

# ris.es.io

## software solutions

Legal Tech: Advanced Analytics Applied  
Verticalized Solution in Legal Domain

13 Dec 2016  
Version: Ver 0.5

# LegalTech

## Similarity Search and Auto Summarization Solution

## Needs and Challenges

- Time consuming and inefficient processes to find similar documents
- Case summarization take lot for efforts in law firms as well as contracts in BFSI
- There is no predictability of the case outcome, it impacts the planning
- Low utilization of analytics and cloud

## Approach

- Advanced analytics techniques to search similar docs
- Use of NLP / Text Analytics to automatically generate the summary of legal doc / contract
- Use of predictive analytics
- Risk solution - Improve the understanding of risk, analyse entire portfolio of documents

## Techniques

- Machine learning / AI based Algos
- Natural Languages Processing / Text Analytics
- Predictive Analytics

## Solution

- e.g. Finding nearest matching docs using various techniques

## Clientele

- LPO/KPO, Legal Departments in BSFI



# The Difference

Same workflows but different and more relevant solutions

# Finding Relevant Documents

## Current Search

- Text based
- A “word” is searched through all documents and ANY document containing the “word” is presented

## Current Outcome

- Slow, mostly manual processing
- Brings many un-important results
- Weeding through them takes time

## Improved Search

- Theme based
- Documents are pre-populated with themes, which are matched with the query “word”

## Improved Outcome

- Faster
- Results based on ranked importance
- The top results are most relevant

# Finding Similar Documents

## Current Method

- Search based, mostly manual
- A “word” is searched through all documents and ANY document containing the “word” is presented

## Current Outcome

- Slow
- Bring many un-important results
- Weeding through them takes time

## Improved Similarity

- All documents are pre-populated with pair-wise similarity scores
- For a given document, similar documents list is readily available

## Improved Outcome

- Faster
- Results based on ranked importance
- The top results are most relevant

# Finding Summary of a Document

## Current Method

- Summary prepared by Experts
- “Head Notes” of a judgement

## Current Outcome

- Slow
- Not scalable
- Expensive and subjective

## Improved Summary

- Sentences are extracted, ranked, ordered and presented
- Domain knowledge is incorporated

## Improved Outcome

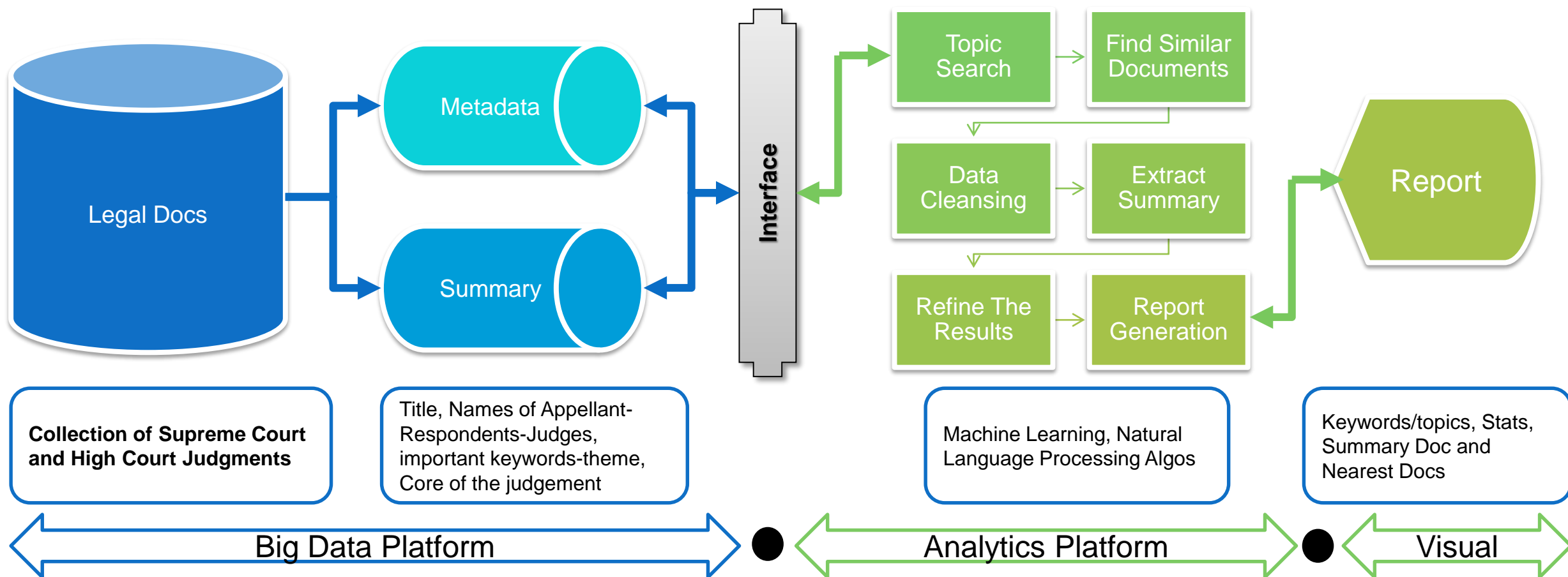
- Faster
- Highly scalable
- Consistent and customizable

