Accelerating clinical data management.

We Deliver Better Business Outcomes, Faster

This document describes our desire to work with your organization to accelerate its clinical data management processes. The TrialTwin™ platform can deliver immediate value on a fixed price, fixed timeline basis.

Fast-changing market, regulatory, and reimbursement realities are forcing both sponsors and CROs to change their old ways of doing things. Our take on how clinical trial data management must change is defined in Our End Goal: Accelerate Clinical Trial Data Flow below.

Our approach is to start at the end: helping build faster, cleaner, defensible, more compliant submissions. Our approach is “digital-at-birth”, with a digital end-to-end integrated solution. And with a single source of truth where changes ripple through all downstream processes.

TrialTwin™ is an integrated suite of modules described in detail below, including:
* Metadata Repository
* Protocol Manager
* Synthetic Health Data
* Open Data Repository

Our software was designed specifically for Life Sciences organizations. We work with a large sponsor to build our software from scratch based on their requirements.

We offer a “done-for-you” service with a very short time to value. Our team performs the heavy lifting. And we make it easy to work with us:
* implement modules one at a time
* small, low-budget PoCs
* incremental build-outs
* no need for IT involvement

An instance of TrialTwin™ specific to your organization can be operational in a few weeks after receiving a signed PO.

Data Santander is a founder-led company. With no external funding. We're solely customer-funded. You'll work directly with our development team. And we're able and willing to customize TrialTwin™ to better fit your organization's unique needs.

Contact Us

José C. Lacal | +34 (674) 88 17 52 | Jose.Lacal@DataSDR.com

Our End Goal: Accelerate Clinical Trial Data Flow

We believe the way data is handled during a clinical trial needs to be re-designed from the ground up. That's our goal with TrialTwin™: to incrementally re-design current processes.

Status Quo

The diagram below presents an overview of the steps currently involved in designing and handling data during a clinical trial. There are multiple data transcoding steps, many manual processes, and a large number of very brittle SAS macros that need to be customized for different trials.

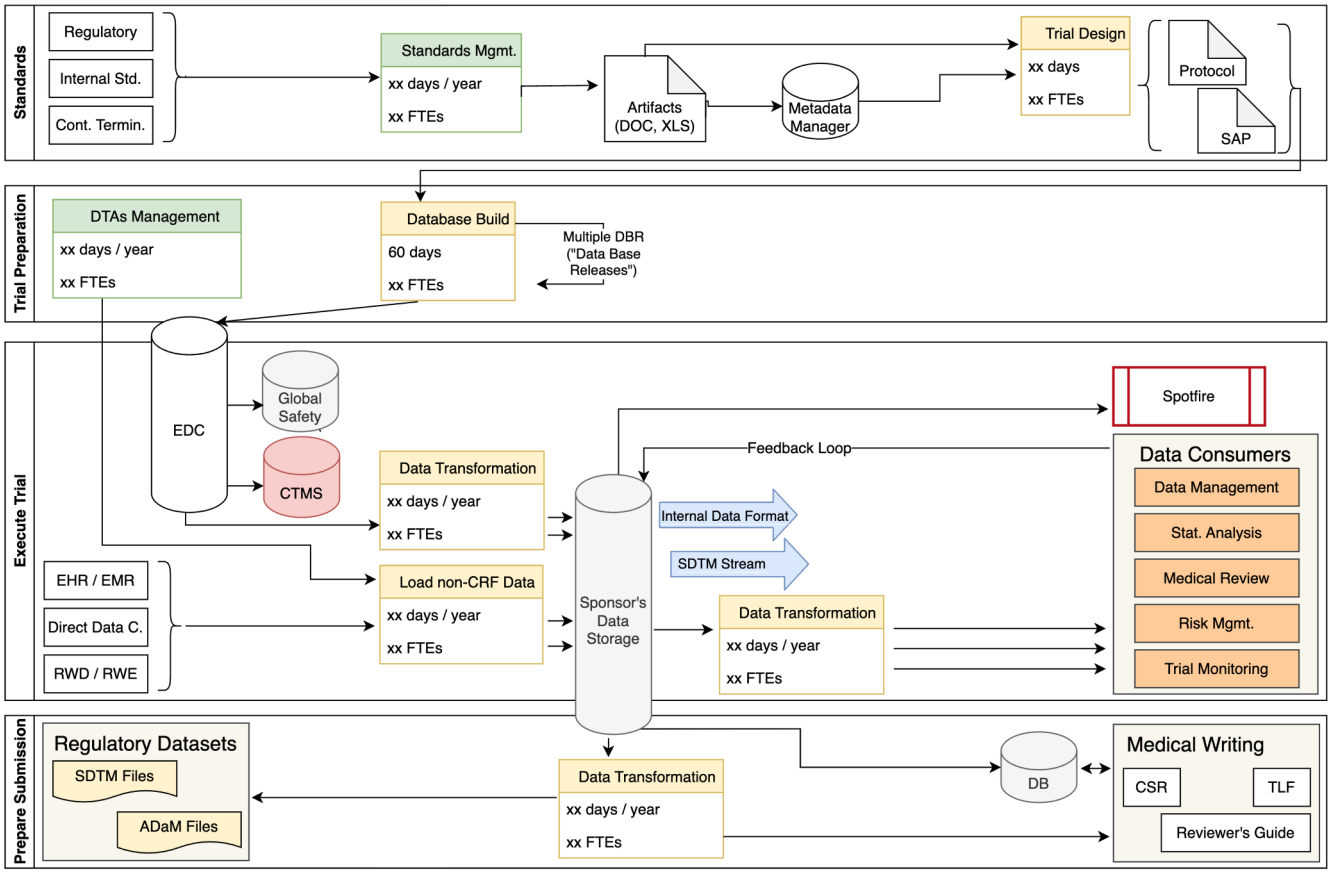


Figure 1: Current state of data flows in a clinical trial.

