trials data at a lower cost
- Achieve unexpected new revenue by making the most of your clinical data

While real-world clinical data is unusually valuable, it is also unusually complex and diverse. Even the most disciplined clinical practices will generate dirty data. That's where Melissa Informatics can help. We offer a unique set of clinical data quality and integration tools that will quickly curate your data, harmonize it to common standards, and link or "integrate" the data in a manner that supports rich searching and reporting.

Using AI to Improve Clinical Data Quality and Interoperability

Taking control of your data and freeing it for transparent access has traditionally been a time consuming and expensive task, requiring manual data curation and creation of new "data warehouses" that aren't suited to new or changing data and research requirements. Melissa Informatics changes all of this by bringing advanced "AI" – including machine reasoning, natural language processing and machine learning using advanced graph databases – to reduce the time and cost for data harmonization and integration, and to deliver a well-defined, interoperable and adaptable AI-enabled database or "knowledgebase" to you. With an improved and future-proof data and knowledge base environment, you can apply research grade quality queries to your newly integrated data and:

- Assess drug efficacy and contraindications for patient populations and individual patients
- Discover and pursue new and innovative treatment regimen by identifying effective health outcomes from historical data
- Monetize your research quality data and improve healthcare by sharing and collaborating with pharma and university researchers

Read our 6-Step Guide to turning your disconnected and dirty laboratory and clinical study data into a clean, unified data resource, and learn how Melissa Informatics can help you tap your under-used data goldmine.

Roadmap to Value

Melissa Informatics can turn your diverse and disconnected laboratory and clinical study data – from electronic medical records (EMR), electronic health records (EHR), laboratory information management systems (LIMS) and other sources – into a unified, high quality, standards-compliant, research-ready data resource. We can turn your expensively gathered and under-valued data into a resource for new knowledge, improved care, and substantial new revenue.

1 Inventory, Access

Inventory, identify and gain searchable, easily repeated access to your data. Apply combined machine learning and machine reasoning to access and extract critical information required to integrated data that was previously hidden or difficult to access. Now you have the "ore" you need to turn clinical data into gold.

2 Curate, Integrate

Dirty, disconnected data isn't useful for improved patient care, research or collaboration. Turn dirty "ore into gold" through more efficient curation and integration. Melissa applies advanced semantic technologies to reduce the time and effort to curate and integrate.

5 Realize Benefits for Community, Patients and Revenue

Turn your valuable but dirty and disconnected "real-world-evidence" into clean, research-quality data aligned with the standards you need to engage high value partnerships.

6 Partner with Melissa Informatics

Melissa Informatics brings over 30 years of successful experience to our data quality and analytical partnerships. We apply the best of traditional data curation technologies with advanced AI, including machine reasoning and learning, to reduce time and cost required to meet your data-driven goals. Partner with Melissa Informatics to improve your data quality, to enable transformative data-enabled applications and to realize unexpected new revenue.

4 Make Data FAIR and Standards Compliant

Realize value from existing clinical data resources by meeting requirements for useful data research and publication. Ensure data is secure, properly de-identified and consented when relevant, and is "findable, accessible, interoperable and reusable".

3 Create Relationships, Apply AI and Act on Patterns

Apply "ontologies" to define semantic concepts and relationships required to apply machine reasoning and pattern recognition with your data. Explore integrated data to identify patterns of interest, whether for additional data curation and for advanced decision support with data that is richly described, research-ready, and "future-proofed."

1 Inventory and Access Your Data

Have you ever asked questions like: "Where is that data resource we created?" "Do we have that patient information?" "How can I find patients who meet criteria for clinical trials or more effective treatment?" If so, you are losing mission-critical time and resources. You need to inventory, identify and gain searchable, easily repeated access to your data.

Melissa Informatics (MI) has experience with efficient inventory of your data and applications environment. We can also help you obtain patient consent to access and apply data towards better treatment and new cures. Once we have identified data sources, our "Sentient" software applies the state of the art AI, including machine learning and machine reasoning, to access and extract critical information required to integrated data that was previously hidden or difficult to access. Now you have the "ore" you need to turn your clinical data into gold.

