TRACE: A Framework to Assist Auditors in Evaluating Regulatory Compliance

Vidyaraman Sankaranarayanan

Divya Venkatraman

DeepDive Labs, Singapore

PROTOTYPE: https://deepdive-labs-trace.streamlit.app/

CODE: https://github.com/svidyaraman/TRACE-Framework/

DeepDive Labs

DEEPDIVE LABS

Singapore startup offering technology consulting, solutioning and training services in Data and AI

SERVICE MODEL

- Train, then consult
- Consult, then train

BESPOKE TRAINING

- Data Analytics
- Data Science for UX designers
- GenAI for educators
- Data Engineering for AI etc.

REGXPERIENCE

This paper is a subset of the work that forms part of our upcoming Micro-SaaS product RegXperience

OUR MOTIVATION

- Vidyaram during is time as Risk Architect at Microsoft, helped their enterprise clients in their audit process.
- Hours of manual effort was spent to link regulations to evidence across different excel sheets!

- Divya worked on creating knowledge graphs to get answers from text verdicts (on femicides) from Hong Kong Legal Information Institute.
- Presented in conference NODES 2023 by neo4j, a graph database company

Agenda

- Regulatory Compliance Process & Challenges
- TRACE framework
- Implementing TRACE
  1. Single-step prompt
  2. Workflow prompting
- Demo & key findings
- Q&A

REGULATORY COMPLIANCE PROCESS

1	**Regulators craft guidelines**	Example: MAS TRM guidelines, DPTM certification by IMDA
2	**Organizations create internal policies and processes to adhere to these regulations**	Example: policies around data protection
3	**Proof of adherence is provided by means of evidence documents**	Example: Software Configurations, interview transcripts, dashboards, etc.
4	**Auditors assess policies and evidence documents to check if they adhere to the regulatory requirements**	Validate letter and spirit of regulations is met.

User Quotes & Insights

Jane
Lead Auditor
I have to review everything manually and read through everything myself.

Jane
Lead Auditor
How am I gonna make sure that all this information that the clients send me is not co-mingling with each other?

JT
Financial Crime Compliance
Dealing with fragmented and scattered compliance documents makes it hard to get a complete picture.

Jeff
Tech Compliance
Everything is quite manual, and we cannot just dump a policy into AI and expect the results we want.

DATA MANAGEMENT & INTEGRATION

- Challenges in managing and ensuring accuracy of compliance documentation
- Desire for AI solutions to streamline workflows

ACCESSING RELEVANCY

- Challenges in assessing control effectiveness & relevance to business process
- Manual review process is time-consuming & burdensome

UNORGANIZED DOCUMENTATION & MANUAL REVIEW

- Struggle with fragmented and scattered compliance documents
- Difficulty in obtaining a complete overview

User Quotes & Insights

Jane
Lead Auditor
I have to review everything manually and read through everything myself.

Jane
Lead Auditor
How am I gonna make sure that all this information that the clients send me is not co-mingling with each other?

JT
Financial Crime Compliance
Dealing with fragmented and scattered compliance documents makes it hard to get a complete picture.

Jeff
Tech Compliance
Everything is quite manual, and we cannot just dump a policy into AI and expect the results we want.

DATA MANAGEMENT & INTEGRATION

- Challenges in managing and ensuring accuracy of compliance documentation
- Desire for AI solutions to streamline workflows

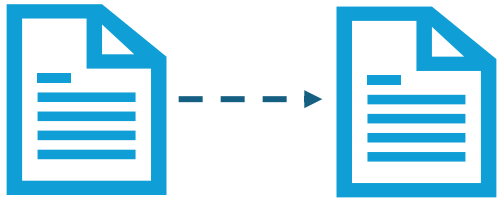
ACCESSING RELEVANCY

- Challenges in assessing control effectiveness & relevance to business process
- Manual review process is time-consuming & burdensome

UNORGANIZED DOCUMENTATION & MANUAL REVIEW

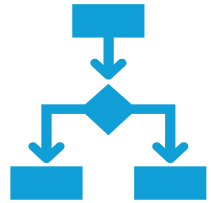
- Struggle with fragmented and scattered compliance documents
- Difficulty in obtaining a complete overview

Key Challenges



Lack of Structure of the different documents to enable mapping

1. **Varied Formats:** Regulatory documents may not follow a standard structure and can vary widely in format.
2. **Diverse Topics:** Include varied subjects like data resilience, encryption (MAS TRM, RMiT), and rights of subjects (GDPR, PDPA).
3. **Multiple Stakeholders:** A single requirement might impact different stakeholders like employees, customers, and partners, complicating compliance.



