



IBM Research – Australia

Cognitive Compliance for Personal Financial Advice

Sue Ann Chen, Wanita Sherchan, Nebula Alam, Simon Harris, Khoi-Nguyen Tran, Chris Butler, Simon Wail
IBM Research Australia

Risk analysis of unstructured documents

Challenge: Transforming unstructured data into insights and risk scores for advisers and auditors

Our data set

	Name	Source	No. of documents	No. of sentences	No. of tables	Document types
TRAINING	Bulk-SoA	Promontory	202	66340	8405	docx
TEST	ASIC-TEST	ASIC	20	4570	489	docx, pdf

Our approach

Mapped and interpreted components of document affecting compliance – tests for three risk indicators

Used bulk-SoA data set for training purposes and ASIC-provided data for performance testing

Our assumptions

KRI assessments based on sentences extracted from the SoA documents

KRI assessments based on textual data only, no numerical calculation

Risk analysis of unstructured documents

Deriving goal-advice insight from sentences within document

We identified several key risk indicators from the problem statement of this symposium.

Goal-advice appropriateness

Projection

Insurance

Table classification

1

Goal-advice appropriateness

What are the client's goals and have they been addressed by appropriate recommendation?

TEST 1: SoA contains recommendations

Yes, OK No, FLAG

TEST 2: SoA contains goals without recommendation

Yes, FLAG No, OK

Extracted domains match ASIC's details

Client name	Details	Extracted domains	Risk score
Wilma Flintstone	Strategy/social security, retirement	Insurance, Centrelink	Low
Keanu Roves	Super, investment, insurance	Insurance, Cashflow management	Low
Roger & Diana Rabbet	Retirement planning	Debt management, TTR, Centrelink	Low
Sean Conneray	Changes to risk profile	Risk profile, investments	Low
Timothy Dixon	Investment in wrap	-	Low
Grace Codd	Geared investment	Banking, risk profile, cashflow	Low
Pierce Brown & Paula Brown	Making a NCC contribution	Superannuation, TTR, investments	Low
Mary Poppins	Change in risk profile	Risk profile, cashflow management	Low
John and Jane Wick	Investments in SMSF	Self-managed super funds	Low
Cindy Rella	Life, TPD, IP, Trauma	Insurance, estate planning, debt	Low
Cruella de Ville	Life/TPD within Super	Insurance, estate planning	Low
Dr. Stephen Strange	IP in & out of super	Insurance	Low
Mrs Ygritte Snow & Mr John Snow	Replacement advice	Risk profile, insurance, taxation	Low
Daniel Cray & Eva Cray	Spousal contribution	Cashflow management, insurance	Low
Mon Gustave	Trauma only	Insurance	Low
Bruce Li	Pension rollover	Pension, insurance	Low
Anthea Saint & Lou Burns	Personal investment	Banking	Low
Jim Jones	Trauma only	Insurances, fees	Low
LeBron Jones	Life, TPD, IP, Trauma	Insurance	Low
Katniss Ye	Transition to retirement	TTR, debt management, insurance	Low

SoAs with none-to-low risk

SoAs with moderate risk of non-compliance

SoAs with high risk of non-compliance

Risk analysis of unstructured documents

Deriving evidence that insurance adequacy and appropriateness has been considered

2

Insurance

Does the advice contain statements that discuss personal, life, or income protection insurance?

TEST 1: *SoA contains recommendations for personal insurance (life, income protection, TPD, Trauma)*

TEST 2: *SoA defers/excludes insurance discussion*

Yes, OK No, FLAG
 Yes, FLAG No, OK

We identified several key risk indicators from the problem statement of this symposium.

