

HOW THE DEPARTMENT OF DEFENSE CAN LEVERAGE SNOWFLAKE DOCUMENT AI TO EASE RECORDS MANAGEMENT

snowflake®

WHITE PAPER

TABLE OF CONTENTS

3 Background

3 Minimizing the burdens of document management

5 Automated categorization, classification and metadata extraction

5 Discovery and decision making

6 Data protection and governance

6 Summary

Background

Today's connected world generates enormous amounts of information. The Department of Defense (DOD) is no exception. This problem is exacerbated by the inclusion of vast stores of traditionally managed records, due to their volume, complexity and unstructured nature. Effective management and inclusion of these records is crucial for decision making, information sharing and critical mission support.

According to the Deputy Secretary of Defense, key elements in the DOD Records Strategy involve leveraging artificial intelligence (AI) and machine learning (ML) to ease records management burden through the use of automation. Of the key elements within this strategy, metadata extraction, document classification and discovery and decision making were listed as targets in which automation would be of benefit.

This paper discusses the capabilities of Snowflake Document AI[1] and explores the Snowflake approach to meeting the AI/ML and LLM requirements as outlined in the DOPSR 23-S-2095 Memorandum for Senior Pentagon Leadership, Commanders of the Combatant Commands and Defense Agency and DOD Field Activity Directors (10 May 2023). This document will address the following key document automation challenges as listed above.

MINIMIZING THE BURDENS OF DOCUMENT MANAGEMENT

Today's advancements in AI and ML are impressive. With over 50,000 AI companies in existence today, the promise of leveraging AI to carry out repetitive and time-consuming manual tasks of records management are finally within grasp. One of the many goals of Snowflake is to responsibly support new technologies, and AI is no different. Snowflake understands the complexities and demands of document processing and the systems that support it. We also understand that easing the burden for some can — and often does — cause added complexity for others. So what makes us special? How does it ease the burden of records management and how can we help with records management?

Lowering the technological barrier for entry

The first step in reducing the records management burden is to simply make things easier. Snowflake's Document AI feature supports a wide variety of document types and offers a single interface that simplifies data extraction from a wide variety of document types. The intuitive interface allows the user to easily train a custom model for their use, on their own documents. Training is as simple as asking a question in natural language, such as, "What is the unit part number?" or "What is the primary topic?" and assigning a label — this includes handwritten documents. Boolean values such as checkboxes are also supported. Why is this important? Records management personnel are the subject matter experts (SMEs). The capability of bringing AI to the document experts allows for better model training, increased accuracy and, more importantly, does not shift this burden to the IT, data scientists or data engineering teams, allowing them to focus on larger data challenges such as automation workflows or pipelines.