# The Approach

Incorporating new technologies, algorithms and Deep Learning

# Solution Design

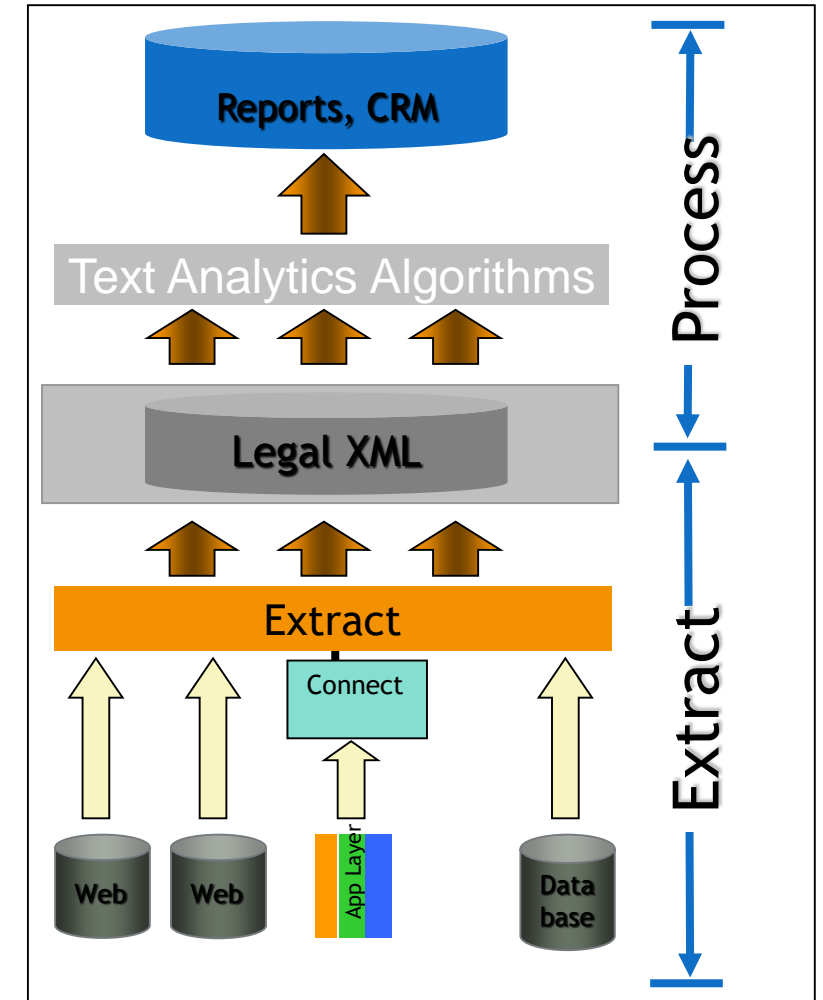
Natural Language Processing (NLP) Pipeline applied to LegalTech