2 Curate and Integrate Your Data

Data as it comes from EMR systems and other clinical applications are uncommonly complex. Dirty, disconnected data isn't useful for improved patient care, research or collaboration. To turn your dirty ore into gold, your data needs to be curated and integrated. MI has decades of data quality experience with global fortune 500 companies. Rather than demanding expensive manual data curation, we apply semantic technology that reduces the time and effort to curate (normalize, harmonize to common standards) and link (connect using the latest graph databases).

Data curation or "cleaning", including harmonization and integration, makes it possible to exploit the knowledge currently hidden within EMR/EHR systems, past and current clinical and research databases, historical data sources and other important data sets. Melissa Informatics "Sentient" software can reduce time and cost to transform a mess of "dirty" but valuable clinical data into an integrated, research quality resource. The resulting resource, or "knowledgebase", makes the best use of the incredible wealth of data that clinics have collected over the years.

MI can help you turn your under-used, inaccessible, disconnected and "dirty" data "ore" into findable, accessible, interoperable and reusable gold. MI's "Sentient" software transforms disconnected data - characterized by entry errors, spelling errors, missing data, conflicting data – into harmonized (different dataset are normalized to common units and terminologies), integrated (linked, richly connected), and deeply searchable data.

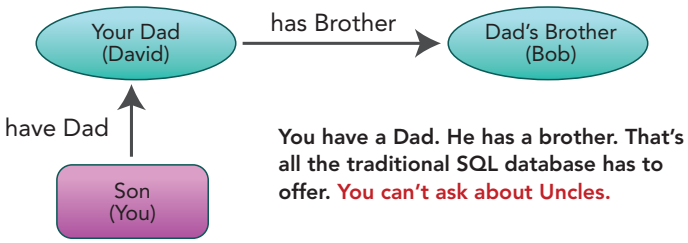
3 Create Relationships, Apply Machine Reasoning (AI), Identify and Act on Patterns

Once data are linked, MI's Sentient product can apply "ontologies" that define the semantic concepts and relationships that required for an advanced "semantic databases" to apply machine reasoning. The system can now explore the full dataset and identify patterns of interest for additional data curation and for advanced decision support. "Linked data" is custom-designed to provide a flexible and extensible master data management (MDM) environment that is research-ready "future-proofed."

Ontologies are meaningful (epistemologically coherent) graphs that provide practically relevant knowledge about concepts and relationships in any area of interest. Because ontologies are explicitly meaningful to both humans and computer, they make reasoning possible, even on incomplete and 'dirty' data. This allows the computer to "reason" about the data, to find patterns, to correct errors and even to fill in blank spots in your data.

For example, let's say you have a database with information about parents and siblings, but no extended family information. You want to know, "Is Bob my Uncle?" Unfortunately, in the current database, there isn't any specified information about who your uncle might be.

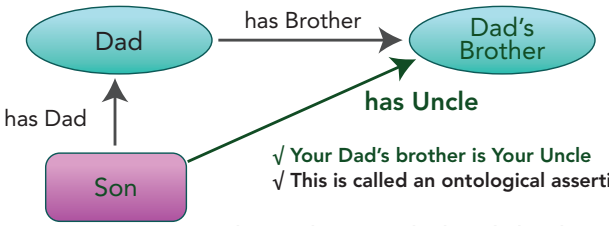
The database doesn't know if you have an Uncle



You have a Dad. He has a brother. That's all the traditional SQL database has to offer. You can't ask about Uncles.

With the semantic ontology-enabled Sentient Platform, you can apply machine reasoning to learn whether Bob is your uncle.

