

AI-Enabled Document Analytics

8 Nov 2023



©2023 Tau Express Pte. Ltd. All rights reserved.

ີ່ເ

DISCLAIMER

By acceptance of this Document, the recipient agrees that this document and the information contained therein, and the recipient's involvement in the process described herein will be kept confidential and not disclosed to third parties unless agreed in writing by the author. All information contained in this Document is Confidential Information.

Use of Confidential Information is strictly limited to the preliminary consideration of an investment in the business described in this Document. No portion of this Document may be reproduced, duplicated, or revealed in any manner without prior written consent of the author. Without limiting the generality of the foregoing, the recipient also agrees that photocopying or other duplication of the Document and the data provided without prior written consent is strictly prohibited.

No information contained in this Document or any other written or oral communication transmitted or made available to a recipient of this Document is, or shall be relied upon, as a promise or representation, whether as to the past or future, and no liability will attach except as may be provided in a definitive agreement when, as if it is executed and subject to such limitations as may be provided in such agreement. In addition, any projections and estimates or other forward-looking statements contained in this Document involve numerous and significant subjective assumptions and are subject to risks, contingencies, and uncertainties (many, if not most of which, are outside the control of the author and the Company being represented) which could and likely cause actual results to differ materially. Such assumptions, and risks, contingencies, and uncertainties relate to, among other matters, general business, economic and market conditions.

OUR COMPANY

TAU Express Pte Ltd is an AI Analytics company founded in 2018 as a spin-off from Nanyang Technological University. With deep expertise in machine learning and natural language processing, we built "**TITAN**", an Intelligence Analytics Platform capable of parsing, extracting and categorizing unstructured text data for intelligent search and analytics.

TITAN sets the foundation for our company's suite of Solutions and has been implemented in projects led by both private and public sectors.



Industries TAU has experience in

- Banking and Financial Institutions
- Aviation
- Legal
- Healthcare
- Logistics
- Government
- Cybersecurity



Initial Project Lead



Prof LAM Kwok Yan

Decades of research excellence, technical expertise and industry experience

Vice President (Strategy & Associate Partnerships) and Professor in School of Computer Science and Engineering at Nanyang Technological University (NTU), Singapore. He is concurrently serving as the Executive Director of the Strategic Centre for Research in Privacy-Preserving Technologies and Systems (SCRIPTS), and Director of NTU's SPIRIT Smart Nation Research Centre.

INTELLIGENT CASE RETRIEVAL SYSTEM FOR DECISION SUPPORT

64] THAT COURTS, SINGAPOR



The Supreme Court and State Courts, together with NTU SPIRIT Centre, have embarked on the development of an Intelligent Case Retrieval System. The aim is to have a more effective retrieval of key legal documents in decided cases for road traffic accidents and personal injury matters. Besides making it easier and guicker for legal research to be performed, it will add to the quality of decision making and judgment writing for judges. The technology developed through this platform may even perform other legal functions, such as Case Summarisation and Logic Reasoning. This will lead the way to a more sophisticated legal research, advisory and decision support system in the future.

FOREWORD BY THE HONOURABLE THE CHIEF JUSTICE

SUPREMECOURI SINGAPORE



and impartiality

less acrimoniously

The judiciary will have to respond to an

increasingly challenging global and economic

climate. We must enhance the quality of our work

even as we look to reduce the cost of litigating by

we have introduced significant changes to our civil

and criminal litigation processes. At the Family

based on actuarial data is also being devised to

so that such disputes can be settled quickly and

guide the resolution of cases involving children

Justice Courts, a child maintenance matrix,

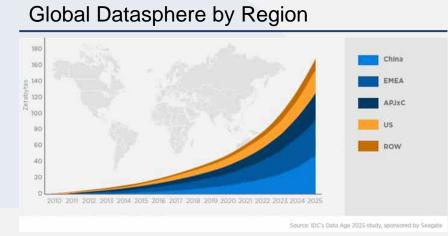


On the international front, we are widenin eives me great pleasure to present for the very first time, a One judiciary annual report boxed set, and strengthening our judicial networks. Ai which showcases the work of the Supreme Court. International Advisory Council which I chair and the State Courts and the Family Justice Courts comprising family judges from various jurisdictions sharing a common mission to superintend the provides the Family Justice Courts with many useful administration of justice and make justice and diverse perspectives that will help us address accessible to all. Themed "Upholding Our Values" present and future challenges. the reports detail the work of the three Courts and their initiatives that are underpinned by the values I am also delighted by the success of the first of fairness, accessibility, independence, integrity