Many software packages in use today provide a thin layer of digitalization on top of otherwise manual, paper-based processes. Most stakeholders in the clinical trial space are awash in "digital paper."

Instead of trying a novel way to optimize innately inefficient processes, the TrialTwin™ approach eliminates many of the existing processes by building a digital-at-the-start flow.

The diagram below outlines TrialTwin™ as a centralized, metadata-driven platform that eliminates most data transcoding processes and SAS macros.

TrialTwin™ leverages the concept of a "Computable Protocol" (such as that defined by CDISC's USDM) to drive all the downstream processes.

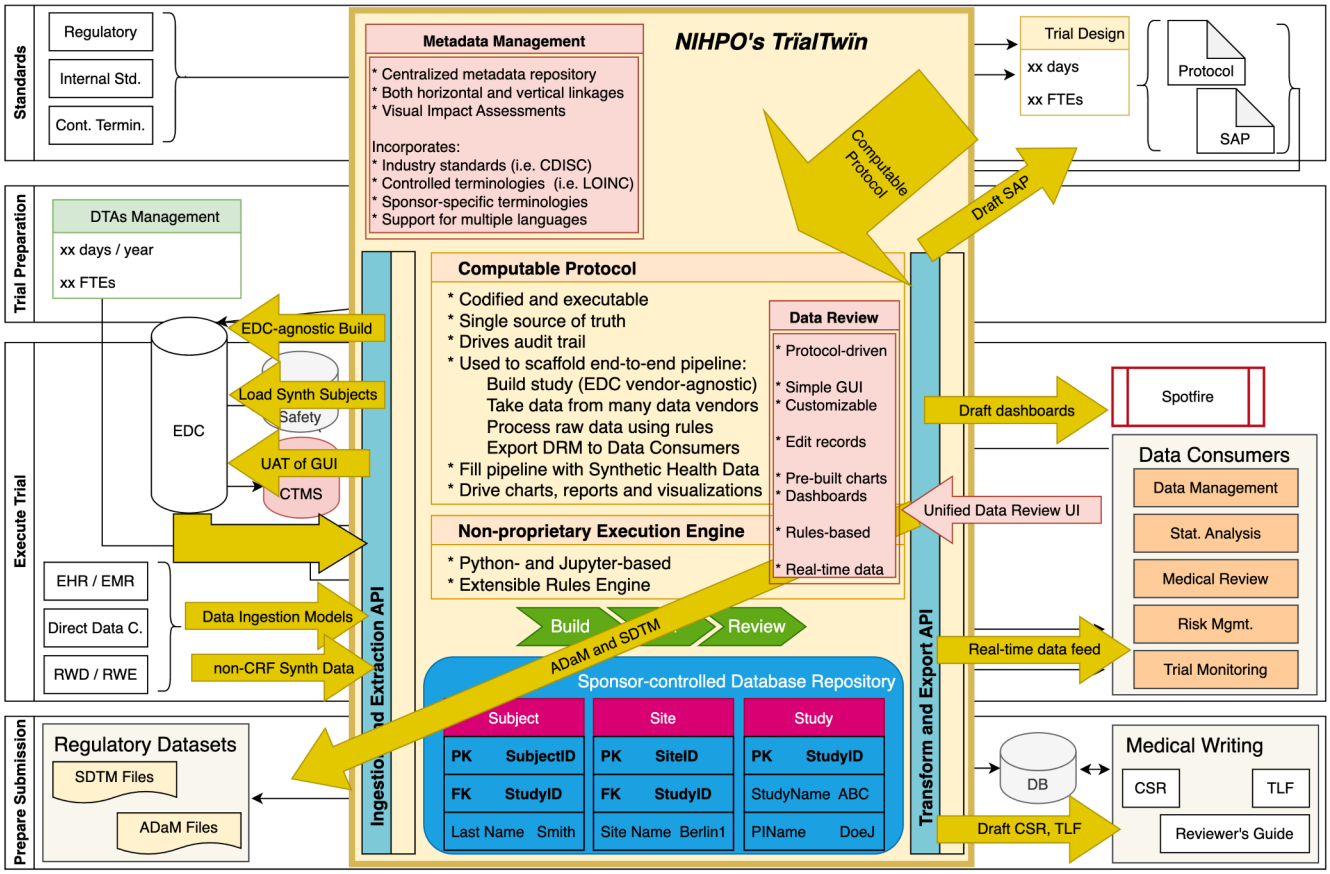


Figure 2: Desired end state of optimized, digital-first data flow.

Metadata Repository

The TrialTwin™ Metadata Repository (“MDR”) module is a visual, integrated manager of data standards & terminologies used in the Life Sciences market.