Diverse Types of Required Documents

1. **Multiple Documents for One Regulation:** Often, fulfilling a regulation needs several policy and evidence documents.
2. **Combined Documents:** Policy and evidence may reside within the same document, complicating extraction and review.
3. **Varied Evidence Forms:** Evidence can include interview transcripts, web updates, or data from analytical dashboards.

FRAMING THE PROBLEM FOR LLMS

- GENERIC: Given a set of policy and evidence documents, determine if they satisfy a regulatory guideline.
- PRESCRIPTIVE: Given a regulatory guideline document* **R**, a set of policy documents* **P**, and a set of associated evidence documents* **E**, determine if all the requirements in the regulations are satisfied through a combination of the policy documents and evidence documents.

*In any format

The TRACE Framework

Technical **R**egulatory **A**ssessment and **C**ompliance **E**valuation

a framework designed to aid auditors in assessing compliance with regulations using LLMs.

STEP #1: Extraction of Requirements

- Extract Requirements: Extract a detailed list of compliance requirements through CoT prompts

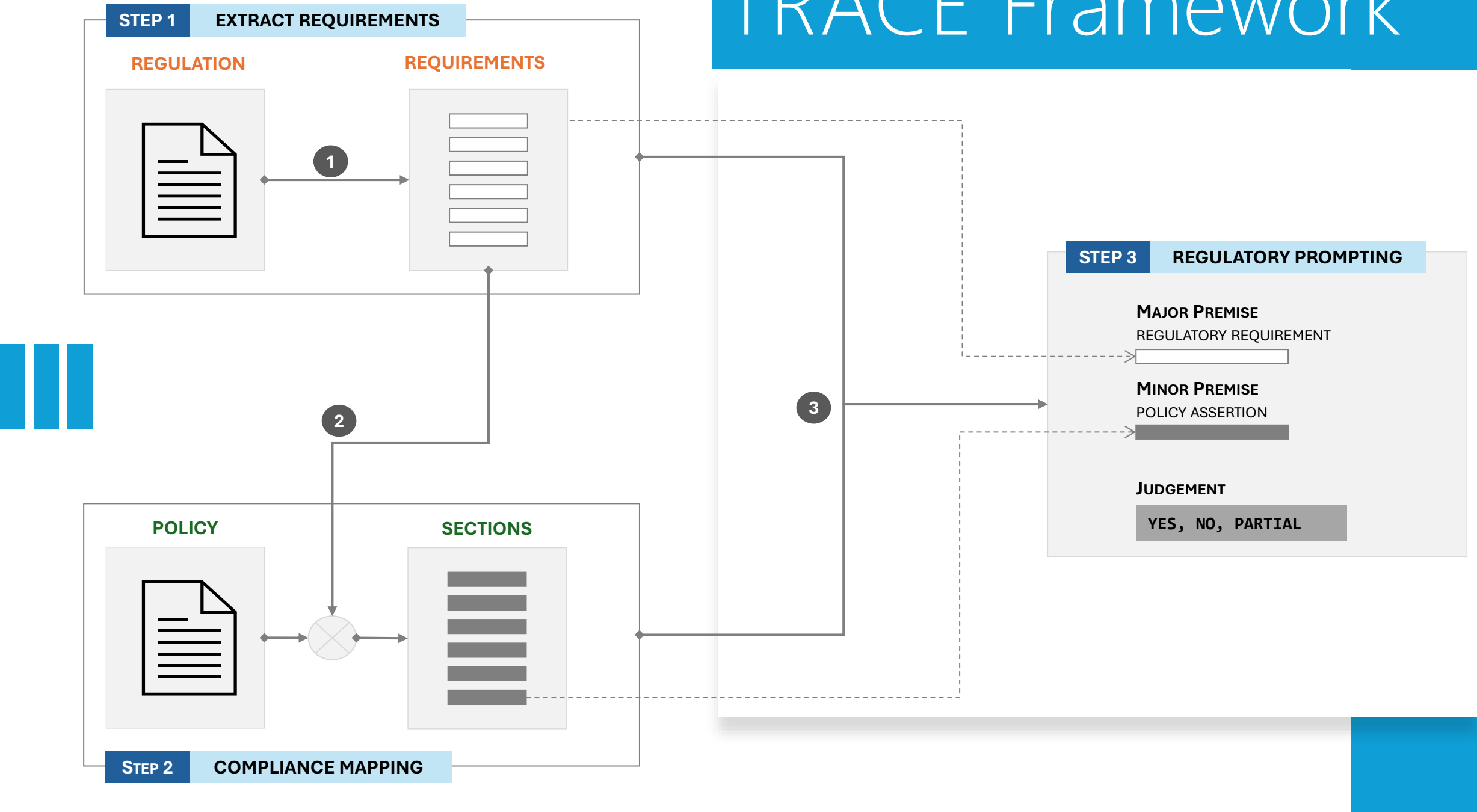
STEP#2: Compliance Mapping

- Regulation-to-Policy Alignment: Match regulatory requirements with relevant sections in policy and evidence documents.

STEP #3: Regulatory Prompting

- “major premise” : the regulatory requirement
- “minor premise”: the policy assertion
- “judgement”: evaluate if the policy assertion satisfies the regulatory requirement.

TRACE Framework



Implementing the TRACE Framework

Testing Specifications

REGULATORY DOCUMENT

- **DPTM certification requirements** published by IMDA, Singapore.
- Voluntary certification demonstrating accountable data protection practices.

POLICY DOCUMENTS

- Policy document: privacy policy from **DBS** and **Grab** website.
- Why? These companies were on IMDA's list of DPTM certified organizations

EVIDENCE DOCUMENTS

- Included **privacy policies specific to job applicants** from career pages of **DBS** and **Grab.**

PROCESS

- **Two-Stage Testing:**
- Stage 1: All regulation vs Policy Mapping followed by Judgement