With an ontology, you can reason using the limited data you already have



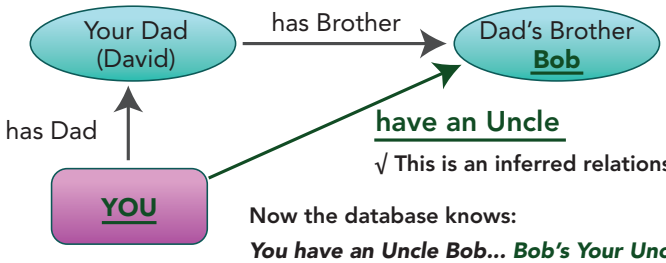
√ Your Dad's brother is Your Uncle
√ This is called an ontological assertion

This ontology provides knowledge that can be used for reasoning => if your Dad has a brother, you have an Uncle.

Challenge solved, through the power of ontology-enabled machine reasoning. Even though it wasn't in the original data source, you now have the information you need to know that Bob's your Uncle.

Your original database only had information about parents and siblings, but by applying ontological reasoning (also called "inference"), you can now infer who your Uncle is. With Sentient, you have been able to apply machine reasoning to learn whether Bob is your uncle and you can now add new data to the database. Adding new or corrected information to a database following the application of reasoning is called "entailment".

NOW I have the data I need to answer my question



√ This is an inferred relationship

Now the database knows:
You have an Uncle Bob... Bob's Your Uncle.

Machine reasoning provides a new and much more efficient way to clean up and enrich your data. Although the example above is extremely simple, in real practice the Sentient system will link billions of data points, to identify and validate information required for new insight.

Melissa's Sentient software brings advanced technology to data quality. With Sentient you can apply AI to find and create new relationships, to refine and grow the hidden value of your data. If you're technically inclined, you may enjoy this section. If not, skip to the end!

Melissa's Sentient software delivers ontology-enabled artificial intelligence to reduce time and cost for effective data harmonization, integration and achievement of research quality data outcomes. To do this, Melissa applies semantic ontologies and machine reasoning.

Ask us to demonstrate how our system can apply an advanced form of artificial intelligence (AI), by application of ontology-enabled Machine Reasoning to create new relationships and new meaning from unstructured and structured data

4 Make Data FAIR and Standards Compliant

If your mandate is to share or collaborate around clinical and published study data, you need to be aligned with standards required for collaboration and interoperability. To meet FAIR guidelines data must be findable, accessible, interoperable and reusable while maintaining regulatory compliance for protected health information, contractual and business compliance, security, and confidentiality (See M. Wilkinson et al, "The FAIR Guiding Principles for scientific data management and stewardship," Scientific Data, 2016.).You can more easily and effectively publish and share (or collaborate around) your own data by meeting FAIR guidelines and all required standards compliance for data publication and sharing.

To ensure Findable data, it is important to add structure and metadata to your data, including searchable globally unique and persistent resource identifiers (URIs). We apply the World-wide-web Consortium's (W3C) Resource Description Framework (like HTML but for data!) standard for URIs for all data and metadata elements in a secure, standards-based data source. To be Accessible, data are made retrievable by their identifier using a standardized communications protocol that is universally implementable with authentication and authorization; and metadata is managed to ensure accessibility. To be Interoperable, all data is defined using accessible, shared, and broadly applicable "semantic" methods for knowledge representation, using standards and lexicons that follow FAIR principles. To be Reusable, data is well described and released with a clear and accessible data usage license that meets domain-relevant requirements.

MI can make it possible for you to elegantly meet standards and guidelines required for useful data. Melissa's Sentient software makes it easy to meet US FDA, UK NHS and other standards required for different applications. MI is EMEA compliant, GDPR and HIPAA certified. We have experience and infrastructure required to work with your highly sensitive protected health information (PHI). We have deep experience with data de-identification and patient consenting and can partner with you to ensure useful data that doesn't put PHI at risk. We can also help ensure your own compliance at the level of software for data management.