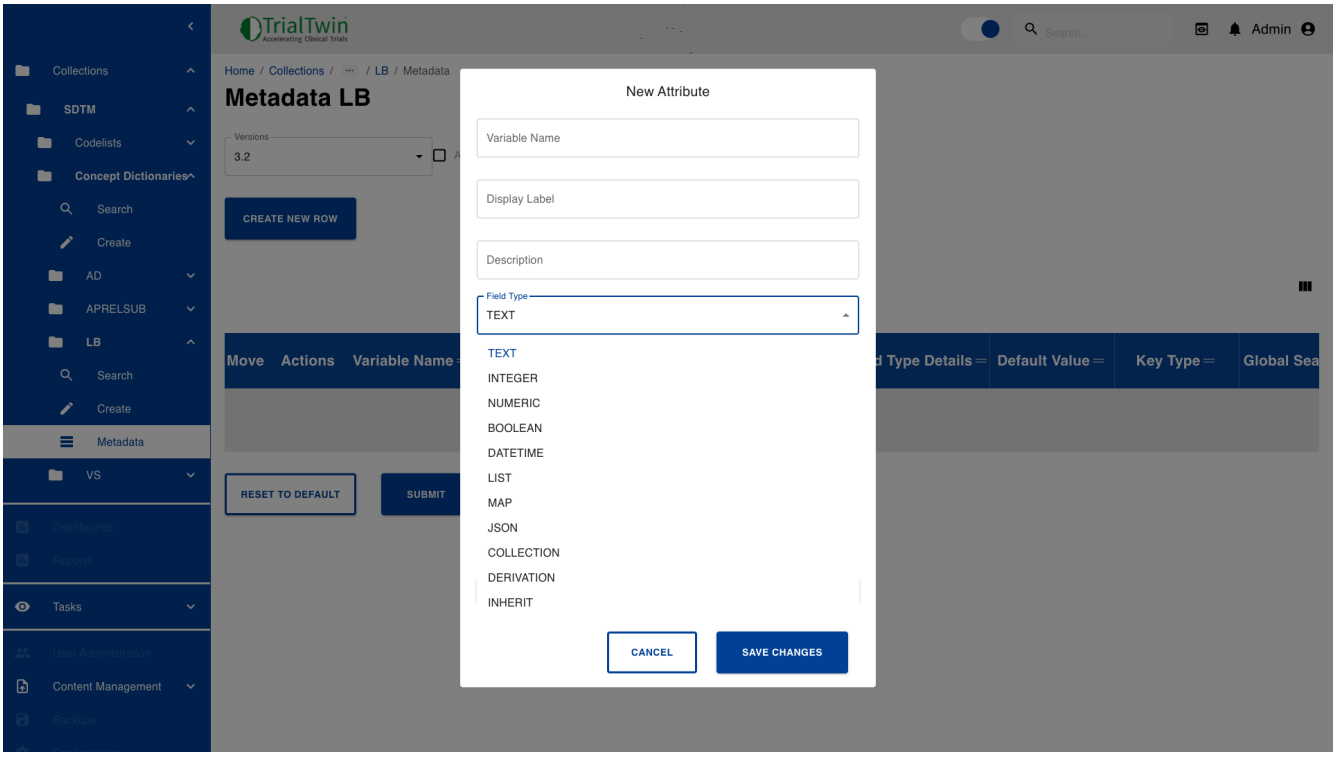


Figure 3: Metadata Repository ("MDR")

These are some of the benefits to sponsors of having access to our MDR:

Table 1: Benefits of MDR.

Benefit	Description
Saves time	No more web searching or looking for Excel trackers.
Centralize Data	Manage terminologies and dictionaries for use across functional areas. With full audit trail.
Item Connectivity	Display and manage relationships between terms, concepts, forms, libraries, up to studies.
User Roles	Access for Searcher, Editor, Approver roles.
Task Pool	Workflow engine for editing and approval.
Impact Assessment	Analyze ripple effect of changes to items.

The TrialTwin™ Metadata Repository (“MDR”) module provides these features:

Table 2: MDR features.

Feature	Description
Hosted	Instant-on, nothing to install and maintain on client’s side.
Pre-loaded industry standards	CDISC; LOINC; and SNOMED-CT are loaded and ready for use. Additional industry standards / terminologies can be loaded on request.
Load client-specific metadata	Client-specific metadata will be loaded to system.
Strict role management	Various roles with different level of permissions to perform actions
Single source of truth	With real time management of controlled terminology and concept dictionaries including audit trail.
Items Connectivity	Display of metadata connectivity in UI.
Impact Assessment	CT update impact to downstream metadata and action based on DISC or sponsor-specific codes.
Task pool management	Mechanism to identify, create, assign, track, and close open Tasks (additions, edits, retirements).
Modern User Interface	Comprehensive search functionality and UI visualization flexibility to rename column names and reposition column as needed.
Quality control process	Workflow for a task and data validation for record creation and maintenance.

Custom Enhancements

Client-specific enhancements can be developed, please ask us for a cost estimate.

Protocol Manager

The TrialTwin™ Protocol Manager module allows sponsors to capture and manage Protocol data in digital format. And to connect the Protocol with sponsor data, once submitted.