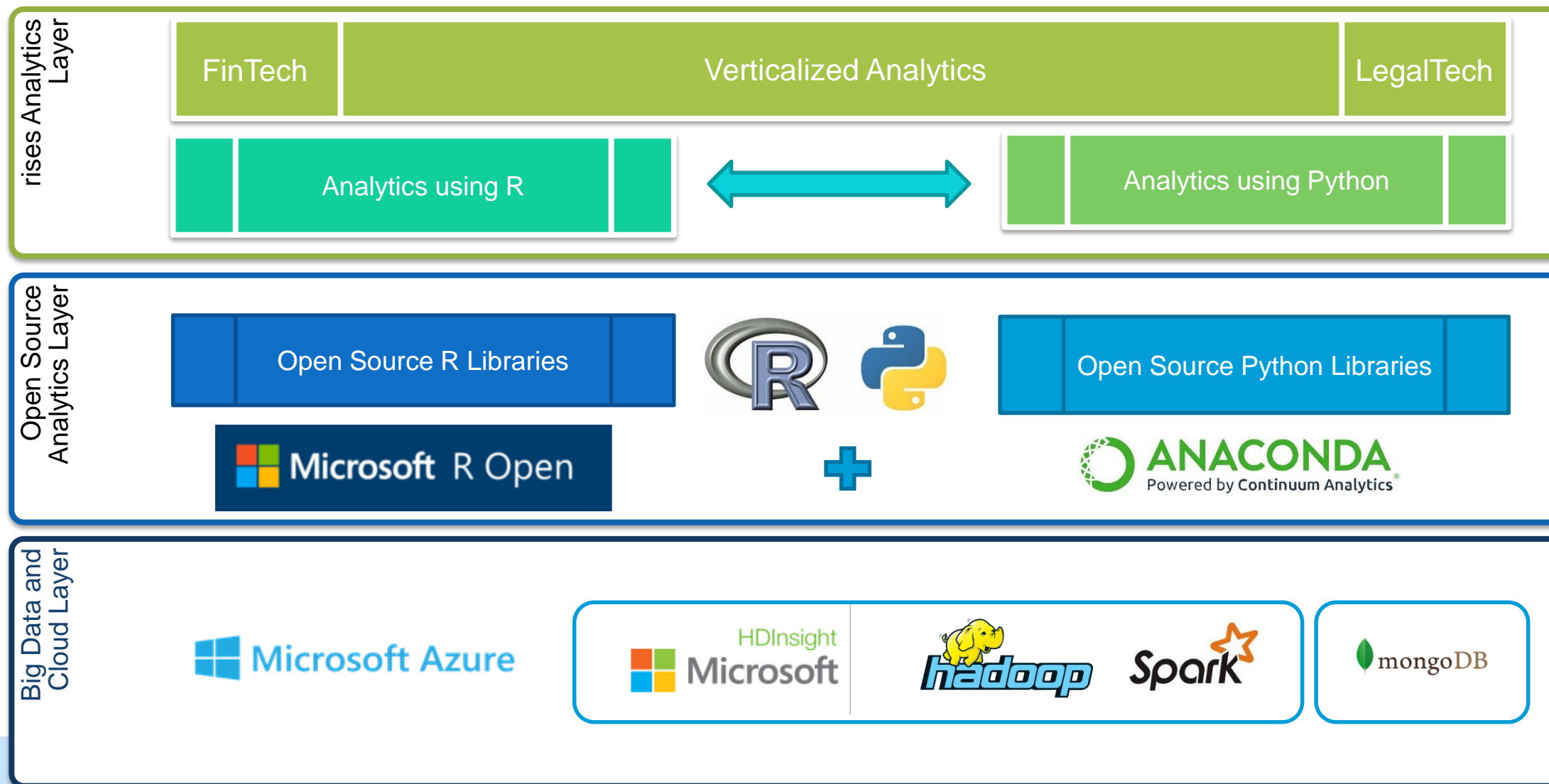
# Legal Analytics Workflows

- Legal Text Sources : Web, Databases, Local
- Pre-processing
  - Html -> Text
  - Db -> Text
- Linguistic Pre-Processing
  - Tokenizing
  - Stop word removal
- Text Algorithms: Search, Similarity, Summary
- Results reporting



# riser Analytics Platform

## Platform Architecture



# Searching

Matching not just a “word” but to the topics or themes of the document

# Searches by Topics

Robot Lawyer

Robot Lawyer

income

Searched Judgements

- supremecourt\_7\_.txt
- supremecourt\_8\_.txt
- supremecourt\_12\_.txt
- supremecourt\_18\_.txt
- supremecourt\_19\_.txt
- supremecourt\_24\_.txt