- Stage 2: Subset of regulations vs Evidence Mapping followed by Judgement

  to evaluate TRACE's effectiveness

- **ChatGPT-4-Turbo Model** from OPENAI was used throughout

DPTM Certification Checklist

This checklist provides a broad outline based on abridged DPTM certification requirements to help organisations gauge their readiness before applying for the DPTM certification. To access the full DPTM certification requirements, organisations would need to apply for the DPTM certification at www.imda.gov.sg/dptm.

Organisations should review their data protection regime using the checklist and having a "yes" answer to all the questions is an indication that the organisation is ready to apply for DPTM.

However, kindly note that answering "yes" to all questions on this checklist **may not necessarily equate to meeting all the DPTM requirements**.

The DPTM assessment will also require the organisation to demonstrate and provide evidence for the following:

- Documented data protection policies and processes; and
- Demonstrate that data protection policies and processes are implemented and practised on the ground.

	Checklist	Yes	PDPC's Reference Advisory Guides/Guides/Templates
Principle 1: Governance and Transparency			
A: Establish data protection policies and practices			
1	Organisation shall have data protection policies and practices approved by management, setting out the organisation's approach to managing personal data (include management of special categories of personal data such as personal data of a sensitive nature) for various stakeholders such as:		• Advisory Guidelines on Key Concepts in the Personal Data Protection Act • Guide to Accountability under the Personal Data Protection Act
	• Employees - Internal data protection policy and notice	☐	• Data Protection Notice Generator (https://apps.pdpc.gov.sg/dp-notice-generator/introduction)
	• Customers, Job applicants, visitors etc - External data protection notices	☐	
	• Third party vendors - Third party agreement for management of the organisation's personal data	☐	• Guide to Managing Data Intermediaries • Guide on Data Protection Clauses for Agreements Relating to the Processing of Personal Data