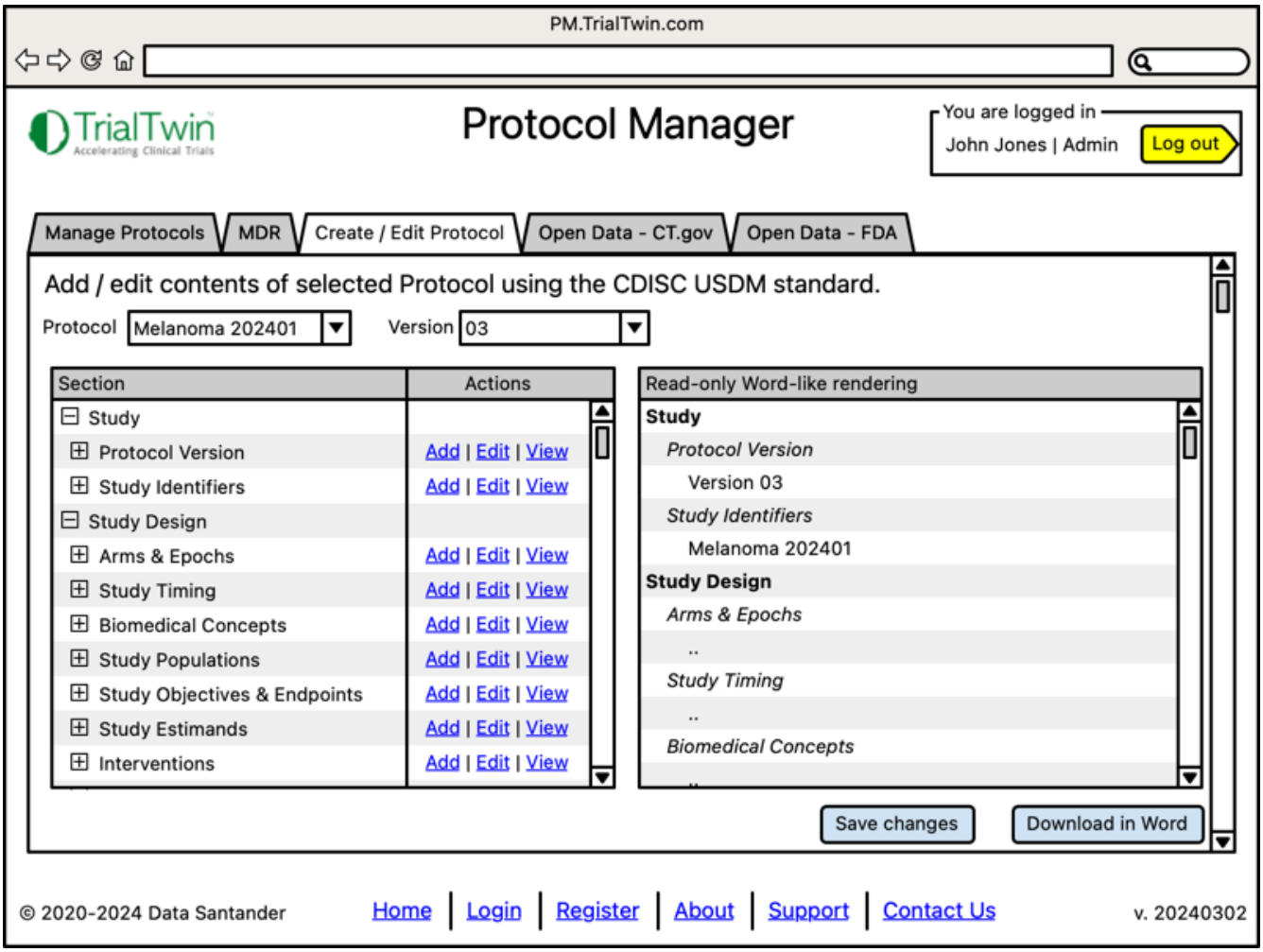


Figure 4: Add / Edit Protocol.

Table 3: Benefits of Protocol Manager

Benefit	Description
End-to-end Digital	Write once, read often across all trial phases. Link to EDC, CTMS
Integrated MDR	Tied to controlled terminologies during protocol design process. Full audit trail.
CDISC USDM	Data is stored natively using USDM standard.

Benefit	Description
Pre-loaded Content	Users access previous Protocols, SAPs, ICFs from ClinicalTrials.gov Used as reference to accelerate protocol creation, management. Reducing user errors.

The CDISC USDM standard is currently supported, We can enhance the system to also support ICH M11 if requested.

User can use the Protocol Manager to view previous protocols and application documents for related products. The Protocol Manager gives medical writer a view into the entire life cycle of previously-approved products.