Similar Judgements

- supremecourt\_8\_.txt
- supremecourt\_62\_.txt
- supremecourt\_38\_.txt
- supremecourt\_97\_.txt
- supremecourt\_72\_.txt

Keywords : income rubber trees

Send comments to [yogeshkulkarni@yahoo.com](mailto:yogeshkulkarni@yahoo.com)

Judgement

PETITIONER:THE COMMISSIONER OF AGRICULTURALINCOME-TAX Vs.RESPONDENT:THE CALVARY MOUNT ESTATES (PRIVATE)LTD.DATE OF JUDGMENT:15/12/1960BENCH:KAPUR, J.L.BENCH:KAPUR, J.L.HIDAYATULLAH, M.SHAH, J.C.CITATION: 1961 AIR 1099 1961 SCR (3) 285ACT:Agricultural Income Tax-Rubber Plantation-Expenditure onimmature trees-Whether permissible deduction-Madras Plan-tations Agricultural Income-tax Act, 1955 (Mad. V of 1955),s. 5(e).HEADNOTE:The assessee owned an Estate of 590 acres out of which 235acres were occupied by immature non-bearing rubber trees,for the maintenance and upkeep of which the respondentclaimed expenses from out of the income, which was allowedboth by the Agricultural Income Tax Tribunal and the HighCourt. The appellant came up by special leave.Held, that the provisions of s. 5(e) of the Madras Planta-tions Agricultural Income Tax Act, 1955 (Mad. V of 1955),applicable to the present case, and those of s. 5(1) of theTravancore-Cochin Agricultural Income Tax Act, 1950 (Tr.Co. XXII of 1950) being the same, the judgment in TravancoreRubber & Tea Co. Ltd. v. The Commissioner of AgriculturalIncome-tax, Kerala, in which the question of deductibilityof sums expended for purposes of forking, manuring etc. ofimmature rubber trees had been decided, will govern thiscase.Travancore Rubber & Tea Co. Ltd. v. The Commissioner ofAgricultural Income-tax, Kerala, [1961] 3 S.C.R. 279,applied.JUDGMENT:CIVIL APPELLATE JURISDICTION: Civil Appeal No, 145 of 1960.Appeal by special leave from the judgment and order datedMarch 18, 1958, of the Kerala High Court in Tax RevisionCase No. 12 of 1957.V.A. Seyid Muhamad and Sardar Bahadur, for the appellant.C.K. Daphtary, Solicitor-General of India, ThomasVellapally, S. N. Andley, J.

Summary

Inplace of "for the purpose of the plantation" in the former,the words "for the purpose of287deriving the agricultural income" are used in the latter.If anything the words of the former Act are more favourableto the respondent.In

# Topics by LDA (Latent Dirichlet Allocation)

- Discovering hidden topical patterns that are present across the collection

**Doc 1:** Income generated from Rent was not disclosed.

**Doc 2:** Present litigation does not come under Rent control act.

**Doc 3:** Appellant violated the Sales agreement.

**Doc 4:** Proceeds of the Sale of the Motor were shown as Income

**Doc 5:** Sales and disclosed Income did not match.

**Topic 1:** 30% Income, 15% Rent,

**Topic 2:** 10% Sales, 10% Motor, 5% Income

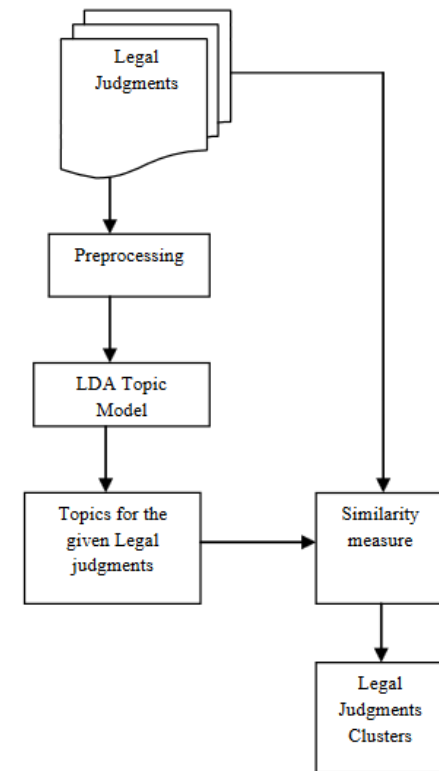
**Docs 1 and 2:** 100% Topic 1

**Docs 3 and 4:** 100% Topic 2

**Doc 5:** 70% Topic 1, 30% Topic 2



International Journal of Applied Information Systems  
Foundation of Computer Science FCS, New York, USA  
Volume 2-No 6, 2012 – www.ijais.org