Understanding DPTM

1. PRINCIPLE 1: GOVERNANCE AND TRANSPARENCY

- Establish data protection policies and practices (10) : *Across different stakeholders*
- Establish queries, complaints and dispute resolution handling processes (2): *Across different stakeholders*
- Establish processes to identify, assess and address data protection (3)
- Establish data breach plan (5): *Different levels of planning*
- Accountability (2)
- Internal Communication & Training (1)

2. PRINCIPLE 2: MANAGEMENT OF PERSONAL DATA

- Appropriate purpose (1)
- Appropriate Consent (2)
- Appropriate Use and Disclosure (4): *Across different process*
- Compliant Overseas Transfer (2)

3. PRINCIPLE 3: CARE OF PERSONAL DATA

- Appropriate protection (4)
- Appropriate Retention and Disposal (5)
- Accurate & complete record (3)

4. PRINCIPLE 4: INDIVIDUAL'S RIGHTS

- Effect withdrawal of consent (2)
- Provide access and correction rights(4)

Single-step prompt

WHAT	ADVANTAGE	DISADVANTAGE
All the three steps in the TRACE framework are provided as a single Chain-of-Thought prompt	1. useful for validation of snippets of regulatory requirements	1. complete compliance mapping is not possible due to limited number of output tokens 2. results are not reproducible over multiple runs

PROMPT

- YOU ARE AN AUDITOR RESPONSIBLE FOR EVALUATING DATA PROTECTION REGULATIONS IN SINGAPORE.
- YOUR TASK IS TO EXTRACT DATA PROTECTION REQUIREMENTS FROM THE DATA PROTECTION TRUSTMARK CHECKLIST AND MATCH THEM TO THE RELEVANT CLAUSES IN THE PROVIDED POLICY DOCUMENT.
- BASED ON THIS MAPPING, YOU NEED TO PROVIDE A COMPLIANCE JUDGEMENT (**MET**, **NOT MET**, **PARTIALLY MET**), AN EXPLANATION FOR EACH JUDGEMENT, AND SUGGEST THE TYPE OF EVIDENCE REQUIRED, ALL FORMATTED IN A SPECIFIC JSON STRUCTURE.

Results of Single-Step Prompt : **Trial 1**

Regulatory_Requirement_Description	Policy_Documentation	Judgement	Judgement_Reasoning	Type_of_Evidence
Organisation shall have data protection policies and practices approved by management, setting out the organisation,Äôs approach to managing personal data (include management of special categories of personal data such as personal data of a sensitive nature) for various stakeholders such as employees, customers, job applicants, visitors, and third-party vendors.	DBS has in place safeguards to protect the personal data stored with us. This policy describes how we may collect, use, disclose, process and manage your personal data.	Partially Met	The policy mentions safeguards and general management of personal data but does not explicitly detail the management of special categories of personal data or the approval by management.	Inquiry
Organisation shall publish and communicate the data protection policies to relevant internal and external stakeholders through appropriate platforms.	This policy applies to any individual,Äôs personal data which is in our possession or under our control.	Partially Met	The policy indicates that it applies to all personal data under the organization's control, suggesting some level of communication. However, there is no specific mention of the platforms used for publishing and communicating	Observation
Organisation shall have documented policies and processes to review, update and monitor compliance of data protection policies and practices.	We may amend this policy from time to time to ensure that this policy is consistent with any developments to the way DBS uses your personal data or any changes to the laws and regulations applicable to DBS.	Completely Met	The policy explicitly states that it is subject to amendments to stay consistent with data usage developments and regulatory changes, indicating a process for review and updates.	Inquiry
The organisation shall have documented policies and processes on how it receives and responds to queries/complaints on the collection, use and disclosure of personal data.	Responding to queries or feedback; addressing or investigating any complaints, claims or disputes.	Completely Met	The policy outlines processes for responding to queries, feedback, and complaints regarding the use of personal data, aligning with the regulatory requirement.	Observation
The organisation shall have documented policies and processes on how it performs risk and impact assessments (e.g., Data Protection Impact Assessment) on its operational functions, business needs and processes which involve personal data.	Conducting credit checks, screenings or due diligence checks as may be required under applicable law, regulation or directive; monitoring products and services provided by or made available through us, including to prevent, detect and investigate crime/fraud.	Partially Met	The policy mentions conducting checks and monitoring for compliance and fraud prevention, which are forms of risk assessment. However, it does not explicitly mention Data Protection Impact Assessments or their documentation.	Inquiry