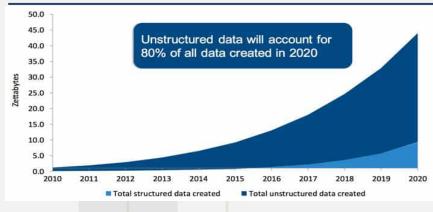
judicial insolvency network confetence involving alose from 10 invitalictions which we bested in October 2016. The conference resulted in the promulgation of guidelines for court-to-court ommunication and cooperation pretocols Following the meeting, Singapore and the United States Bankruptcy Court for the District simplifying processes where possible. On this front, of Delaware formally implemented the guidelines with other jurisdictions expected to similarly adop the guidelines.

> Technological advancements will require the logal sector to rethink entire areas of practice. While this will likely prove to be disruptive, we should embrace and welcome the change as technology can potentially improve processes, resulting in greater effectiveness at a lower cost for society. The Courts of the Future Taskforce will make recommendations to get our Courts "future-ready while the Singapore Academy of Law has launched the Legal Technology Vision to help the legal sector adopt and implement technological changes.

The annual growth of unstructured data has increased by 55-65%



Capacity Growth by Data Type



WHY IS THIS A PROBLEM?



Huge Amount of Unstructured Data - Many key operations depend on personnel handling large amounts of unstructured data in their day-to-day activities and require decisions to be made based on such unstructured data.



Manual Human-dependent Analysis - It is relatively easy to build systems around structured data, but humans are typically required to decipher and make sense of unstructured information.



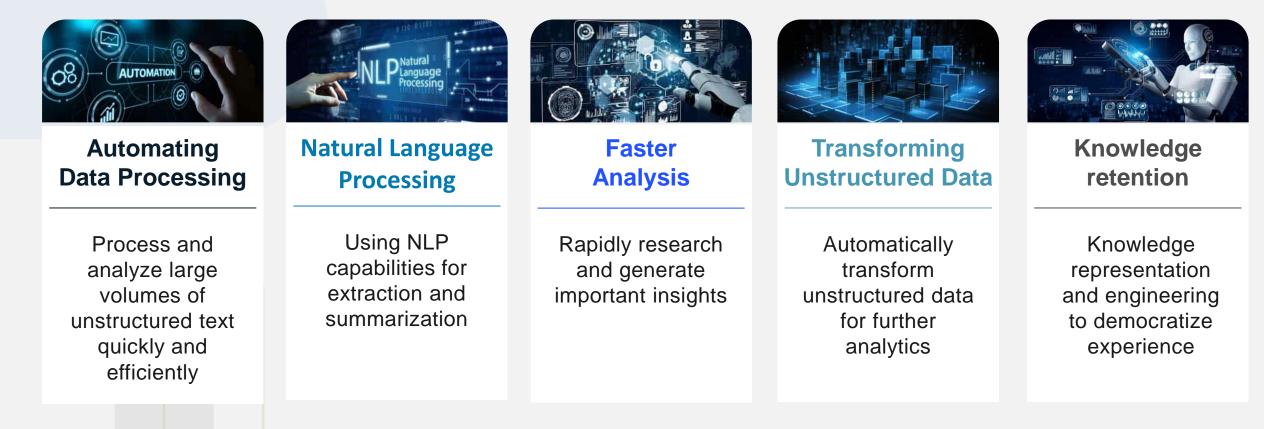
Opportunity Cost from Unused Data - A lot of unstructured data is stored and rests dormant due to the difficulty of deciphering, understanding, sorting, and searching through the data, which presents a huge opportunity for value extraction.



Inefficient and Prone to Human-error - This results in inefficiencies, lack of consistency in task execution across departments, and human-error.