# Topics extraction – presented as keywords

The screenshot displays the 'Robot Lawyer' application window. On the left, there is a search bar with the text 'income' and a 'Search' button. Below it, a list of 'Searched Judgements' includes 'supremecourt\_7\_.txt' (highlighted), 'supremecourt\_8\_.txt', 'supremecourt\_12\_.txt', 'supremecourt\_18\_.txt', 'supremecourt\_19\_.txt', and 'supremecourt\_24\_.txt'. Below this is a 'Similar Judgements' list with 'supremecourt\_8\_.txt', 'supremecourt\_62\_.txt', 'supremecourt\_38\_.txt', 'supremecourt\_97\_.txt', and 'supremecourt\_72\_.txt'. At the bottom left, a box labeled 'Keywords : income rubber trees' is circled in blue, with a red arrow pointing from the highlighted 'supremecourt\_7\_.txt' file to it. The main area on the right is titled 'Judgement' and contains the full text of a legal case: 'PETITIONER:THE COMMISSIONER OF AGRICULTURALINCOME-TAX Vs.RESPONDENT:THE CALVARY MOUNT ESTATES (PRIVATE)LTD.DATE OF JUDGMENT:15/12/1960BENCH:KAPUR, J.L.BENCH:KAPUR, J.L.HIDAYATULLAH, M.SHAH, J.C.CITATION: 1961 AIR 1099 1961 SCR (3) 285ACT:Agricultural Income Tax-Rubber Plantation-Expenditure onimmature trees-Whether permissible deduction-Madras Plan-tations Agricultural Income-tax Act, 1955 (Mad. V of 1955),s. 5(e).HEADNOTE:The assessee owned an Estate of 590 acres out of which 235acres were occupied by immature non-bearing rubber trees,for the maintenance and upkeep of which the respondentclaimed expenses from out of the income, which was allowedboth by the Agricultural Income Tax Tribunal and the HighCourt. The appellant came up by special leave.Held, that the provisions of s. 5(e) of the Madras Planta-tions Agricultural Income Tax Act, 1955 (Mad. V of 1955),applicable to the present case, and those of s. 5(1) of theTravancore-Cochin Agricultural Income Tax Act, 1950 (Tr.Co. XXII of 1950) being the same, the judgment in TravancoreRubber & Tea Co. Ltd. v. The Commissioner of AgriculturalIncome-tax, Kerala, in which the question of deductibilityof sums expended for purposes of forking, manuring etc. ofimmature rubber trees had been decided, will govern thiscase.Travancore Rubber & Tea Co. Ltd. v. The Commissioner ofAgricultural Income-tax, Kerala, [1961] 3 S.C.R. 279,applied.JUDGMENT:CIVIL APPELLATE JURISDICTION: Civil Appeal No, 145 of 1960.Appeal by special leave from the judgment and order datedMarch 18, 1958, of the Kerala High Court in Tax RevisionCase No. 12 of 1957.V.A. Seyid Muhamad and Sardar Bahadur, for the appellant.C.K. Daphtary, Solicitor-General of India, ThomasVellapally, S. N. Andley, J.

Below the judgment text is a 'Summary' section with the text: 'Inplace of "for the purpose of the plantation" in the former,the words "for the purpose of287deriving the agricultural income" are used in the latter.If anything the words of the former Act are more favourableto the respondent.In

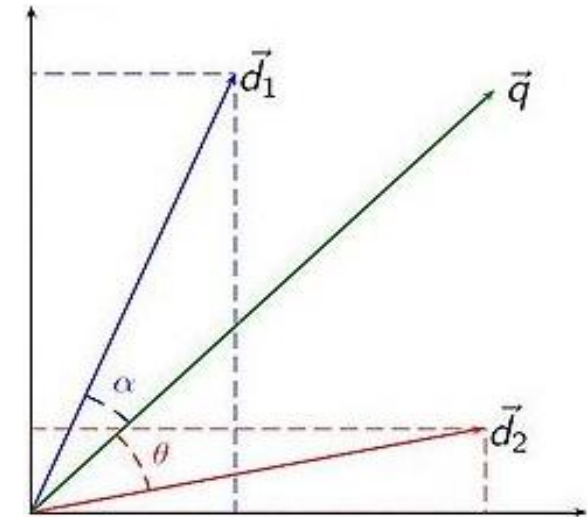
At the bottom left of the window, it says 'Send comments to yogeshkulkarni@yahoo.com'.