Results of Single-Step Prompt : **Trial 2**

Regulatory_Requirement_Description	Policy_Documentation	Judgement	Judgement_Reasoning	Type_of_Evidence
practices approved by management, setting out the organisation,Äôs approach to managing personal data (include management of special categories of personal data such as personal data of a sensitive nature) for various stakeholders such as employees, customers, job applicants, visitors, and third-party vendors.	DBS has in place safeguards to protect the personal data stored with us. This policy describes how we may collect, use, disclose, process and manage your personal data.	Partially Met	The policy document mentions that DBS has safeguards to protect personal data and outlines how personal data is managed. However, it does not explicitly state that these policies are approved by management or detail the management of special categories of personal data.	Inquiry
Organisation shall publish and communicate the data protection policies to relevant internal and external stakeholders through appropriate platforms.	This policy applies to any individual,Äôs personal data which is in our possession or under our control.	Partially Met	The policy document states that the privacy policy applies to individuals whose personal data is under DBS's control, indicating some level of communication. However, there is no specific mention of the platforms used to communicate these	Observation
Organisation shall have documented policies and processes to review, update and monitor compliance of data protection policies and practices.	We may amend this policy from time to time to ensure that this policy is consistent with any developments to the way DBS uses your personal data or any changes to the laws and regulations applicable to DBS.	Partially Met	The policy mentions that it is updated periodically to reflect changes in data usage or legal requirements, which shows a process for reviewing and updating the policy. However, there is no explicit mention of monitoring compliance with these policies.	Inquiry
The organisation shall establish a data breach management plan and communicate it to relevant employees and external stakeholders. The data breach management plan should include roles and responsibilities of data breach management team, timeline for reporting data breach incidents, processes for notifying affected individuals/organisations and relevant regulators/enforcement authorities, processes for third parties to notify organisation in the event of a data breach,	In addition, should we be aware that your data has been compromised and is, or likely to result in significant harm to you, we will notify you and advise on the possible steps to protect yourself from further potential harm.	Partially Met	The policy document mentions notifying individuals in the event of data breaches that could result in significant harm, which partially addresses the requirement for a data breach management plan. However, it lacks detailed information on the full scope of the data breach management plan, including roles, responsibilities, and processes for third-party notifications.	Inquiry

CONTEXT WINDOW	MAX OUTPUT TOKENS
128,000 tokens	4,096 tokens

- This approach is limited by the max output token of ChatGPT-4-turbo model is 4096 tokens ~3000 words (also applicable to most LLMs)
- Because of the long regulatory document, every trial picks a different set of regulation from the DPTM checklists.
- However, this was constrained when specific category of requirement were accessed

WORKFLOW PROMPTING

DEMO

WHAT

Each step in the TRACE framework is elaborated as a Chain-of-Thought prompt

ADVANTAGE

1. Produces reproducible results
2. Creates an audit trail of the steps & reasoning

DISADVANTAGE

1. Does not extract all requirements and policy clauses if the documents are large. Needs logic for context chunking

PROMPT

STEP #1 **REQUIREMENTS EXTRACTION PROMPT**

- AS AN AUDITOR ASSESSING PRIVACY REGULATIONS IN SINGAPORE, YOUR TASK IS TO EXTRACT AND ANALYZE ALL REGULATORY REQUIREMENTS FROM THE DATA PRODUCTION TRUSTMARK (DPTM) CHECKLIST PROVIDED IN THE DOCUMENT.
- YOU NEED TO IDENTIFY EACH CHECKLIST ITEM'S CATEGORY, SUBCATEGORY, APPLICABLE STAKEHOLDER, AND THE CORRESPONDING POLICY DOCUMENT DETAILS FOR THE CATEGORY **"{CATEGORY_OF_INTEREST}"**.
- THE EXTRACTED INFORMATION SHOULD BE FORMATTED INTO A STRUCTURED JSON FORMAT, ADHERING STRICTLY TO THE SPECIFIED GUIDELINES.