- Goal-advice appropriateness
- Insurance**
- Projection
- Table classification

Client name	Scope of advice	Risk score
Wilma Flintstone	Full service	Orange
Keanu Roves	Full service	Orange
Roger & Diana Rabbet	Full service	Green
Sean Conneray	Investment only	Green
Timothy Dixon	Investment only	Green
Grace Codd	Investment only	Orange
Pierce Brown & Paula Brown	Superannuation only	Pink
Mary Poppins	Superannuation only	Green
John and Jane Wick	Superannuation only	Green
Cindy Rella	Superannuation and insurance	Orange
Cruella de Ville	Superannuation and insurance	Orange
Dr. Stephen Strange	Superannuation and insurance	Green
Mrs Ygritte Snow & Mr John Snow	Superannuation and insurance	Orange
Daniel Cray & Eva Cray	Full service	Green
Mon Gustave	Insurance only	Orange
Bruce Li	Superannuation only	Orange
Anthea Saint & Lou Burns	Investment only	Orange
Jim Jones	Insurance only	Orange
LeBron Jones	Insurance only	Orange
Katniss Ye	Full service	Orange

SoAs with none-to-low risk

SoAs with moderate risk of non-compliance

SoAs with high risk of non-compliance

Risk analysis of unstructured documents

Deriving evidence of client's position after implementation of advice

We identified several key risk indicators from the problem statement of this symposium.

Goal-advice appropriateness

Projection

Projection

Table classification

3

Projection

Does the advice contain statements describing the long term capital position/10 year projection?

TEST: SoA contains statements that evidence the adviser has considered the financial position of client post-advice implementation

Yes, OK

No, FLAG

Client name	Risk score
Wilma Flintstone	
Keanu Reeves	
Roger & Diana Rabbit	
Sean Connery	
Timothy Dixon	
Grace Codd	
Pierce Brown & Paula Brown	
Mary Poppins	
John and Jane Wick	
Cindy Rella	
Cruella de Ville	
Dr. Stephen Strange	
Mrs Ygritte Snow & Mr John Snow	
Daniel Cray & Eva Cray	
Mon Gustave	
Bruce Li	
Anthea Saint & Lou Burns	
Jim Jones	
LeBron Jones	
Katniss Ye	

All SoAs exhibit evidence of client's financial projection

Risk analysis of unstructured documents

Deriving insights from tables within document

We identified several key risk indicators from the problem statement of this symposium.

Goal-advice appropriateness

Projection

Insurance

Table classification

4

Table classification

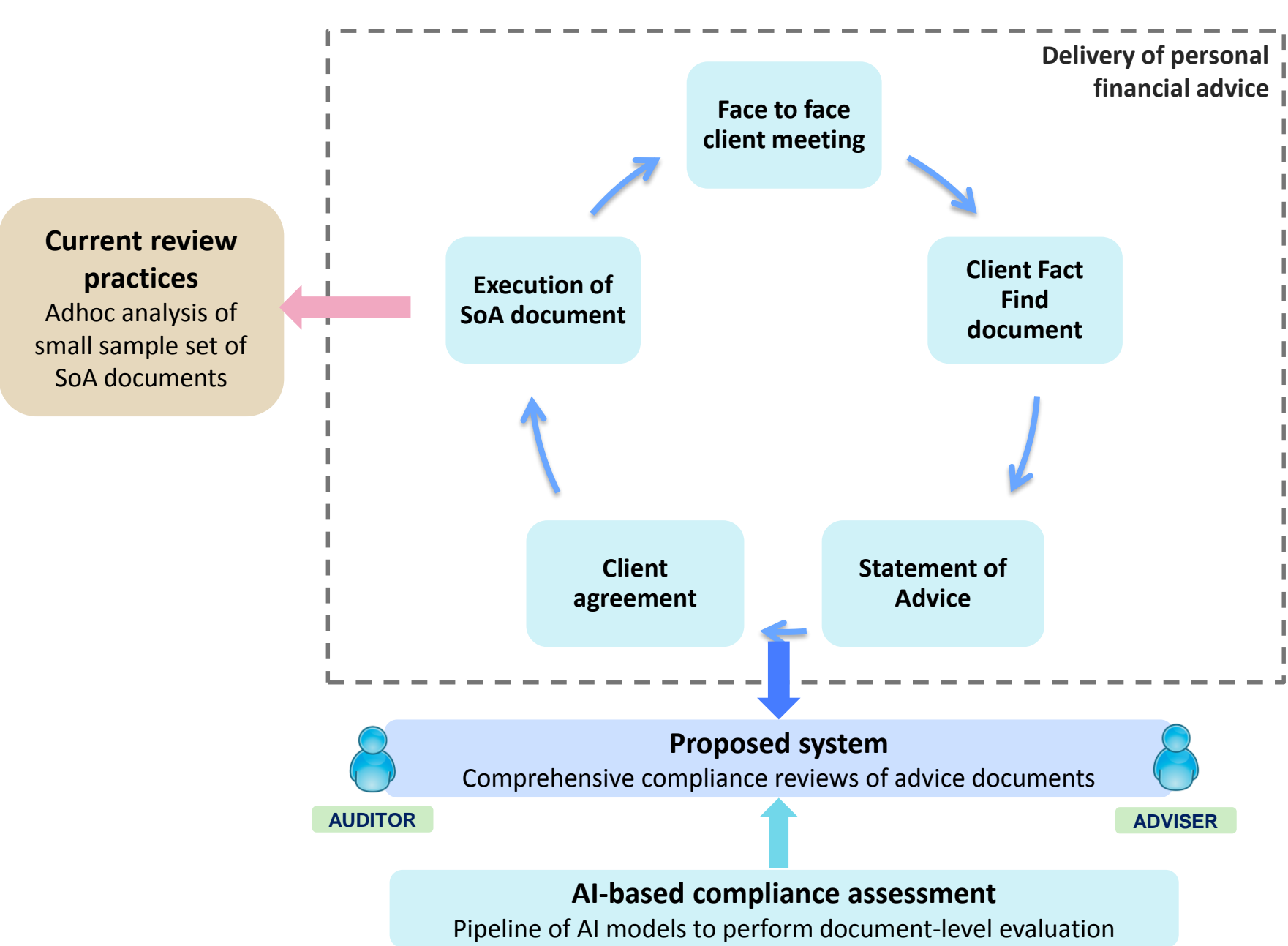
Does the advice contain tables relating to asset classes, cash flow analysis and client’s projection?