We can also apply our Sentient system and its Lexicon resources for Drugs, Diseases, Genes and Proteins, to transform your data from one lexical Standard to another. Does FDA want a submission aligned with the CDISC Standard? Done. Does AstraZeneca Pharmaceutical company need the data aligned with the SDTM Standard? Done. MI is familiar and can help you apply our tools to rapidly align with different standards - ranging from UMLS down to icd9 and 10, HL7, SnoMed, MeSH, FHIR, and many others - according to your needs.

5 Realize Benefits for Community, Patients and Revenue

Where data is previously un-usable, after exposure to Melissa's Sentient application, data become research quality. Structure and content required for research is provided. Perhaps more important, the data is ready for rewarding uses, for patient care, and for collaboration. Clean and integrated clinical data environments make it possible for hospitals

and clinics to improve care at lower costs, to share data with leading research communities, and to realize substantial revenue from data brokering collaborations and more efficient clinical trials. Once data are managed using Sentient, it becomes easy to search and report different content for different uses. Need all of the information about a specific drug? No problem. Need a Control Group? No problem. Need data prepared for analysis in statistical packages like SAS or R? Done. Need to identify a patient cohort? Easily done. De-identified? Done. Securely accessible identified data? Provided.

Clinical Uses – Improved clinical outcomes through integrated patient data access

More intelligently and cost effectively treat patients by providing a comprehensive "360 degree view" of gold standard information about a patient. The ability for a clinician or clinical researcher to search for information that is meaningfully "linked" across previously connected data resources is one key to moving patient care forward faster. An integrated knowledge base makes it possible to:

- More effectively treat patients through comprehensive insight into each patient's unique symptoms, treatment history and responses
- Deepen understanding of causes and treatments for disease
- More accurately qualify clinical trial participants
- More cost effectively construct innovative clinical trials
- Establish new and alternative endpoints for treatments and trials
- Improve healthcare by making personalized, precision medicine a reality

Clinical, administrative and scientific users benefit by the ability to easily ask virtually any question across all permitted data within a clinic, with results in seconds. This will provide critical information to more quickly and cost effectively evaluate treatment options for patients.

Clinicians can avoid poring through different, disconnected software and printed reports during clinical visits, losing valuable face time with patients. Researchers can ask critical questions that used to take months, getting clear reports back in seconds.

Community Assets – benefit from published data and give back to the community

In addition to helping you with monetizing proprietary data, Melissa Informatics makes it possible for you to more effectively use public / published data resources.

Below are a just a few examples for data sources that Melissa Data has integrated for clinical customers interested in understanding their disease specialization at a molecular biological level:

Dataset Name	Dataset URL	Dataset Description
String-DB	http://string-db.org/	Protein-protein and gene-protein interaction database
Reactome	http://www.reactome.org/	Open-Source, curated and peer reviewed pathway database
Biogrid	http://thebiogrid.org/	Protein-gene interactions from major model organism species
Gene Ontology	http://geneontology.org	Controlled vocabulary of genes, gene products, biological roles
Entrez Gene	http://www.ncbi.nlm.nih.gov/gene	Deep gene-specific information
UniProt	http://www.uniprot.org/	Protein sequence and functional information, various organisms
Ensembl	http://www.ensembl.org/	Genetic data of vertebrates and other model organisms
MyGene	http://mygene.info/	Gene and protein annotation and identification service

Revenue – clinical trials, biotechnology and pharma partnerships

IIn addition to more quickly identifying and qualifying clinical trials participants, a clean, integrated data resource makes it possible to engage data-driven partnerships with biotechnology and pharmaceutical companies. Research quality, integrated clinical data is in high demand within biotechnology and pharma industries, including your data such as:

- Electronic Medical Records / Electronic Health Records (EMR/HER – clinical outcomes)
- Blood and body chemistry, including metabolomic data and laboratory assessments
- Genomic and proteomic information
- Tissue bank information

Access to properly de-identified and searchable patient data (or, "real world evidence") for life science research remains challenging. Innovative methods are required to clean and valuable but under-used internal and external data resources.