STEP #2 **COMPLIANCE MAPPING PROMPT**

- AS AN AUDITOR ASSESSING PRIVACY REGULATIONS IN SINGAPORE, YOUR TASK IS TO EXTRACT CLAUSES FROM THE ORGANIZATION'S POLICY DOCUMENT THAT CORRESPOND TO SPECIFIC REGULATORY REQUIREMENTS.
- YOU MUST READ THROUGH THE ENTIRE POLICY DOCUMENT AND IDENTIFY THE EXACT CLAUSES THAT ALIGN WITH EACH REGULATORY REQUIREMENT LISTED.
- THE RESULTS SHOULD BE PRESENTED IN A STRUCTURED JSON FORMAT, DETAILING THE REGULATORY REQUIREMENT DESCRIPTION, THE CORRESPONDING POLICY CLAUSES, AND THE POLICY CATEGORY.

STEP #3 **REGULATORY PROMPTING**

- AS AN AUDITOR ASSESSING PRIVACY REGULATIONS IN SINGAPORE, YOUR TASK IS TO DETERMINE IF AN ORGANIZATION'S POLICY DOCUMENTATION ALIGNS WITH SPECIFIC REGULATORY REQUIREMENTS.
- USING THE PROVIDED REGULATORY REQUIREMENT AS THE MAJOR PREMISE AND THE CORRESPONDING POLICY DOCUMENTATION AS THE MINOR PREMISE, YOU NEED TO MAKE A JUDGEMENT (**NOT MET**, **PARTIALLY MET**, **MET**) AND PROVIDE AN EXPLANATION FOR IT.
- ADDITIONALLY, SUGGEST THE TYPE OF EVIDENCE NEEDED TO VERIFY COMPLIANCE, AND PRESENT THE FINDINGS IN A STRUCTURED JSON FORMAT.

Results of Stage 1
Regulation vs Enterprise Policy

INPUT	OUTPUT
STEP #1: Regulatory requirements were extracted using the DPTM document as an input	**35 requirements** were extracted from the DPTM
STEP #2: For Compliance Mapping, DPTM requirements list that were manually extracted (50 requirements) were provided	**All the 50 requirements were mapped to appropriate policy sections.** • Some of the clauses around **CONSENT** were not found in the privacy policy for Grab
STEP #3: The Regulatory Prompt was executed • regulatory requirements as the major premise, and • the compliance mapped policy assertions as the minor premise, and • the output as a judgement if the (policy) assertion satisfied the (regulatory) requirement.	JUDGEMENTS for most regulatory requirements were **`PARTIALLY MET`** because only the external-facing privacy policy document was provided. • For Grab privacy policy: **`MET: 1, NOT MET: 8`** and **`PARTIALLY MET: 41.`** • For DBS privacy policy, they were **`NOT MET: 3`** and **`PARTIALLY MET: 47.`** • **`NOT MET`** included regulatory requirements corresponding to Training, Data Breach management, DPIA.

DeepDive into results (Requirements POV)

REGULATION CATEGORY	DBS	GRAB
DATA PROTECTION OFFICER (2 REQUIREMENTS) 1. Appointment of competent DPO 2. DPO should have defined roles & responsibilities, with contact information easily available	**NOT MET** **PARTIALLY MET**	**PARTIALLY MET** **PARTIALLY MET**
TRAINING (1 REQUIREMENT) 1. Training programme for staff for awareness of data protection obligation	**NOT MET**	**PARTIALLY MET** because *We will take reasonable legal, organisational and technical measures to ensure that your Personal Data is protected.*
DATA PROTECTION IMPACT ASSESSMENT DPIA (2 REQUIREMENTS) 1. Documented policies & process on how it performs risk & impact assessments 2. Shall document DPIA conducted and ensure appropriate action plans are implemented for identified risks	**NOT MET** **NOT MET**	**PARTIALLY MET** Because of clause **"Conducting credit checks, screenings or due diligence checks as may be required under applicable law, regulation or directive; complying with all applicable laws, regulations, rules, directives, orders, instructions, guidance and requests from any local or foreign authorities, including regulatory, governmental, tax and law enforcement authorities or other authorities"**

