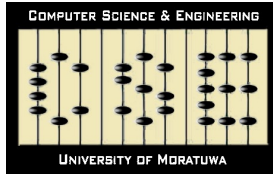


idstack.one

The Common Protocol for Document Verification built on Digital Signatures

<http://www.idstack.one>



idstack^{one}

The Common Protocol for Document Verification built on Digital Signatures

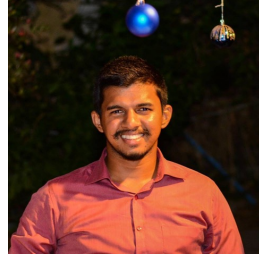
Our Team



Chanaka Lakmal



Sachithra Dangalla



Chamin Wickramaratna



Chandu Herath

Supervisors



Prof. Gihan Dias

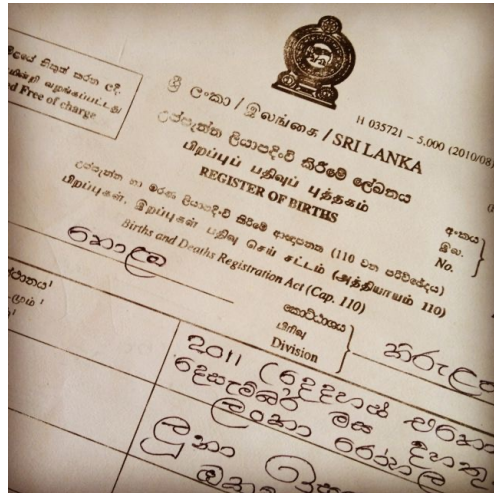
PhD (UCD), MSc (UCSB),
BSc Eng. (Hons)
(Moratuwa), MIE (SL),
CEng



Eng. Dr. Shantha Fernando

PhD (TU Delft), MPhil
(Moratuwa), BSc Eng. (Hons)
(Moratuwa), MIE (SL), MIEE
(UK), CEng

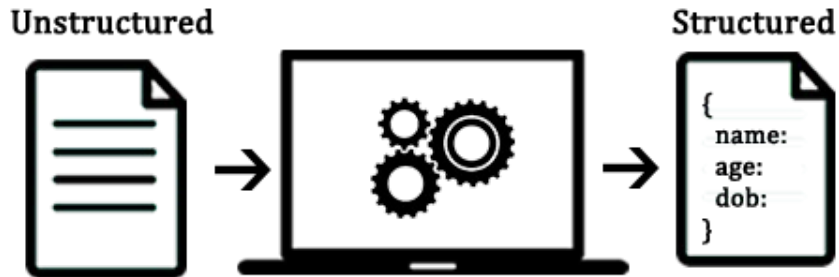
Problem



In countries like Sri Lanka the document verification process is highly based on the **printed copies**.

Problem : 1/4

Documents have **unpredictable layouts** that have raw information and there is **no mechanism** in the current research field to **extract them into common format of machine readable**.



Problem : 2/4

The document is subjected to a **process of verification** by an authority; using a **hand written signature**.

Another verified copy?

The whole process is **repeated**.



Problem : 3/4

Non existence of a mechanism in the current research field that can **calculate trustworthiness** of a individual document.

Sometimes the **fraudulent activities** that are followed by certain citizens **leading to doubtfulness of the authenticity** of the content.



Problem : 4/4

Non existence of a mechanism in the current research field that can **calculate correlation** among set of documents belonging to a person.



Revolutionize





Document Processing

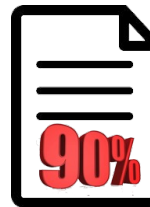
idstack.one

The Common Protocol for Document Verification built on Digital Signatures

Features



-  **Extraction**
Extract and create structured document
-  **Validation**
Validate document using digital signatures
-  **Confidence score**
Measure **trustworthiness** of a individual document
-  **Correlation score**
Measure **correlation** among set of documents



Functionalities



➔ Extraction

Provide an application to extract data

- from a paper document
- from a digital document (PDF)

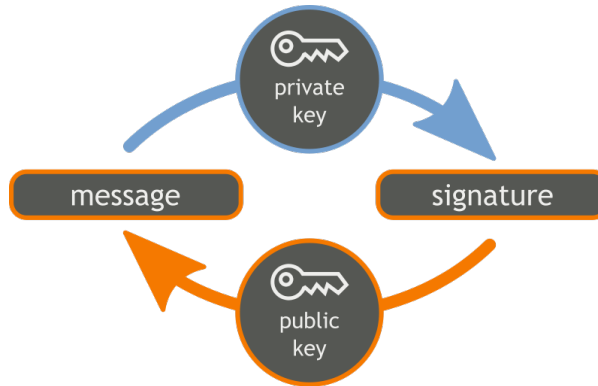
and convert into a Machine Readable (.idstack) format and verify it.

Functionalities



➔ Validation

Provide an application to validate and verify a document with the use of digital signatures.



