The Eras of Clinical Data: Fearless Collaboration Between CDM and CDS for a Bejeweled Future

Laura Fazio; Director, Clinical Data Management

Melanie Hullings; Director Clinical Data Analytics & Programming

July 25, 2024

Formation Bio

Laura Fazio

Director, Clinical Data Management

Melanie Hullings

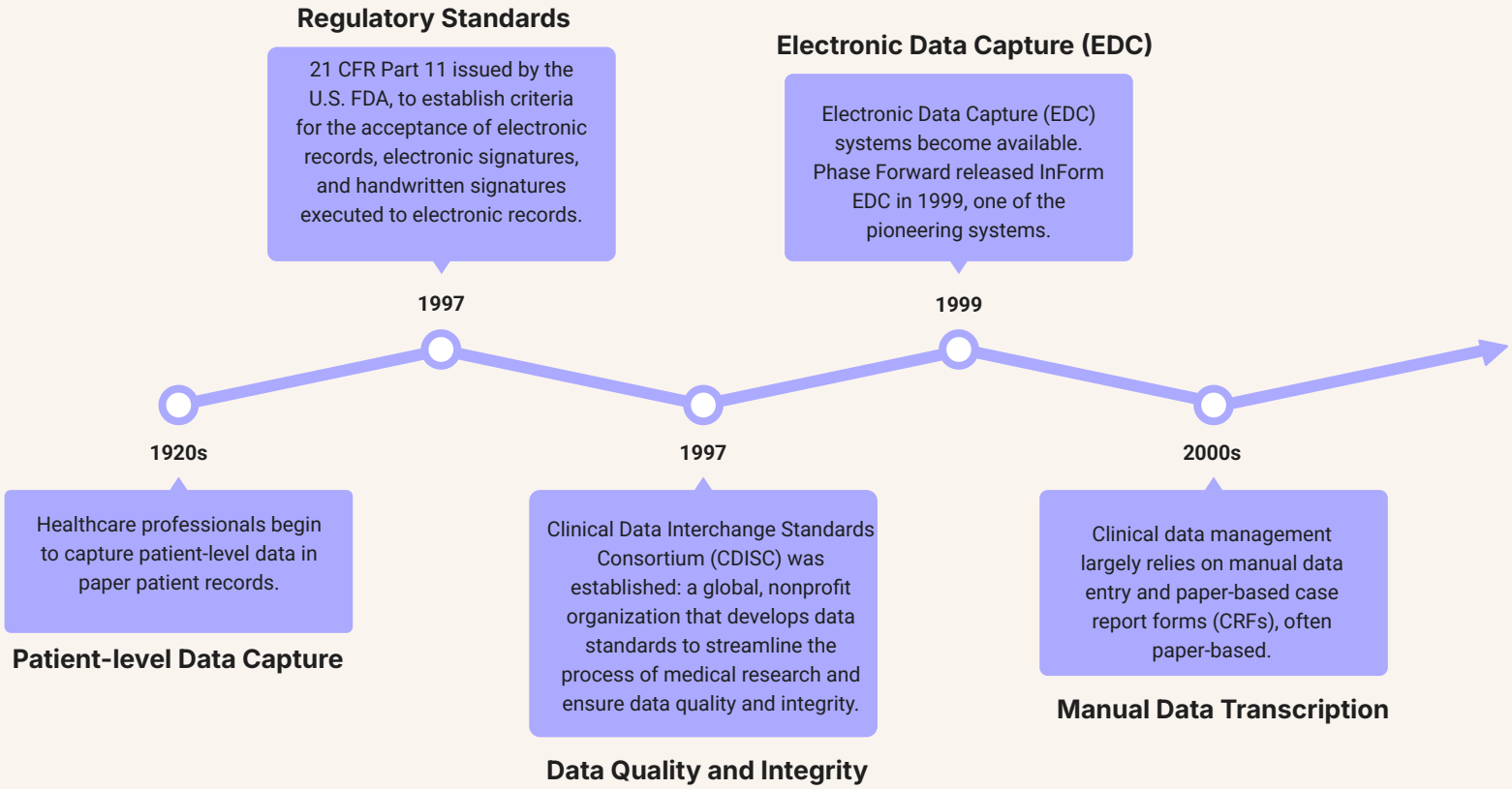
Director, Clinical Data Analytics & Programming

Agenda

- Where have we been?
- Where are we now?
- Where are we going?