Power and efficiency

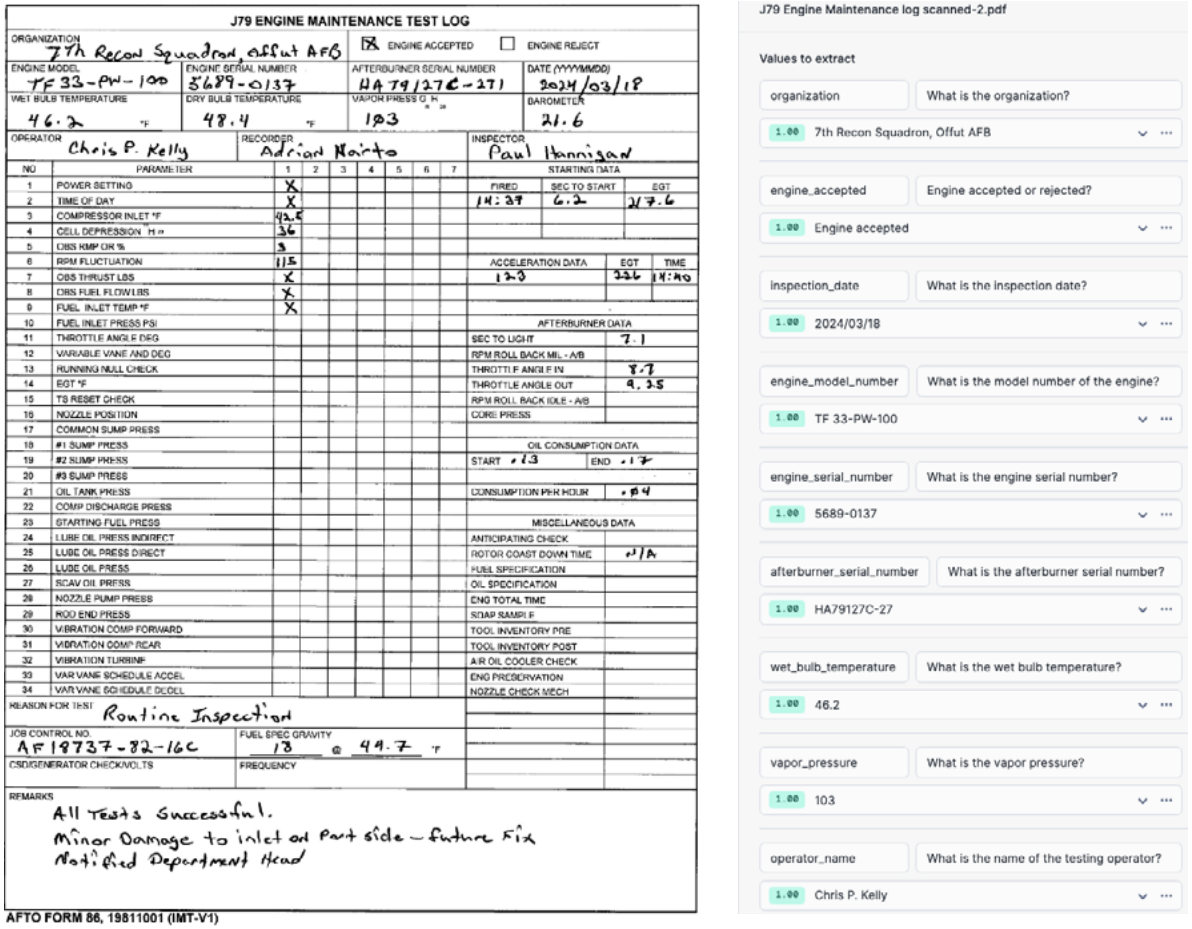
One of the biggest challenges to any AI system is the computing resources necessary to make the technology accessible to those who need it, when they need it. Not accounting for system limitations, costs and complexity can impair the value of an effort such as this. Snowflake deeply understands this, and it is one of the reasons Document AI was created. Leveraging the speed, power and efficiency

[1] Contact your dedicated Snowflake account representative to confirm feature availability in government regions.

of a powerful purpose-built model means more can be done, quicker, with less. At the heart of Document AI is Snowflake's Arctic-TILT (Text Image Layout Transformer) model. This is a Snowflake grown LLM that leverages a proprietary and unique transformer architecture, tailored to understand and extract data from documents. By combining multiple data modalities, Arctic-TILT offers unparalleled versatility and performance in document-understanding tasks. Coupled with the virtually limitless elastic processing and storage capabilities of Snowflake, no task is too large or complex.

KEY FEATURES AND CAPABILITIES

- **Multimodal understanding:** Arctic-TILT can understand, analyze and extract information from text, images and spatial layouts simultaneously, providing a holistic understanding of content and its context.
- **State-of-the-art performance:** On benchmarks such as DocVQA, Arctic-TILT demonstrates Visual Question Answering capabilities on par with if not better than models like GPT-4 with orders of magnitude more parameters.
- **Extended context window:** Arctic-TILT features an exceptionally large context window of 375,000 tokens. This capability is crucial for grasping the full context of multimodal content.
- **Efficient inference:** Arctic-TILT is designed to handle both small- and enterprise-scale document volumes, while maintaining performance and, more importantly, accuracy — both of which are critical when it comes to business document processing.
- **Adaptability:** Designed for a wide range of applications and industries, Arctic-TILT requires no previous knowledge of a given document or format and is easily fine-tuned if needed.