# Document Similarity

Preparing pair-wise similarity scores

# Similarity by TF-IDF and Cosine

- TF-IDF (Term Frequency Inverse Document Frequency)
- Input: Set of documents, each Vectorised by TFIDF
- Similarity between documents  $d_1$  and  $d_2$  is angle between them
- If angle is 0, they are similar (exact)
- If angle is 90 they are different.
- Query  $q$  can also be evaluated to see if its closer to  $d_1$  or  $d_2$  then return the closer one.





# Pair-wise Similarity Result

The screenshot shows the 'Robot Lawyer' application window. On the left, there is a search bar with the text 'income' and a 'Search' button. Below it, a list of 'Searched Judgements' includes 'supremecourt\_7\_.txt' (highlighted in blue), 'supremecourt\_8\_.txt', 'supremecourt\_12\_.txt', 'supremecourt\_18\_.txt', 'supremecourt\_19\_.txt', and 'supremecourt\_24\_.txt'. A red arrow points from 'supremecourt\_7\_.txt' to a 'Similar Judgements' list below it, which includes 'supremecourt\_8\_.txt', 'supremecourt\_62\_.txt', 'supremecourt\_38\_.txt', 'supremecourt\_97\_.txt', and 'supremecourt\_72\_.txt'. The 'Similar Judgements' list is circled in blue. Below the lists, the 'Keywords' are 'income rubber trees'. On the right, the 'Judgement' text is displayed, starting with 'PETITIONER:THE COMMISSIONER OF AGRICULTURALINCOME-TAX Vs.RESPONDENT:THE CALVARY MOUNT ESTATES (PRIVATE)LTD.' and ending with 'Daphtary, Solicitor-General of India, ThomasVellapally, S. N. Andley, J.'. Below the judgement, there is a 'Summary' section with the text 'Inplace of "for the purpose of the plantation" in the former,the words "for the purpose of287deriving the agricultural income" are used in the latter.If anything the words of the former Act are more favourableto the respondent.In'. At the bottom left, there is a link to 'Send comments to yogeshkulkarni@yahoo.com'.

Robot Lawyer

income Search

Searched Judgements

- supremecourt\_7\_.txt
- supremecourt\_8\_.txt
- supremecourt\_12\_.txt
- supremecourt\_18\_.txt
- supremecourt\_19\_.txt
- supremecourt\_24\_.txt

Similar Judgements

- supremecourt\_8\_.txt
- supremecourt\_62\_.txt
- supremecourt\_38\_.txt
- supremecourt\_97\_.txt
- supremecourt\_72\_.txt

Keywords : income rubber trees

Judgement

PETITIONER:THE COMMISSIONER OF AGRICULTURALINCOME-TAX Vs.RESPONDENT:THE CALVARY MOUNT ESTATES (PRIVATE)LTD.DATE OF JUDGMENT:15/12/1960BENCH:KAPUR, J.L.BENCH:KAPUR, J.L.HIDAYATULLAH, M.SHAH, J.C.CITATION: 1961 AIR 1099 1961 SCR (3) 285ACT:Agricultural Income Tax-Rubber Plantation-Expenditure onimmature trees-Whether permissible deduction-Madras Plan-tations Agricultural Income-tax Act, 1955 (Mad. V of 1955),s. 5(e).HEADNOTE:The assessee owned an Estate of 590 acres out of which 235acres were occupied by immature non-bearing rubber trees,for the maintenance and upkeep of which the respondentclaimed expenses from out of the income, which was allowedboth by the Agricultural Income Tax Tribunal and the HighCourt. The appellant came up by special leave.Held, that the provisions of s. 5(e) of the Madras Planta-tions Agricultural Income Tax Act, 1955 (Mad. V of 1955),applicable to the present case, and those of s. 5(1) of theTravancore-Cochin Agricultural Income Tax Act, 1950 (Tr.Co. XXII of 1950) being the same, the judgment in TravancoreRubber & Tea Co. Ltd. v. The Commissioner of AgriculturalIncome-tax, Kerala, in which the question of deductibilityof sums expended for purposes of forking, manuring etc. ofimmature rubber trees had been decided, will govern thiscase.Travancore Rubber & Tea Co. Ltd. v. The Commissioner ofAgricultural Income-tax, Kerala, [1961] 3 S.C.R. 279,applied.JUDGMENT:CIVIL APPELLATE JURISDICTION: Civil Appeal No. 145 of 1960.Appeal by special leave from the judgment and order datedMarch 18, 1958, of the Kerala High Court in Tax RevisionCase No. 12 of 1957.V.A. Seyid Muhamad and Sardar Bahadur, for the appellant.C.K. Daphtary, Solicitor-General of India, ThomasVellapally, S. N. Andley, J.