AI-POWERED TOOLS TO STREAMLINE TASKS, ENHANCE EFFICIENCY AND DRIVE PRODUCTIVITY GAINS



"Data

44

in the 21st Century is like oil in the 18th Century: an immense, untapped valuable asset... those who learn to extract and use it will see huge rewards."

- "Data Is the New Oil of the Digital Economy", 2020

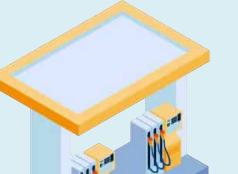
2023 © TAU Express Pte Ltd – All Rights Reserved

TITAN is a refinery that

turns raw unstructured data into fuel

for driving business decisions.





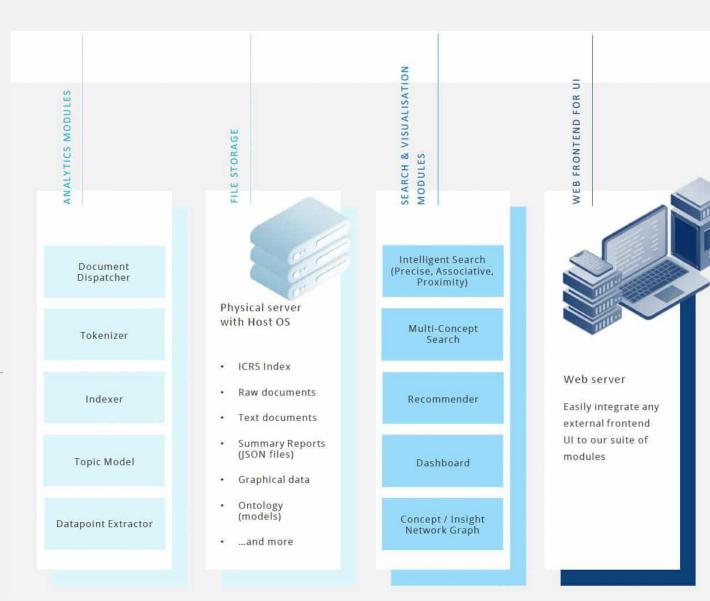
7



OUR SOLUTION TITAN

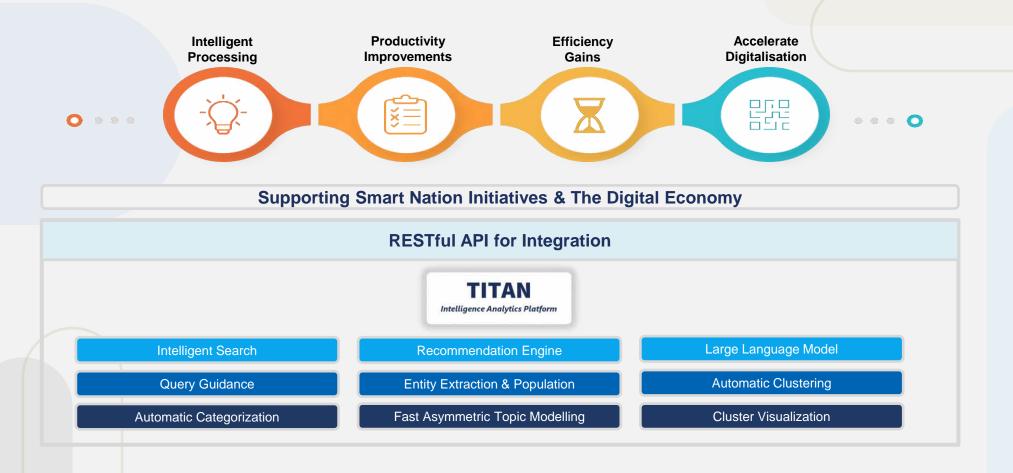
An AI-enabled Platform **proven to achieve** >90% accuracy in retrieving relevant results. An intelligent retrieval system that can parse, extract, and categorize unstructured data to enable AI-driven data analytics & insights

- Backed by strong R&D from world class research center
- Industry agnostic
- Modular and scalable architecture
- Designed to support combination of engines/modules to meet evolving requirements



TITAN – Modular & Scalable System Architecture

TECHNOLOGY PLATFORM & PRODUCT MODULES



CASE STUDY 1 CYBER THREAT INTELLIGENCE

One of the Top Cyber Security Agency in ASEAN

Business Challenges

- **Growing volume & disparate sources** of unstructured CTI data from different vendors
- Inconsistent terminologies across different reports
- **Tedious & time-consuming** to manually contextualize cyberattacks
- Manual evaluation of CTI data leading to inconsistent and less reliable threat assessments.