EXAMPLE: VALUE EXTRACTION UTILIZING SNOWFLAKE DOCUMENT AI AGAINST A COMPLEX LEGACY MAINTENANCE FORM. NOTE THAT THIS FORM IS AVAILABLE PUBLICLY AND THE EXAMPLE USES SYNTHETIC, HANDWRITTEN DATA FOR THE PURPOSES OF ILLUSTRATION.

AUTOMATED CATEGORIZATION, CLASSIFICATION AND METADATA EXTRACTION

An additional burden to records management is the categorization, classification and metadata extraction necessary to group and understand what each document represents and has to offer. This is a very labor-intensive process which can introduce risk due to human factors and decrease the efficiencies that could be gained through repeatable, automated processes. Leveraging AI within an automated and finely tuned pipeline can relieve many of these issues. With its multimodal understanding and efficient inference, Snowflake's Document AI, using Arctic-TILT, provides the ability to quickly understand, analyze and extract information from text, images and spatial layouts simultaneously. This, coupled with a refined processing pipeline, can decrease risk and increase the speed and efficiency of document processing.

DISCOVERY AND DECISION MAKING

One of the key promises of AI technology is the ability to quickly analyze vast amounts of information to gain insights previously unknown. Snowflake is quite adept at utilizing the best models available today, including Snowflake Arctic. To achieve the greatest efficiency, accuracy and repeatability typically means the development of processes and models that best suit the needs of the operation. With the capability to easily adapt and scale to any workload without contention, Snowflake is well-suited to processing the hand-written forms typical of legacy systems — records that pose challenges for the DOD and public sector records management in general. When coupled with an environment such as Snowflake Snowpark, data engineers can easily and securely run the most popular data science languages such as Python, Java and Scala directly within Snowflake, allowing powerful AI solutions to be developed directly within the platform, offering significant and unique advantages in the development of your AI pipeline needs, such as:

- **Process data where it resides**
  Traditionally, data is moved out of a system to allow processing by other tools. Snowpark eliminates this need, allowing developers to write code directly on data within Snowflake, increasing speed and efficiency.

- **Use of familiar languages and libraries**
  Snowpark lets you use your preferred programming languages (Java, Scala, Python) and libraries for AI tasks, allowing developers and data scientists to leverage existing skills rather than learning a new system.

- **Custom logic and UDFs**
  Snowpark allows you to create powerful custom user defined functions (UDFs) and logic for all your needs which is helpful for tasks such as data transformation, feature engineering, or deploying machine models as UDFs within Snowflake.

- **Maintaining security and governance**
  With Snowflake Snowpark, data never has to leave your security and governance boundary. Data remains under your control and can adhere to your existing governance policies, eliminating concerns of data movement outside the platform.

With its ease of use and powerful development environment, we can see that Snowflake Document AI helps eliminate the burdens associated with traditional records management. The power, speed and efficiency of Snowflake's AI Data Cloud platform provide the resources necessary to drive even the most complex automated document processing workloads, while its robust analytic capabilities reveal previously unknown trends and patterns, allowing you to quickly make the accurate strategic decisions necessary for operations support.