Where have we been?

History of Clinical Data Management (CDM)

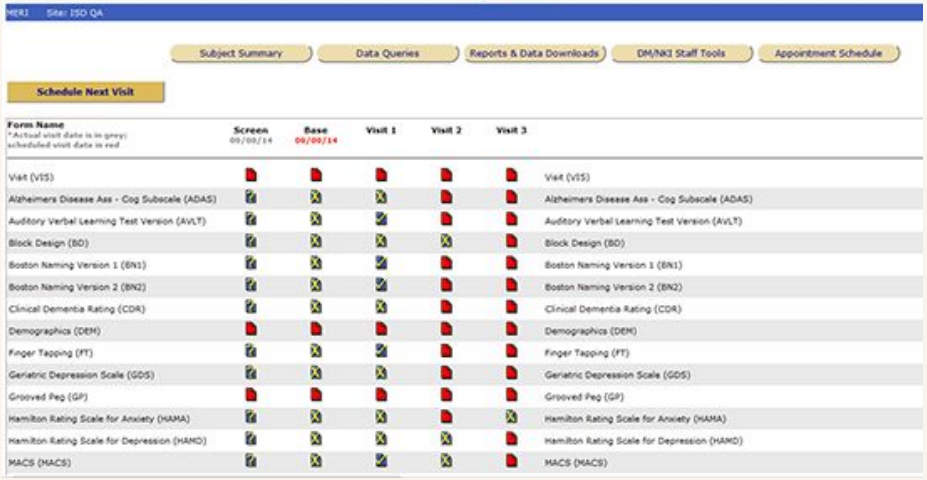


Reference: https://novotech-cro.com/faq/clinical-data-management-past-present-and-future

Early success

Early EDC success factors:

- Improved data quality
- Real time data access
- Increased efficiency
- Regulatory compliance



Positive Outcomes:

➔ **Case Study 1**: A pharmaceutical company conducting a multi-center clinical trial reported a 30% reduction in data entry errors and a 20% decrease in the time required to lock the database after the last patient visit.

➔ **Case Study 2**: An academic research institution conducting a large-scale epidemiological study noted improved data quality and faster access to preliminary results, facilitating more timely publications and presentations.

Reference: Reference: Kush, R. D., & Bleicher, P. A. (2001). Electronic data capture and clinical trials: the complexity of implementation in the USA. *Drug Information Journal*, 35(4), 1241-1249.

Challenges

Introduction of EDC was positive progress but also brought challenges:

- Technology barriers
- Cost
- User acceptance
- Training and support
- Regulatory concerns

Reference: Reference: Kush, R. D., & Bleicher, P. A. (2001). Electronic data capture and clinical trials: the complexity of implementation in the USA. *Drug Information Journal*, 35(4), 1241-1249.

Digital transformation

- Improved data quality and accuracy
  - Transition to Electronic Data Capture (EDC) systems streamlined data collection and reduced errors
  - EDC systems have been shown to reduce data entry errors by up to 80% compared to manual methods
- Efficiency and speed
  - Streamline workflow and reduces time for data entry, cleaning, and analysis
  - Increasing overall trial process
- Advanced systems
  - Electronic EHRs: Affordable Care Act required electronic source that demonstrated meaningful use
  - Proliferation of electronic data capture vendor systems beyond EDC

Reference: The Evolution of Clinical Data Management Systems - PharmiWeb. Transforming Clinical Data Management Systems for Biotechnology Trials - Pharmaphorum.

Data proliferation

Drivers	Challenges	Solutions
• Digital transformation • Internet of things • Cloud computing • Big data analytics • Social medial and user generated content • Compliance and regulation	• Data management • Storage • Security • Privacy • Quality	• Data governance • Advanced analytics • Scalable infrastructure • Data minimization • Security measures • Training and awareness