Table type	Examples	Tables in ASIC-Test
Asset class	Distributions of equity, share, stock, property	24
Cash flow	Details of salary, inflow, outflow and income	1
Projection	Estimated earnings or cashflow in the future	0
Other	Tables not fitting the above definitions	464

Possible future extensions: derive further insights from tabular data for additional risk indicator assessment

Key features of our system

Key differentiators of our adviser-auditor interactive compliance review system



Adviser-auditor usability
Increases trust and transparency in review process through dual-user capability, prior to advice delivery

AI techniques for assessment
Ability to extract and transform information into insights on key risk indicators from diverse data set

Unparalleled review speed
Microservice-based architecture allows for independent deployment, retraining and updating of AI models

Agile environment
Takes an average of 143 seconds to perform complete review of a SoA document, in comparison to 2 – 6 hours required manually

Continuous learning
Facilitates continuous learning and improves performance over time through gathering of user feedback

Challenges faced (and lessons learnt)

Some of the observations/challenges/barriers of how our solution worked on the ASIC data set

Data preparation

Clean extraction of sentences from SoA documents, especially from PDFs

Documents from ASIC contain less information and are less detailed from documents in bulk-SoA data set

More robust extraction approaches expected to improve data quality

Model development

KRI assessments are based only on textual data from SoAs, not including tables, images

Small training data set with different interpretation by annotators

Distribution of data between training and test data sets differs – e.g. explicit insurance discussions in training data not similarly observed in test data

Deployment

Balancing precision of models through high confidence threshold for the KRIs

Difficulties in determining the KRI rating for each SoA, as well as overall SoA risk score

Scaling the system to support large-scale adoption – model pipeline

Risk analysis of unstructured documents: a demo

Key takeaway messages and conclusions

Summary of the insights delivered by our solution

202 anonymised SoAs for training of deep learning models, applied to ASIC test data

Goal-advice appropriateness

Insurance

Projection

Table classification

Leveraged AI models to perform KRI assessments

Discussed insights and risk scores at SoA-level

Possible extension to derive implication of sentences on client's projection

Table classifier capable of determining presence of tables relating to cash flow, projection and asset allocation

Possible extension to derive further insights from tabular data

Challenges

Distribution of data between training and test data sets differs – e.g. explicit insurance discussions in training data not similarly observed in test data

Observations

Small training data set with different interpretation by annotators

Performance on ASIC data set

Complete and thorough analysis of SoA documents in minutes – taking 2 minutes per document

Ability for both adviser and compliance officer to review document compliance prior to advice delivery

Brings to the fore the documents that pose the highest risks of non-compliance