Summary

Inplace of "for the purpose of the plantation" in the former,the words "for the purpose of287deriving the agricultural income" are used in the latter.If anything the words of the former Act are more favourableto the respondent.In

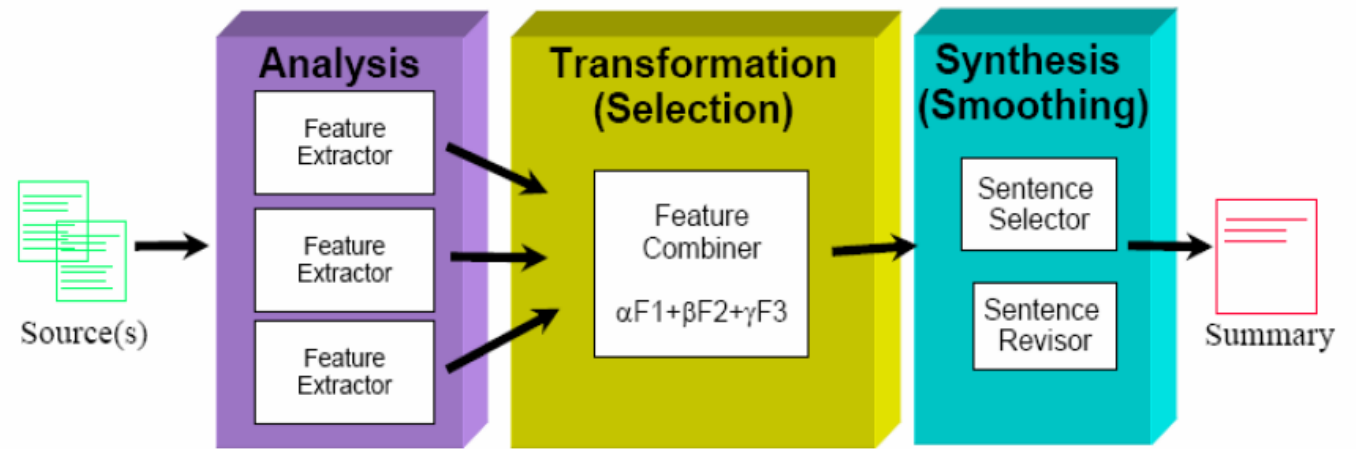
Send comments to yogeshkulkarni@yahoo.com

# Documents Summarization

Judgements to Headnotes, Case reports to abstracts

# Extractive Rank-based Summarization

- The process of condensing a source text into a shorter version preserving its information content is called summarization.
- The motivation Here is to build such tool which is computationally efficient and creates summaries automatically.
- Current Applications
  - Multimedia news summaries
  - Abstracts of Research papers
  - Minutes of the meetings
  - Book reviews



# Rank-based Extractive Summarization

The screenshot shows a web application titled "Robot Lawyer". It features a search bar with the text "income" and a "Search" button. Below the search bar, there are two lists of "Searched Judgements" and "Similar Judgements". The "Searched Judgements" list includes "supremecourt\_7.txt" (highlighted), "supremecourt\_8.txt", "supremecourt\_12.txt", "supremecourt\_18.txt", "supremecourt\_19.txt", and "supremecourt\_24.txt". The "Similar Judgements" list includes "supremecourt\_8.txt", "supremecourt\_62.txt", "supremecourt\_38.txt", "supremecourt\_97.txt", and "supremecourt\_72.txt". A red arrow points from the highlighted "supremecourt\_7.txt" to the "Judgement" text area. The "Judgement" text area contains a detailed legal case summary. Below the judgement, there is a "Summary" section which is circled in blue. The summary text is: "Inplace of 'for the purpose of the plantation' in the former, the words 'for the purpose of deriving the agricultural income' are used in the latter. If anything the words of the former Act are more favourable to the respondent. In". At the bottom of the application, there is a footer that says "Send comments to yogeshkulkarni@yahoo.com".