DeepDive into results (Stakeholder POV)

REGULATION CATEGORY	DBS	GRAB
JOB APPLICANTS (3 REQUIREMENTS)	**PARTIALLY MET** in stage 1	**PARTIALLY MET** in stage 1
1. Data protection policies and practices for customer, job applicants, visitors via external DP notices	**PARTIALLY MET**	**PARTIALLY MET**
2. Publish & communicate data protection policies to job applicants via privacy notice on webpage or career page	**PARTIALLY MET**	**PARTIALLY MET**
3. Document policies & processes on receiving & responding to queries/complaints on collection, use and disclosure of personal data and how individuals submit these queries	**PARTIALLY MET**	**PARTIALLY MET**

Results of Stage 2
Regulation vs Recruitment Policy

INPUT	OUTPUT
STEP #4: For Compliance Mapping, subset of 3 DPTM requirements corresponding to job applicants	**The 3 requirements were mapped to appropriate recruitment policy sections.**
STEP #5: The Regulatory Prompt was executed • regulatory requirements as the major premise, and • the compliance mapped recruitment policy (in this case evidence) assertions as the minor premise, and • Context was added to state were that the evidence was picked from. • the output as a judgement if the (evidence) assertion satisfied the (regulatory) requirement.	JUDGEMENTS for most regulatory requirements were **`PARTIALLY MET`** because only the external-facing privacy policy document was provided. • For Grab privacy policy: **`MET: 1`** and **`PARTIALLY MET: 2.`** • For DBS privacy policy, **`MET: 1`** and **`PARTIALLY MET: 2.`** • **`PARTIALLY MET`** because of the elaborate (multiple?) requirements.

DeepDive into results (Stakeholder POV)

REGULATION CATEGORY	DBS	GRAB
JOB APPLICANTS (3 REQUIREMENTS)	**PARTIALLY MET** in stage 1 STAGE 2:	**PARTIALLY MET** in stage 1 STAGE 2:
1. Data protection policies and practices for customer, job applicants, visitors via external DP notices	**PARTIALLY MET**	**PARTIALLY MET**
2. Publish & communicate data protection policies to job applicants via privacy notice on webpage or career page	**COMPLETELY MET**	**COMPLETELY MET**
3. Document policies & processes on receiving & responding to queries/complaints on collection, use and disclosure of personal data and how individuals submit these queries	**PARTIALLY MET**	**PARTIALLY MET**

Limitations

- Human in the loop required to interpret the context for nuanced requirements.
- Regulation specific context needed to split single regulation requirements into multiple operational requirements.

COMPLETENESS & VERIFICATION

- Long form documents require different approaches
- Document parsing with embedded images are not yet supported by LLMs
- Different workflow for different regulatory compliance.

ENGINEERING CHALLENGES

Future Work

1. **COMPARATIVE TESTING**
   Testing was conducted on two well-known companies, DBS and Grab. The methodology could be extended to compare privacy policies of multiple companies across different industries.

2. **INTEGRATION OF ADVANCED TECHNIQUES**
   Integrate Retrieval Augmented Generation (RAG) with prompting to accurately extract complete and distinct sets of requirements from long form regulatory documents.

3. **UPDATING REGULATORY PROMPTING TECHNIQUES**
   Refine Regulatory Prompting to include specific standards for judgment to enhance the precision of compliance assessments

   E.g.. "Sufficient Appropriate Audit Evidence" by the Singapore Standards on Auditing (SSA 500)

Q&A

Thank you!

Contact:

Vidyaraman: vidyaraman@deepdivelabs.tech
Divya: divya@deepdivelabs.tech

We are looking for feedback & collaborators to enhance the technology and expand into new areas!

PROTOTYPE: https://deepdive-labs-trace.streamlit.app/

CODE: https://github.com/svidyaraman/TRACE-Framework/