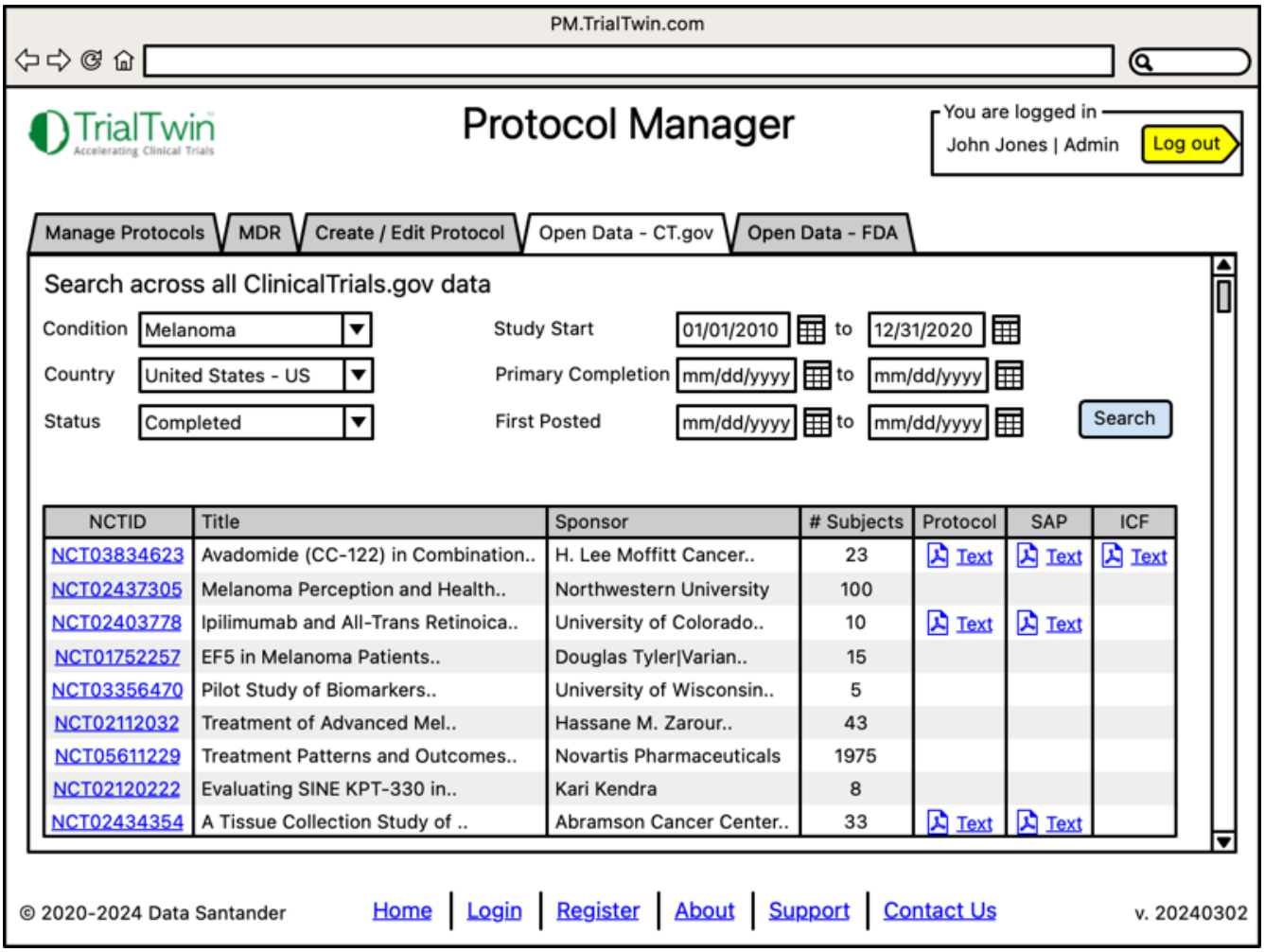


Figure 5: Open Data - ClinicalTrials.gov

In the above wireframe user can search for and review previous studies targeting the same indication. If available, user can see the actual Protocol, SAP, and/or ICF of each of those previous studies.

In the wireframe below user can review the full application documents, Orange Book, and SPL of previously-approved products.