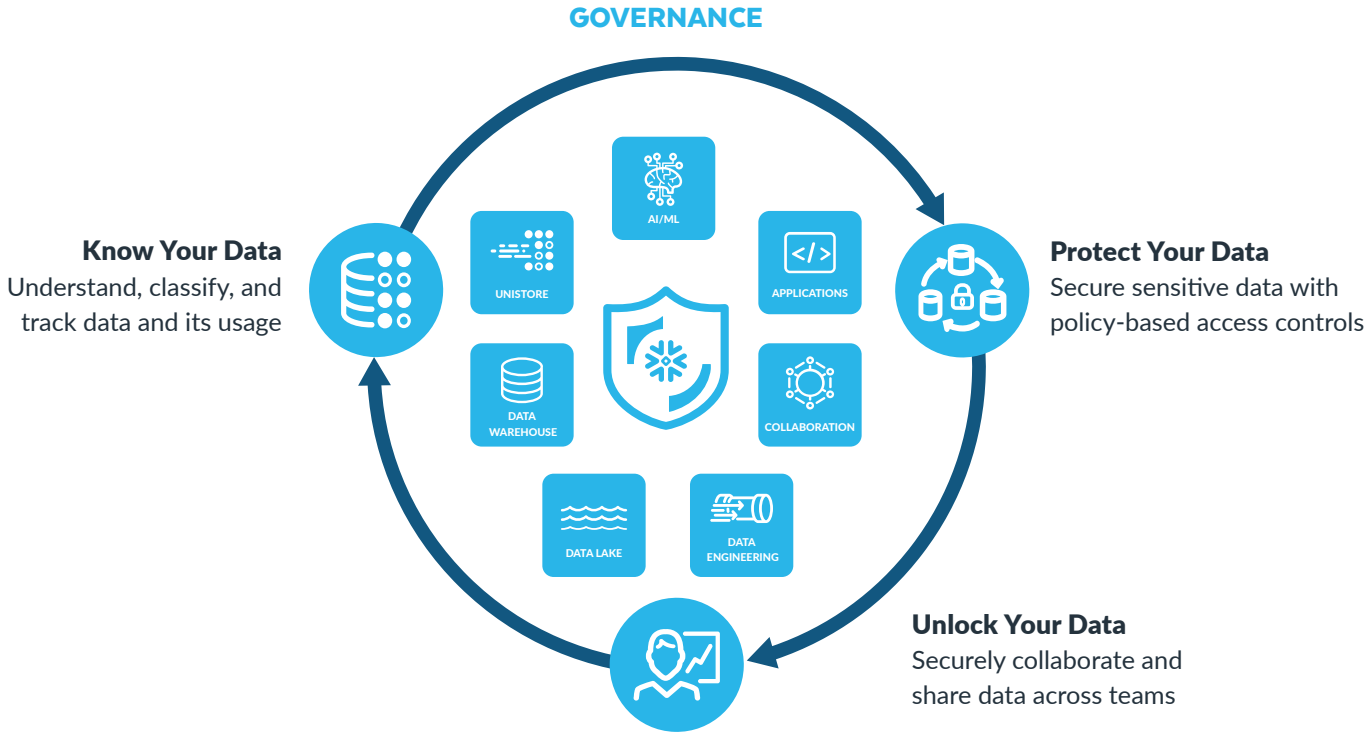
DATA PROTECTION AND GOVERNANCE

Transparent and non-impacting to the user, data in Snowflake is encrypted with at least AES-256 at rest and TLS 1.2 in transit over untrusted networks. Encryption keys are rotated automatically every 30 days by default, minimizing the burden of maintaining keys. If higher security requirements are necessary, Snowflake offers Tri-Secret Secure, allowing customer key management within a cloud provider that, coupled with Snowflake keys, creates a composite key structure, so that if the customer-managed key is revoked, the data cannot be decrypted.

Snowflake's governance capabilities support role-based access controls (RBAC) and provide the ability to implement dynamic permissions down to the column level. Coupled with its dynamic data masking and robust policy creation, Snowflake allows creation of fine-grained controls over your data allowing:

- Column-level security
- Row-level security
- Dynamic data masking
- External tokenization

Additionally, Snowflake offers dynamic data tagging, which allows the assignment of metadata to Snowflake objects to further enhance data governance. Tagging can be centralized or decentralized and allows for the monitoring of sensitive data for compliance, discovery, protection and resource usage.



SUMMARY

Records management and processing is an arduous challenge in and of itself. With its ease of use, power and robust AI development capabilities all within a single governance boundary, Snowflake Document AI and AI Data Cloud platform provide the essential, secure automation capabilities needed to help ease the burden of records processing management today and beyond.



ABOUT SNOWFLAKE

Snowflake makes enterprise AI easy, efficient and trusted. Thousands of companies around the globe, including hundreds of the world's largest, use Snowflake's AI Data Cloud to share data, build applications, and power their business with AI. The era of enterprise AI is here.

Learn more at **snowflake.com** (NYSE: SNOW).









© 2024 Snowflake Inc. All rights reserved. Snowflake, the Snowflake logo, and all other Snowflake product, feature and service names mentioned herein are registered trademarks or trademarks of Snowflake Inc. in the United States and other countries. All other brand names or logos mentioned or used herein are for identification purposes only and may be the trademarks of their respective holder(s). Snowflake may not be associated with, or be sponsored or endorsed by, any such holder(s).