

## Introduction



### Data silos

The most complete set of data is siloed and uses different storage and access technologies



### Data types

Data can be structured or unstructured. Unstructured data forces users to utilize inconsistent manual processes to draw conclusions



### User perspectives

Curated data typically has multiple disciplines of users with disparate needs

*Having current and related information stored in a variety of sources creates a challenge for CTP personnel responsible for the Agency's tobacco activities*

## Materials & Methods

Data was **harvested** from :

- CTP funded Research Tracking System (RTS) and project documentation
- CTP's internally developed Tobacco Constituents Knowledgebase (TCKB)
- Truth Tobacco Industry Documents from University of California San Francisco (UCSF)
- Industry Analysis Document Tool (iDAT) records

Custom APIs were developed to perform business rules for data **ingestion** and **harmonization** of each data source

Scaled **Agile** methodology employed for implementation

Modern technologies (NoSQL, Angular, REST, etc) leveraged to build an integrated, search and retrieval application that offers advanced features yet is easy to use

# C I R D S

## CTP Integrated Research Data system

CIRDs is a proof of concept that allows users to search, explore, discover and disseminate information from nearly 15 million documents, across multiple data sources.

The platform harvests, ingests, enriches and harmonizes tobacco data to present a state-of-the-art integrated search and retrieval application.



Scan the QR code to access additional information online

## Results

Proof of Concept released in April 2018 with half a million records

Version 2 released in May 2019 with 15 million records

CIRDs consolidates and centralizes tobacco data helping accelerate tobacco research

CIRDs caters to all disciplines within DRSI – staff, scientists, toxicologists, leadership, researchers

Secure, segmented user access and roles for enhanced data security and integrity

## Discussion/Conclusions

CIRDs is a feature rich platform catering to multiple disciplines within CTP. It started off as a simple vision of modernizing tobacco research and making it efficient. Providing a fully functional and fast yet accurate tool, CIRDs has surpassed the original vision.

CIRDs has allowed CTP to consolidate multiple applications through rationalization and reducing duplication and continues to evolve as one of most scalable and extensive platforms.

### Acknowledgments

CTP Office of Informatics, Toxicology, CFSAN SDC, SAIC

### References

MITRE – [www.mitre.org](http://www.mitre.org) / MarkLogic – [www.marklogic.com](http://www.marklogic.com) / CIRDs – [cirfds.fda.gov](http://cirfds.fda.gov)