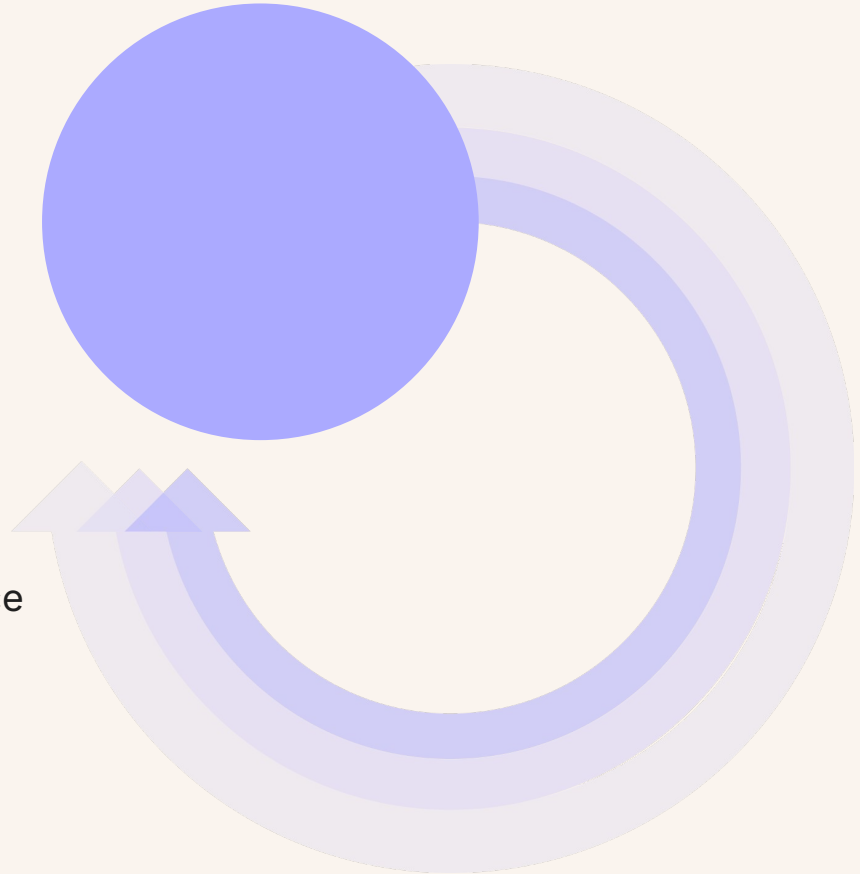
Where are we now?

Overcoming technological challenges for Clinical Data Managers (CDMs):

	Challenge	Solution
Data Integration and Interoperability	Integrating data from multiple sources and ensuring interoperability.	Utilizing data integration platforms and standards like CDISC (Clinical Data Interchange Standards Consortium) to harmonize data across different systems.
Data Quality and Consistency	Maintaining high data quality and consistency.	Implementing robust data validation and cleaning protocols, automated quality checks, and regular data audits.
Data Capture Systems	Effective use of EDC systems.	Comprehensive training on EDC systems, selecting user-friendly EDC tools, and ensuring continuous support and troubleshooting.
Regulatory Compliance	Adhering to regulatory requirements (e.g., FDA, EMA).	Keeping updated with regulatory guidelines, utilizing compliance-checking software, and engaging in regular training on regulatory standards.
Data Security	Ensuring data security and privacy.	Implementing advanced encryption methods, secure access controls, and regular security audits to protect sensitive clinical data.