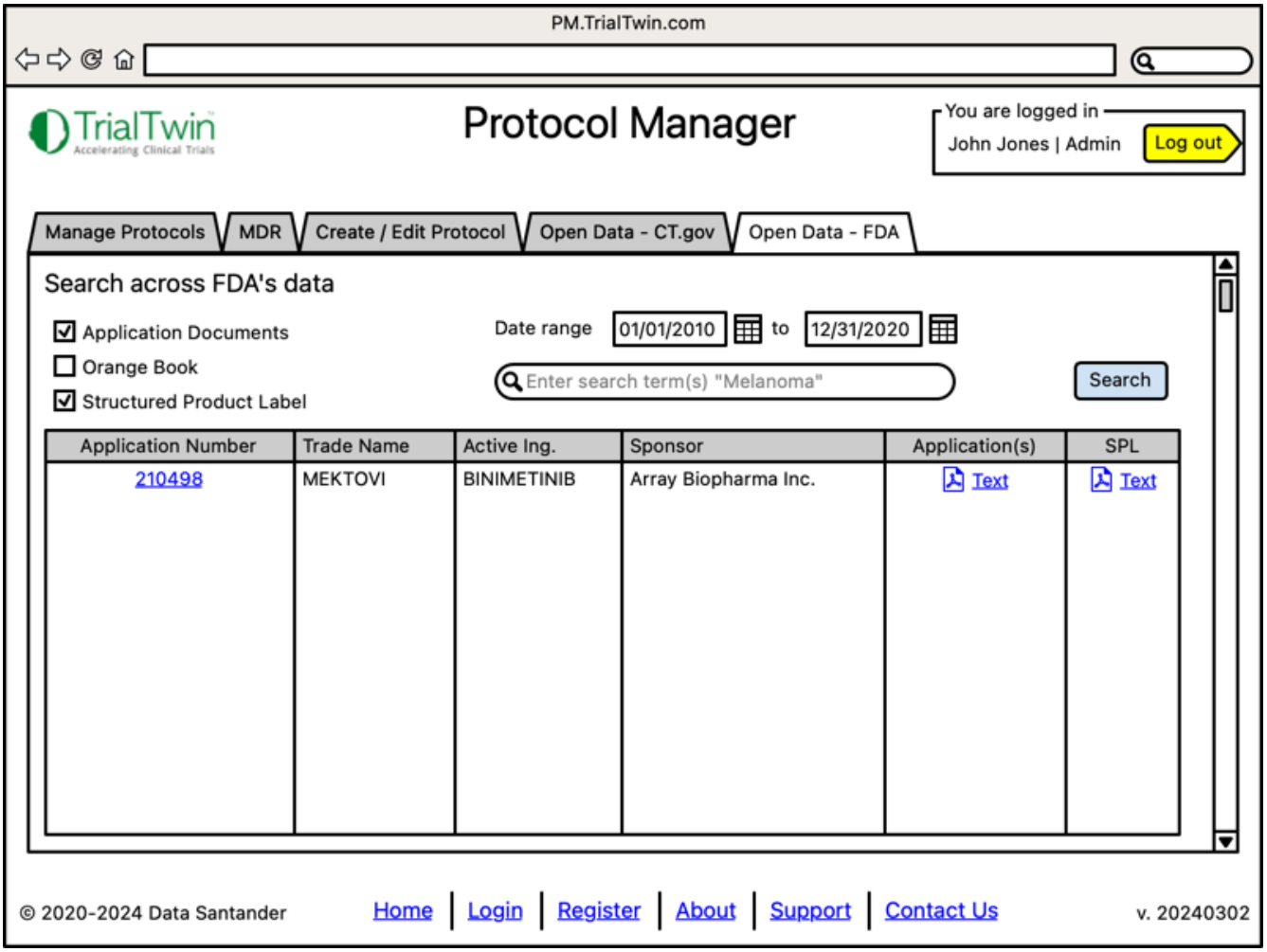


Figure 6: Open Data - FDA

Synthetic Health Data

The TrialTwin™ Synthetic Health Data module allows user to generate realistic yet fake data for testing purposes. From the Protocol Manager, user can create study-specific test data in minutes.

Sponsor obviously has access to massive amounts of real data. But there are many use cases (such as when dealing with outside vendors / contractors) when there is a need to request a data waiver. That process usually takes time, and there are limits and restrictions placed on the data.
With Synthetic Health Data the test data is immediately available to all stakeholders, in large numbers, and without any restrictions. Thus accelerating the execution of desired tests and validation processes.