Value Creation

- **Increase efficiency** by easily retrieving relevant reports using intelligent search
- Automated correlation analysis to help visualizing links and relationships between entities
- Automated, structured summary to present users with key indicators/entities at once glance

Intelligent Search

	AN Q 20000				
•	(particular)		0 9 1 4 1 1 m	-+ menerices + 0 D (
	Teen hucker agreet to 3 years in provint for 1	witter bitum scarts 🛛 🔒	C Control of the state of the s	the test suspects charged for roles in Twitter back, Bitenin sc	
	4. Alamba regelator has principal galling in A paid chan bank of right people Transmissioners in your a regel posterior mapping XT/2000 sectors of manageme.			First supervised starting to the test again and the sound formation of a starting starting to the sound of th	- C - C - C
	REAL PROPERTY				100
	Tetl acted information				inthen.
	IP ABDYSS	~		 Secondary Protect (1) A province on the same of second size (1) and the same start (1) have related to the second size of the same start (1) and (1) and (1) and (1) and	-
				The site of a second advantation of Property Theory Report I Travers Readinger Ray, 1981 Rayma, Doors, A sub-REC from West, Arma, all a research and they work a mendal to be been ad-	100
	Hanhes	~			-
	UMLA	¥ .		 Support on Paylout Angel 4 attended again, excluded in a familier 4 to Paylout Callers, account and again 	
	Emails			¹ The second statement is a second seco	
	Countries			I have an only for the design hand the local fact has been oblighten in account of a many links contrast, or a state of the second state of the se	-
	Contraction of the second seco	100 B			
	Vulleer and men.	~		(b) the end of the	
	Threat Actors	~		The set is the local momentum of the same transmission of the set of the bit spin with the set of t	
				Incoding a finite gang into the 100 hits injugate an ensure the same of a super-initial agent on a second s	1010
	Campaign Names	~		To be all out that I have a closer for he cannot any hearing out a particular sector of the sector o	144

Concept Network Graph



Statistics Dashboard



2023 © TAU Express Pte Ltd – All Rights Reserved

CASE STUDY 2LEGALTECH & DOCUMENT INSIGHTS

One of the Top Public Prosecutor Office in ASEAN

Business Challenges

- Growing number of legal documents to be reviewed / processed
- Laborious process to look through scanned documents to search for required text
- Single document broken up into multiple files
- Documents in the form of scanned PDFs are **not searchable or editable**

Value Creation

- Automatic conversion of scanned PDFs into textual files to enable intelligent search and document analysis
- Facilitate efficient drafting of legal documents
- **Recommend similar clauses** from documents in the repository
- Reduce time and cost in looking for relevant documents

Intelligent Search

1.2		C. most favoured to	ation	00
•	Query Expansion			
5	Original Words	Associated Words		
e i	most ferrored ratios	most favoured nation status	, most favoured nation, trade egreement	
	Search Results			
ŝ				
2	Highlight legend	tavoores ration matur insuit Tamar	tain agreement	
2	Rigillight legens	faround ration status		Rankpergage 28 - Header1-2014105 - 1-1
5	Rigillight legens	Document Trille		Romi per page - 20 Romin 1 - 20 al 210 - 1 - 1 Date el Repontere
2	aligilight legend most favysynd nation	Dataset Tile		
à	Highlight legend Incl. Ground Autor, man File Name EFTR - Hang Koog,	Datament Title APPENDIX 2 TO ANNEX X LIST OF RESERVATIONS	Srepat	Date of Eignature