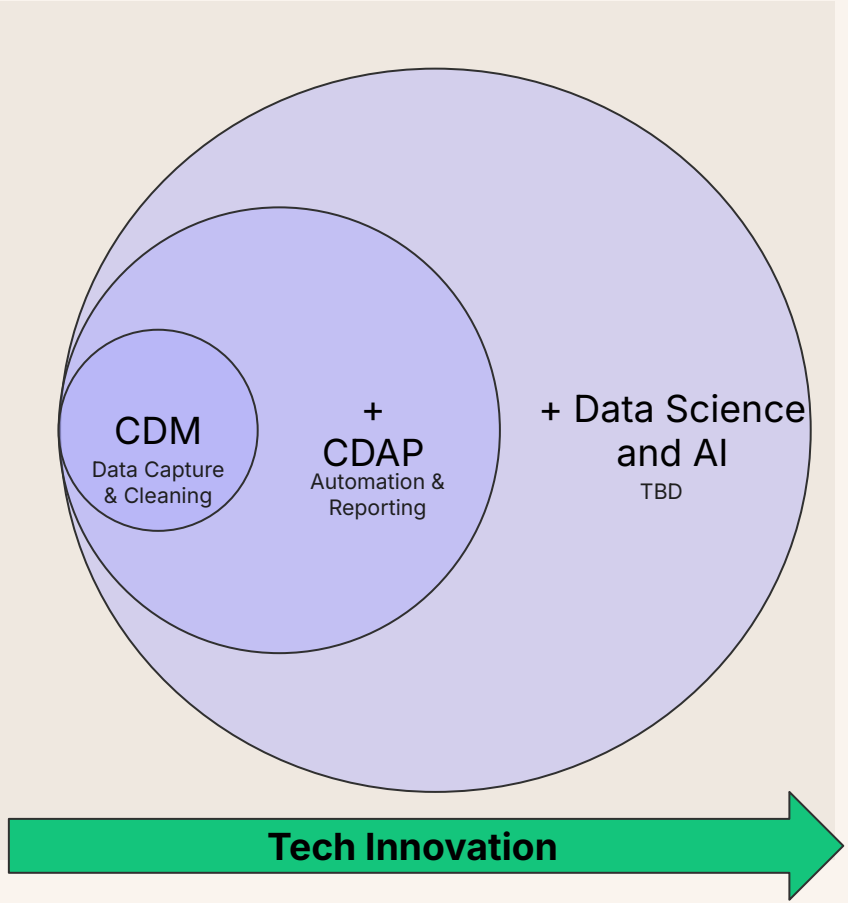
Moving on from the past

- **People and process**
  - Maintain focus on data quality and integrity
  - Continuous training and education
  - Regulatory compliance and data security
  - Collaboration and partnership
- **Innovative data capture**
  - Patient-centric data capture
  - Agile and adaptive trial designs
- **Adopt advanced technologies**
  - Big data and analytics
  - Leverage programming, including open-source



Adopting advanced analytics in CDM requires long-term vision

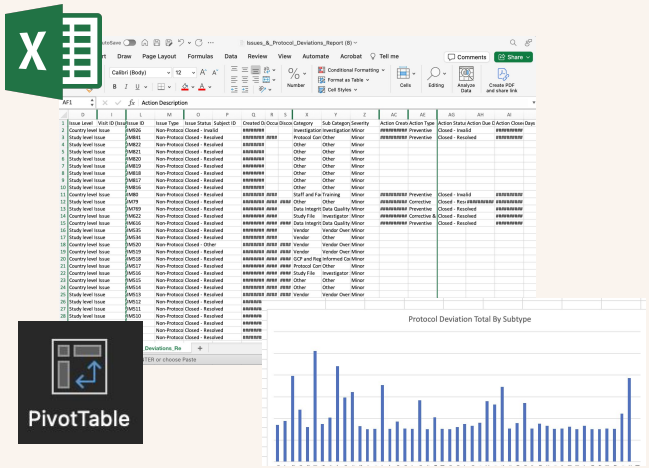
- Started with a vision to automate clinical trial execution and data management
- Over the last 3 years, we have built out technical & operational infrastructure to support automation of clinical trial data cleaning and reporting
- Clinical Data Analytics and Programming (CDAP) team evolved from Clinical Data Management (CDM)
- We are now expanding our infrastructure and capabilities for data science to support end-to-end drug development



Programming methods can automate CDM processes

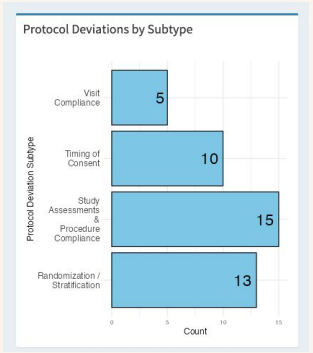
Export data

5+ Steps



VS

1 Click



A new programming function was created at Formation Bio from CDM

Typical Biometrics Team

- Biostatistics
- Clinical Data Management
- Statistical Programming

New Function

Clinical Data Analytics & Programming (CDAP)