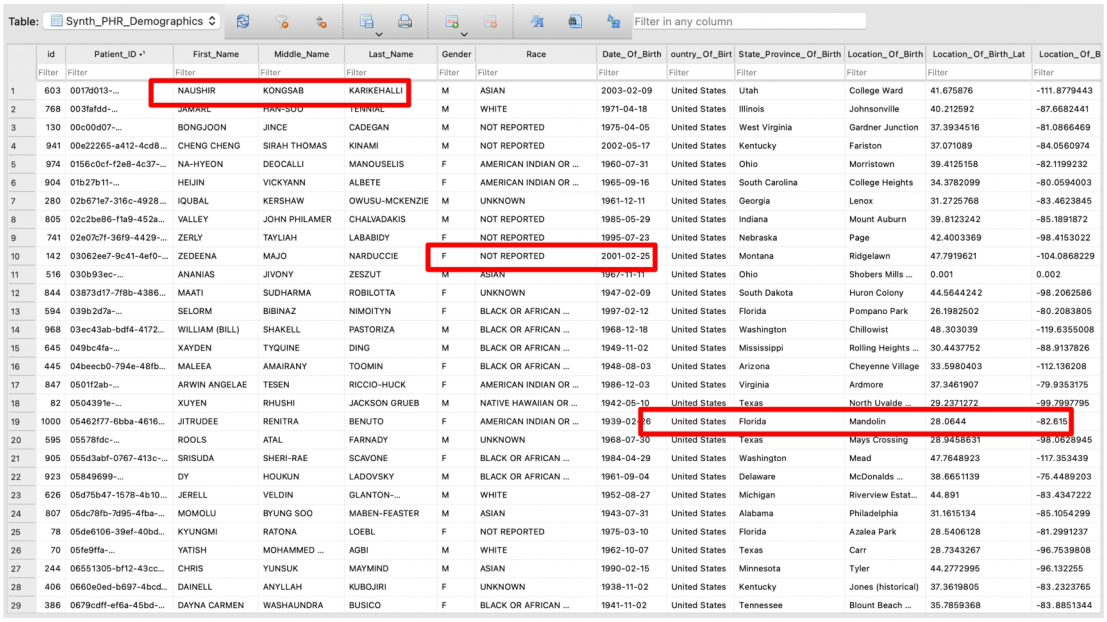


Figure 7: Synthetic Health Data - Demographics

Table 4: Benefits of Synthetic Health Data

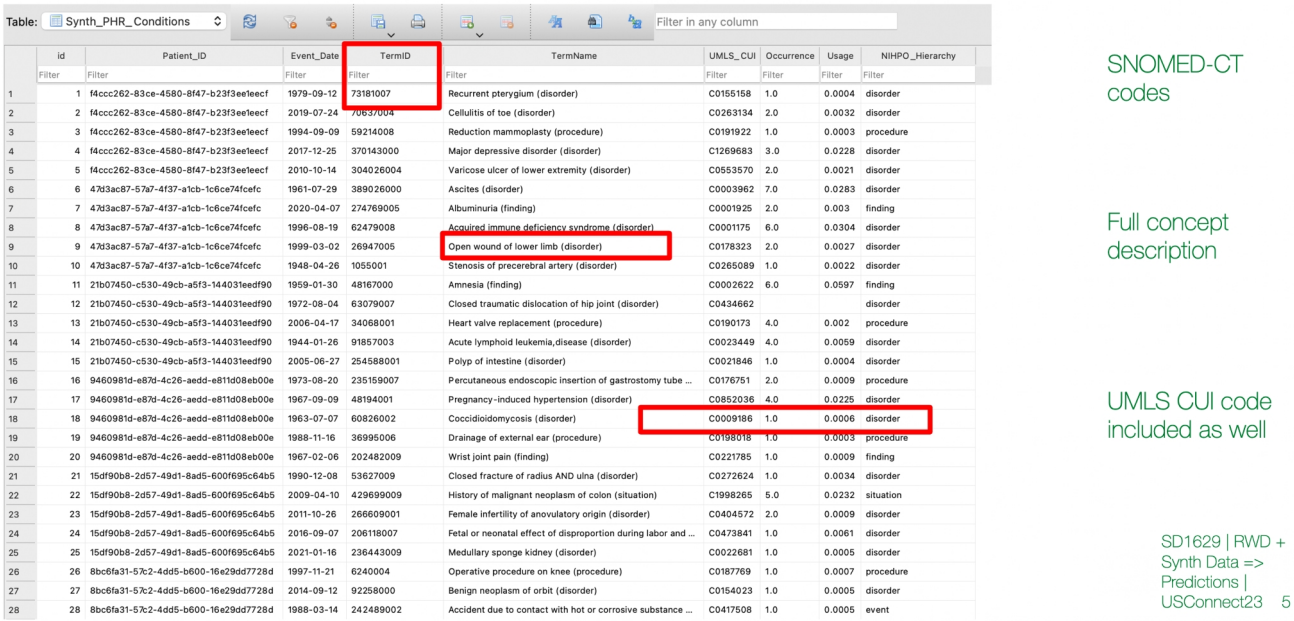
Benefit	Description
Realistic yet Fake	Data is realistic (using terminologies like LOINC, SNOMED) yet totally fake. No de-anonymization risk is a significant benefit.
Test Earlier	Synthetic Health Data helps teams to test code and processes earlier. And cover corner cases.
Fast, Low-cost	Large quantities of data can be created quickly and at low cost. Use

Benefit	Description
	multiple times.
CDISC Formats	SDTM-formatted data is available now. ADaM datasets can be easily generated on request.

The Synthetic Health Data can also be used to test tools for changes in rules. As well as creating tools and visualizations to make users' work more effective

Furthermore, as sponsors are expected to ingest more Real World Data ("RWD"), test data can be used to test and validate ingestion tools. For example, large amounts of test data in the FHIR format will be useful to validate of current tools can ingest these new types of data.

Synthetic Health Data can be used to safely and quickly develop new data management activities.