**Robot Lawyer**

income Search

Searched Judgements

- supremecourt\_7.txt
- supremecourt\_8.txt
- supremecourt\_12.txt
- supremecourt\_18.txt
- supremecourt\_19.txt
- supremecourt\_24.txt

Similar Judgements

- supremecourt\_8.txt
- supremecourt\_62.txt
- supremecourt\_38.txt
- supremecourt\_97.txt
- supremecourt\_72.txt

Keywords : income rubber trees

Send comments to yogeshkulkarni@yahoo.com

**Judgement**

PETITIONER:THE COMMISSIONER OF AGRICULTURALINCOME-TAX Vs.RESPONDENT:THE CALVARY MOUNT ESTATES (PRIVATE)LTD.DATE OF JUDGMENT:15/12/1960BENCH:KAPUR, J.L.BENCH:KAPUR, J.L.HIDAYATULLAH, M.SHAH, J.C.CITATION: 1961 AIR 1099 1961 SCR (3) 285ACT:Agricultural Income Tax-Rubber Plantation-Expenditure onimmature trees-Whether permissible deduction-Madras Plan-tations Agricultural Income-tax Act, 1955 (Mad. V of 1955),s. 5(e).HEADNOTE:The assessee owned an Estate of 590 acres out of which 235acres were occupied by immature non-bearing rubber trees,for the maintenance and upkeep of which the respondentclaimed expenses from out of the income, which was allowedboth by the Agricultural Income Tax Tribunal and the HighCourt. The appellant came up by special leave.Held, that the provisions of s. 5(e) of the Madras Planta-tions Agricultural Income Tax Act, 1955 (Mad. V of 1955),applicable to the present case, and those of s. 5(1) of theTravancore-Cochin Agricultural Income Tax Act, 1950 (Tr.Co. XXII of 1950) being the same, the judgment in TravancoreRubber & Tea Co. Ltd. v. The Commissioner of AgriculturalIncome-tax, Kerala, in which the question of deductibilityof sums expended for purposes of forking, manuring etc. ofimmature rubber trees had been decided, will govern thiscase.Travancore Rubber & Tea Co. Ltd. v. The Commissioner ofAgricultural Income-tax, Kerala, [1961] 3 S.C.R. 279,applied.JUDGMENT:CIVIL APPELLATE JURISDICTION: Civil Appeal No, 145 of 1960.Appeal by special leave from the judgment and order datedMarch 18, 1958, of the Kerala High Court in Tax RevisionCase No. 12 of 1957.V.A. Seyid Muhamad and Sardar Bahadur, for the appellant.C.K. Daphtary, Solicitor-General of India, ThomasVellapally, S. N. Andley, J.

**Summary**

Inplace of "for the purpose of the plantation" in the former, the words "for the purpose of deriving the agricultural income" are used in the latter. If anything the words of the former Act are more favourable to the respondent. In

# The Road ahead

Challenges to get to the target

# Needs

- Most NLP results are subjective
- Need “training data” (input-output pairs) to train the Deep Learning model.
- Need Cases, Judgements, Reports along with their correct summaries.

# Challenges

- Parsing Text : Finding citations, rules, decisions, issues, meta data like dates, judges – appellant - respondent names, etc. Formats vary thus extraction becomes complex.
- Algorithmic complexities: Need to build more optimized searches, theme extraction, Deep Learning model development, etc.
- Infrastructure: Soft (Hadoop/Spark, Databases) and Hard (Cloud/Servers)





rises.io  
software solutions

Company: rises software solutions

- Website: [rises.io](https://rises.io)

Contact Info: Pune, India

- Ajit Deshpade, Founder
- e-mail: [ajit.deshpande@rises.io](mailto:ajit.deshpande@rises.io)
- Phone: +91 98224 48602

 Thank you!