- Advanced Analytics
- Data Science

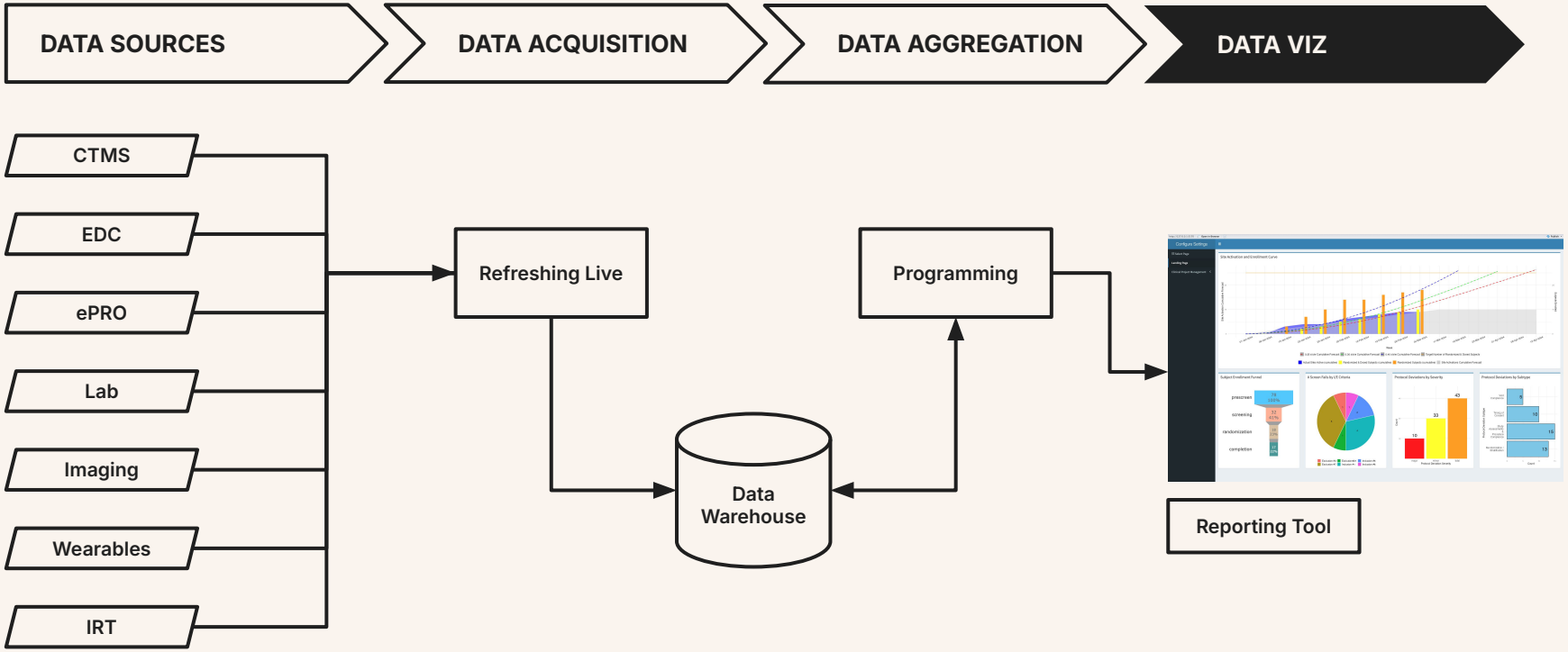
Collaboration Model

	Clinical Data Manager	Clinical Data Analyst
Data Collection and Quality Assurance	Ensures accurate data collection and quality.	Provides input on data collection needs for analysis, identifies potential data issues, and suggests quality control measures.
Database Design and Management	Designs and maintains the clinical trial database.	Advises on database structure to facilitate data extraction and analysis.
Data Cleaning and Preparation	Performs initial data cleaning to ensure data consistency and completeness.	Conducts further data preprocessing to prepare data for analysis, identifying and addressing any anomalies.
Data Analysis and Interpretation	Provides cleaned and validated data sets.	Analyzes the data to extract insights, interprets the results, and ensures that findings are statistically sound.
Regulatory Compliance	Ensures that all data management processes comply with regulatory standards.	Ensures that analytical methods and reporting comply with regulatory requirements.
Communication and Reporting	Documents and reports on data management activities.	Prepares and presents analysis reports to stakeholders, providing interpretations and recommendations based on the data.
Continuous Improvement	Identifies areas for improvement in data management processes.	Suggests enhancements in data collection and analysis techniques based on latest methodologies and technologies.