	id	Patient_ID	Event_Date	TermID	TermName	UMLS_CUI	Occurrence	Usage	NIHPO_Hierarchy
1	1	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	1979-09-12	73181007	Recurrent pterygium (disorder)	C0155158	1.0	0.0004	disorder
2	2	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2019-07-24	70637004	Cellulitis of toe (disorder)	C0263134	2.0	0.0032	disorder
3	3	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	1994-09-09	59214008	Reduction mammoplasty (procedure)	C0191922	1.0	0.0003	procedure
4	4	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2017-12-25	370143000	Major depressive disorder (disorder)	C1269683	3.0	0.0228	disorder
5	5	f4ccc262-83ce-4580-8f47-b23f3ee1eecf	2010-10-14	304026004	Varicose ulcer of lower extremity (disorder)	C0553570	2.0	0.0021	disorder
6	6	47d3ac87-57a7-4f37-a1cb-1c6ce74fcefc	1961-07-29	389026000	Ascites (disorder)	C0003962	7.0	0.0283	disorder
7	7	47d3ac87-57a7-4f37-a1cb-1c6ce74fcefc	2020-04-07	274769005	Albuminuria (finding)	C0001925	2.0	0.003	finding
8	8	47d3ac87-57a7-4f37-a1cb-1c6ce74fcefc	1996-08-19	62479008	Acquired immune deficiency syndrome (disorder)	C0001175	6.0	0.0304	disorder
9	9	47d3ac87-57a7-4f37-a1cb-1c6ce74fcefc	1999-03-02	26947005	Open wound of lower limb (disorder)	C0178323	2.0	0.0027	disorder
10	10	47d3ac87-57a7-4f37-a1cb-1c6ce74fcefc	1948-04-26	1055001	Stenosis of precerebral artery (disorder)	C0265089	1.0	0.0022	disorder
11	11	21b07450-c530-49cb-a5f3-144031eedf90	1959-01-30	48167000	Amnesia (finding)	C0002622	6.0	0.0597	finding
12	12	21b07450-c530-49cb-a5f3-144031eedf90	1972-08-04	63079007	Closed traumatic dislocation of hip joint (disorder)	C0434662			disorder
13	13	21b07450-c530-49cb-a5f3-144031eedf90	2006-04-17	34068001	Heart valve replacement (procedure)	C0190173	4.0	0.002	procedure
14	14	21b07450-c530-49cb-a5f3-144031eedf90	1944-01-26	91857003	Acute lymphoid leukemia,disease (disorder)	C0023449	4.0	0.0059	disorder
15	15	21b07450-c530-49cb-a5f3-144031eedf90	2005-06-27	254588001	Polyp of intestine (disorder)	C0021846	1.0	0.0004	disorder
16	16	9460981d-e87d-4c26-aedd-e811d08eb00e	1973-08-20	235159007	Percutaneous endoscopic insertion of gastrostomy tube ...	C0176751	2.0	0.0009	procedure
17	17	9460981d-e87d-4c26-aedd-e811d08eb00e	1967-09-09	48194001	Pregnancy-induced hypertension (disorder)	C0852036	4.0	0.0225	disorder
18	18	9460981d-e87d-4c26-aedd-e811d08eb00e	1963-07-07	60826002	Coccidioidomycosis (disorder)	C0009186	1.0	0.0006	disorder
19	19	9460981d-e87d-4c26-aedd-e811d08eb00e	1988-11-16	36995006	Drainage of external ear (procedure)	C0198018	1.0	0.0003	procedure
20	20	9460981d-e87d-4c26-aedd-e811d08eb00e	1967-02-06	202482009	Wrist joint pain (finding)	C0221785	1.0	0.0009	finding
21	21	15df90b8-2d57-49d1-8ad5-600f695c64b5	1990-12-08	53627009	Closed fracture of radius AND ulna (disorder)	C0272624	1.0	0.0034	disorder
22	22	15df90b8-2d57-49d1-8ad5-600f695c64b5	2009-04-10	429699009	History of malignant neoplasm of colon (situation)	C1998265	5.0	0.0232	situation
23	23	15df90b8-2d57-49d1-8ad5-600f695c64b5	2011-10-26	266609001	Female infertility of anovulatory origin (disorder)	C0404572	2.0	0.0009	disorder
24	24	15df90b8-2d57-49d1-8ad5-600f695c64b5	2016-09-07	206118007	Fetal or neonatal effect of disproportion during labor and ...	C0473841	1.0	0.0061	disorder
25	25	15df90b8-2d57-49d1-8ad5-600f695c64b5	2021-01-16	236443009	Medullary sponge kidney (disorder)	C0022681	1.0	0.0005	disorder
26	26	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	1997-11-21	6240004	Operative procedure on knee (procedure)	C0187769	1.0	0.0007	procedure
27	27	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	2014-09-12	92258000	Benign neoplasm of orbit (disorder)	C0154023	1.0	0.0005	disorder
28	28	8bc6fa31-57c2-4dd5-b600-16e29dd7728d	1988-03-14	242489002	Accident due to contact with hot or corrosive substance ...	C0417508	1.0	0.0005	event

Figure 8: Synthetic Health Data - Conditions.

Our Synthetic Health Data helps your organization to
* Test earlier: start testing as soon as the first drafts of the Protocol are ready
* Test often: re-run the test data through all your processes as often as necessary
* Test fully: our data is large enough to cover all possible corner cases and possible choices

Open Data Repository

Users can track every drug, manufacturer, and the individuals behind the documents submitted to government agencies.

The Repository allows users to trace an approved drug's lifecycle:
* Starting with chemical compounds (NLM's PubChem)
* Through clinical trials (ClinicalTrials.gov, WHO's ITPR)
* Documentation on regulatory pathway (IND, NDA, etc.)
* Reported adverse events (FDA's FAERS, MAUDE)
* Manufacturer payments to providers (HHS' OpenPayments)
* Medicare reimbursement data (CMS' Provider Utilization, Payment Data)

Figure 9: Open Data Repository.

The Repository offers users a 360 degrees view of each previously-cleared drug.

Open Data to Train AI, ML

The data stored in the Open Data Repository module can be used to train AI / ML models. Think about these data as the "Ground Truth" of what has happened in the pharma space in the US for the last 30 years.

Users can leverage Life Sciences-specific Open Data including:
• 40,000+ Protocols, SAPs, ICFs
• over 70,000 FDA application files
• 110,000 full FDA labels ("SPL")

Users can use this data to train their Models with the text extracted from all those documents, containing 600+ million words.

We can also include additional Open Data from other US agencies, including:
• CMS – Medicare
• HHS – healthcare
• NLM – references

About TrialTwin™

TrialTwin™ is developed by:

Data Santander, SL
San Fernando 16, 6C
39010 Santander, Cantabria
Spain

Data Santander is a small software development company that provides outside, unbiased suggestions and innovative approaches to tackle data management challenges for our Life Sciences clients.

Compared with larger vendors, at Data Santander we are:
* ***fast*** – clients work directly with our developers to reduce cycle times
* ***fearless*** – we want to eliminate, rather than improve, most processes in clinical data management
* ***flexible*** – we're able to make progress while requirements are defined on the fly
* ***friendly*** – we love what we do, and we bring our quirky personalities to the job
* ***fun*** – life's too short so we make working with us fun and enjoyable

Contact:

José C. Lacal, CTO
Jose.Lacal@DataSDR.com
+34 (674) 88 17 52