We can also make it possible for you to create data assets for collaboration and revenue. Melissa Informatics can help you turn your valuable but under-used data resources into gold!

Partner with Melissa Informatics for monetization

Melissa has experience moving under-used data from clinics into fiscally rewarding data sharing and collaboration partnerships. Melissa Informatics can turn your valuable but dirty and disconnected "real-world-evidence" into clean, FAIR, research-quality data. This is the key step required to engage high value partnerships.

Satisfied Melissa Informatics customer and partner relationships

Depending on the quality and depth of your data, revenue can range from six to seven figure deals, with recurring revenue from one or multiple pharma and biotech partnerships. Read our Customer Case Studies for example success stories. Melissa Informatics has made it possible for customers to realize valuable new partnerships, intellectual property and revenue from their data.

Before working with Melissa Informatics, many of our customers had to manually curate and integrate data to answer clinical or research questions or to engage beneficial data-driven collaborations. It would often take days, weeks, even months to create the dataset required to address narrow sets of questions. With the data quality and management Melissa Informatics provides, customers can ask virtually any question from all of their data in seconds. Just as important, our customers' data is delivering completely new and unexpected revenue.

Working with Melissa Informatics makes it possible to create and apply high quality data resources for improved immediate care, ground breaking discoveries, and profitable partnerships. Melissa Informatics can help you realize knowledge and revenue from clinical data, to improve medical care and meet critical business goals.

6 Partner with Melissa Informatics

Deep Experience

- Benefit from over 30 years in business with 15 years in Healthcare / Life Science Data Harmonization and Search
- MI has seen it all, and solved every data quality and integration problem we've faced

Transformational Technology and Methods

- Drive new discoveries, better science and business
- Publications in Nature, International Journal of Systems Biology, W3C and many others
- CIO Review's Top 20 Most Promising Integration Providers, Bio-IT Best-In Show Award, Peer-reviewed competition / CHI Best Practices Award, HIPAA and W3C compliance

World-class Customers and Partners

- Partnerships and alignment with technology and industry accelerate our technology and your business
- W3C, Stanford BMIR and NCBO, UBC, PROOF Centre, St. Paul's Hospital, Max Planck, Roche, AZ, Merck, …

Global Presence and Support

- Stable and growing business
- Globally distributed, 24/7 server and API access options
- Globally located Software, Support and Data Engineering Teams

Melissa Informatics makes a difference

- "[Tools and team from MI] was absolutely fantastic at getting data harmonized, connected and into our system."- Team Leader, tranSMART Datathon, June 2015
- "Questions that used to take my team six months to answer… are now answered in six seconds." - Dr. S. Tracy, AstraZeneca, Patient Safety Science Lead (CSHALS, February 2015)

Ready to Take the Next Step?

Interested in learning more?

Visit **melissainformatics.io/gold**, email **sales@melissainformatics.io** or call us at **800-MELISSA** to discuss your business and data needs, request a demo or engage a project.

melissa
INFORMATICS

www.melissainformatics.io

About Melissa Informatics

Melissa Informatics extends the capabilities of Melissa's global intelligence software and services to support world leaders in life sciences, biotechnology, pharmaceutical, and medical industries by harnessing the entire data lifecycle for business, pharmaceutical and clinical data. Our software and services bring data quality and machine reasoning together for insight and discovery by intelligently cleaning, connecting and harmonizing multiple sources to offer interoperable data. Melissa Informatics reduces time and cost to benefit from clean, richly connected data, and reveals deeper data relationships from complex, dynamic data through machine reasoning operations for reliable information in mission critical healthcare and life science informatics.

US

22382 Avenida Empresa
Rancho Santa Margarita, CA 92688-2112

800.MELISSA (635.4772)

UK

+44 (0)20 7718 0070

GERMANY

+49 (0)30 7978 8829

INDIA

+91 (0)80 4854 0142

AUSTRALIA

+61 02 8091 6000