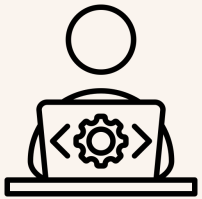
Overcoming technological challenges for Clinical Data Analysts (CDAs):

	Challenge	Solution
Real-time Data Processing	Processing real-time data efficiently.	Adopting real-time data processing frameworks, streamlining data warehouse, and connecting to downstream tools to analyze data as it arrives.
Big Data and Analytics	Managing and analyzing large volumes of complex data.	Leveraging advanced analytics tools, big data platforms, and machine learning algorithms to handle and derive insights from large datasets.
Data Visualization	Effective data visualization and reporting.	Utilizing sophisticated data visualization tools (e.g., Tableau, R, Python) to create clear, impactful, and interactive visual representations of data.
Advanced Statistical Methods	Implementing advanced statistical and computational methods.	Continuous education and training in the latest statistical techniques, using statistical software (e.g., SAS, R), and collaborating with biostatisticians.
Data Standardization	Standardizing data from diverse sources.	Applying data standardization techniques and tools and using common data models and ontologies to ensure consistency.

Data aggregation is crucial to automating programming



Results of embedding programming on the clinical study team



Automate away hours of manual data cleaning and review to identify potential data issues, including deviations

Improve speed and accuracy of clinical insights so the study team can focus on resolving issues rather than finding them

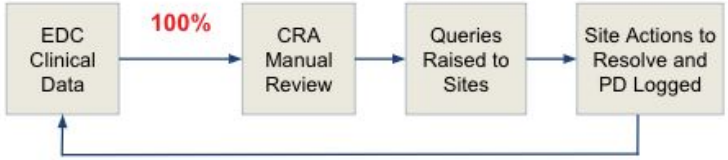
Improve collaboration with the entire clinical study team having access to real-time reports

Example: R program to detect protocol deviations

Process Improvement

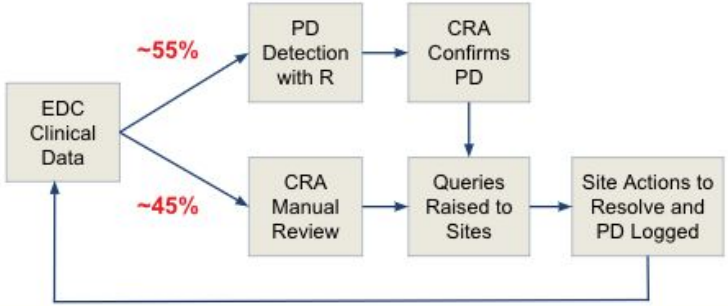
Industry Standard:

- relies on manual review by CRAs to identify deviations from clinical trial protocol requirements



Program Assisted:

- more efficient and ensures holistic issue identification



Results

- Script identified 33 PDs that were not identified by CRAs
- 63% of PDs identified by script were false positives
- False positives have value in identifying data abnormalities

Conclusion

- R scripts can be used to programmatically detect issues in clinical trial datasets including PDs and participant eligibility
- Programmatic issue detection has the potential to decrease manual review burden of study teams, as well as cost, and improve issue identification and clinical trial data quality overall
- Future directions include refining scripts to make issue identification more accurate and expanding the number of data issues that scripts are detecting

CRA = Clinical Research Assistant
EDC = Electronic Data Capture
PD = Protocol Deviation

Where are we going?

Learning from the past to embark on new eras

- Learn from mistakes
  - Resistance to adoption
  - Data & team silos
- Learn from successes
  - Quality
  - Efficiency

Stage is set for AI to transform CDM but…

we need to proceed with caution!

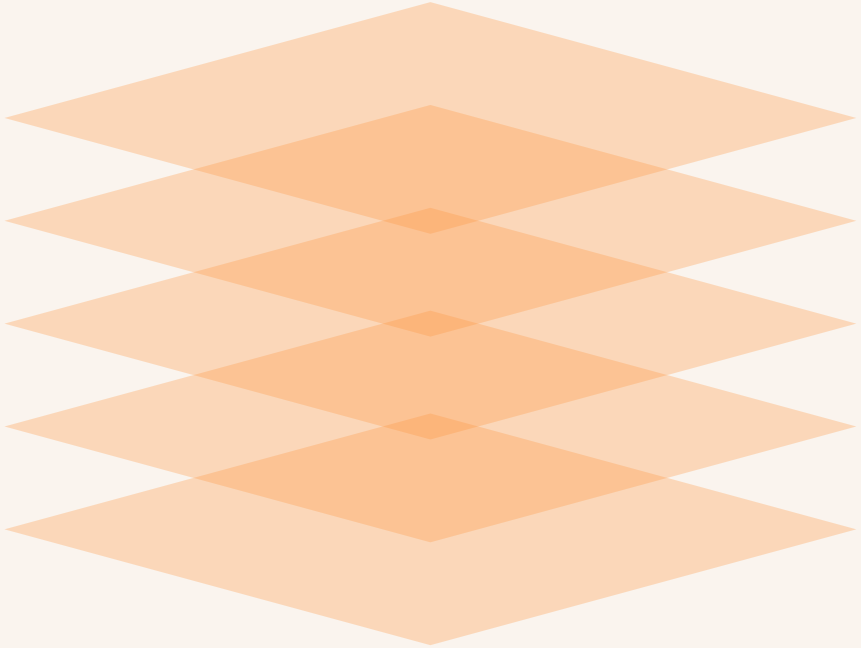
Advanced Generative Pre-trained Transformer (GPT) tools overview