Document/Clause Comparison



Clause Recommender

I	and a second				
	ause Recon				
• •••	erch the repositor	y for clausse used in any	lar context documents		
			a and restrations, which in mutual bads provides for the nonuse of antidupping, comp		on the beats of pu
	uniform policy in ta	will and noritaril regulation	uniform rules of competition, application of subsidies and other forms of state support		
					-
				Casa	Q. Record
				finner per pages (20) - Result	N-2-50 H-10
				from per page (10) - theory	MAR
		GREEMENT TITLE	RECOMMENDED CLAUSE	Nove per page - 30 Result MARK RELEVANCE	MAR AS MOS HELI
	9	CORSEMENT TITLE	RECOMMENDED CLAUSE		MAR
			· establishing a free trade area without exclusions and restrictions.		MAR AS MOS HELI
		Somman Economic	 establishing a free trade area without exclusions and real/closs, which is mutual tode provides for the nonuce of antidumping, compensativity and search provider executes on the basis of provide the search of the		MAR AS MOS HELI
		Symmon Economic	 establishing a free trade area without exclusions and real-totons, which is mutual trade providee for the noruse of astribumping. 	MARN RELEVANCE	MAR AS MOS HELI
		Symmon Economic	 estigitioning a the trade area without exclusions and realifctions, which in mutual tode provides to the nonuce of estidations; a unternewing values if which examines on the basis of parsing a unterney paties; in tode and unternet regulation, underne roles of immediation, setticing and and and and and and a differentiated start support; executing a set of measures to compare to the diredepresent of a free- 	MARN RELEVANCE	MAR AS MOS HELI
	0	Symmon Economic	 estigationing a fine train encode acculations and tradictions, which is mutual tasks provides to the consist of estimationspin, compresenting and specific protective measures on the basis of provide a universe particle, in ordinary and encode acculation under the compactions, application of autoaties and after formal of state support; 	MARN RELEVANCE	MAR AS MOS

2023 © TAU Express Pte Ltd – All Rights Reserved

Retrieval Augmented Generation

Open-Source CTI Reports

what does the PDPA govern and what is the penalty for breach of user privacy

Submit

Output

The Personal Data Protection Act (PDPA) governs the collection, use, disclosure, and protection of personal data in Singapore. The Act applies to all organizations that collect, use, or disclose personal data, including government agencies and private companies.

Under the PDPA, there are specific provisions for the handling of sensitive personal data, such as information about an individual's race, religion, health, or sexual orientation. Organizations must obtain explicit consent from individuals before collecting, using, or disclosing their sensitive personal data, and they must also ensure that such data is protected with appropriate security measures. The penalty for breach of user privacy under the PDPA can be severe. If an organization fails to comply with the Act's provisions and causes harm to individuals as a result, it may be liable for damages and face fines of up to \$1 million. In addition, the Commissioner of Personal Data Protection has the power to impose penalties of up to \$50,000 on individuals who breach the Act's provisions. Repeat offenders may also be subject to criminal prosecution and imprisonment.

References

*** source: raw_pdfs\Personal Data Protection Act 2012.pdf*** page: 89
 *** source: raw_pdfs\Personal Data Protection Act 2012.pdf*** page: 32
 *** source: raw_pdfs\Personal Data Protection Act 2012.pdf*** page: 58
 *** source: raw_pdfs\Personal Data Protection Act 2012.pdf*** page: 111

 Selected documents stored in repository for access and retrieval of relevant knowledge during model's inferencing phase.

- 2 User-entered prompt. A more detailed response will be generated if "Detailed Reply" is checked.
- 3 Model learns to extract relevant information from the documents and uses it to generate answers.
- 4 Knowledge base in the repository can be updated regularly to ensure that the language model remains up-to-date with the latest information.

Dataset:

Input Query

under the act?

Detailed Reply

IMDA Regulations & Guidelines

The context retrieval ensures that the model generates responses that are grounded in information contained within the knowledge base.

TAU EXPRESS	
Dataset:	Output
IMDA Regulations & Guidelines Open-Source CTI Reports Statutes • Pharmaceutical Drugs	References
Input Query	
Detailed Reply	
Low Similarity Threshold	
Submit	



THANK YOU

We look forward to working with you

 \bigcirc 77 Science Park Drive #04-14 Singapore, 138632

info@tauexpress.com







©2022 Tau Express Pte. Ltd. All rights reserved.