Tool	Context of use	Choose this tool when…	Data retention	Example(s)
ChatGPT	Conversation with the internet of information	There is no need to upload information (you could ask the general public)	N/A	➔ Questions about FDA guidelines ➔ Specs for a single edit check
Enterprise Custom GPT	A tool trained on tasks and topics by combining specific instructions, knowledge, and access to systems	This is a repeatable task or topic trained on specific information (you could ask the intern)	Uploaded data is transferred to OpenAI and retained and/or used to train the model as defined in the agreement	➔ Questions about a protocol ➔ Generate potential edit checks from SDS
Application Programming Interface (API)	Used by an internal applications to create an assistant that can have access to specific information and instructions and respond to user queries within the application	You need a bot within an application such as Slack trained on specific information (you could ask a previous intern hired as an assistant)	Data is transferred to OpenAI via an API and retained and/or used to train the model as defined in the agreement	➔ Questions about company HR policies you can ask a Teams channel

Reference: https://help.openai.com/en/articles/8673914-gpts-vs-assistants

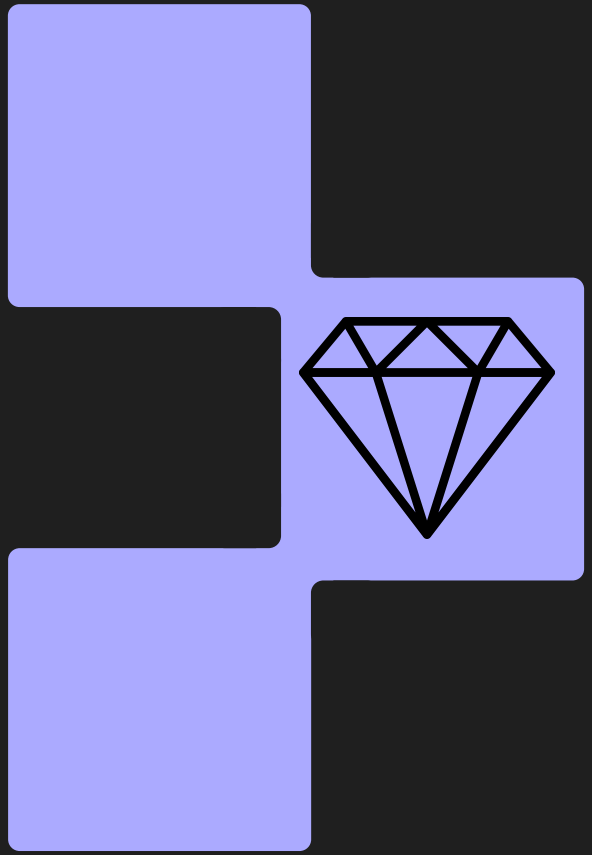
Opportunities for AI in Biometrics

- Areas of opportunity:
  - Documentation & TMF
  - Validation & UAT
  - CRF and edit check configuration
  - Data analysis & visualization
  - New vendor capabilities
  - Statistical programming
  - and so many more...



What AI means for the Biometrics teams of the future

- All roles supported by AI to become more efficient
- Opportunities to streamline and automate processes to become more efficient
- Ability to triage issues immediately to significantly improve data quality
- Proactively informing study operations and success
- Continuous learning



Questions?

Thank You!

Feel free to contact us:

Laura Fazio (lfazio@formation.bio)

Melanie Hullings (melanie@formation.bio)