Functionalities



Confidence Score

Provide an application with algorithm to calculate the trustworthiness of a individual document



Functionalities



➔ Correlation Score

Provide an application with algorithm to calculate the correlation among a set of documents



Design

All the **citizens** are grouped into one of the followings.

Anyone can play any role at **anytime**.



Owner



Extractor



Validator



Relying Party

Design

PDF or Paper Document



Citizens have their **own documents** as PDF or Paper Document.

This is for human beings to look at it.

VS

Machine Readable Document

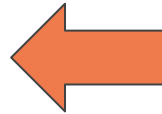


All the documents are converted into **.idstack** format by **Extractor** and given back to **Owner** with **Digital Signature**.

This is for automatically processing by the computers.

Design

```
1 {
2   "meta_data": {
3     "name": "IDStack",
4     "version": "1.0",
5     "document_id": "",
6     "document_type": "",
7     "issuer": {
8       "organization": "",
9       "email": "",
10      "url": ""
11    }
12  },
13  "content": {},
14  "extractor": {
15    "id": "",
16    "signature": {
17      "message_digest": "",
18      "url": ""
19    }
20  },
21  "validators": [
22    {
23      "id": "",
24      "signature": {
25        "message_digest": "",
26        "url": ""
27      },
28      "signed_content": "",
29      "signed_signatures": []
30    }
31  ]
32 }
```



Machine Readable Document

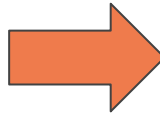


All the documents are converted into `.idstack` format by **Extractor** and given back to **Owner** with **Digital Signature**.

This is for automatically processing by the computers.

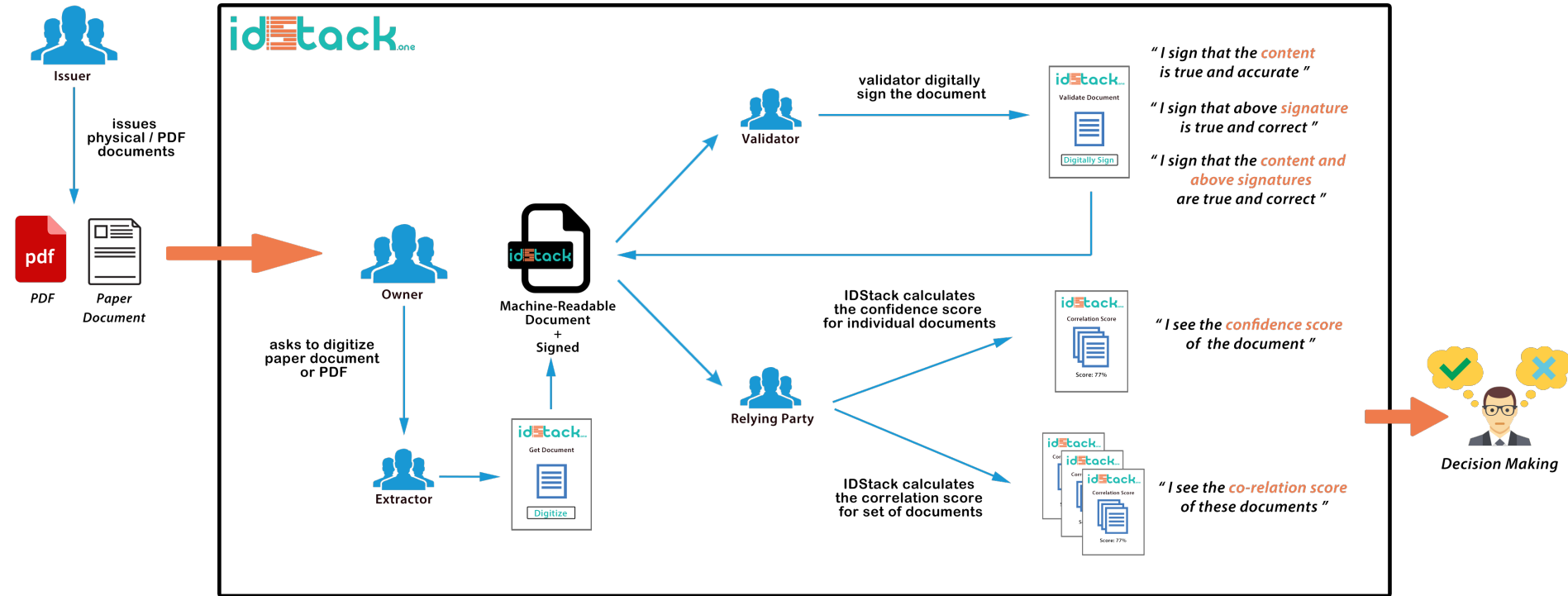
Design

```
1 {
2   "meta_data": {
3     "name": "IDStack",
4     "version": "1.0",
5     "document_id": "",
6     "document_type": "",
7     "issuer": {
8       "organization": "",
9       "email": "",
10      "url": ""
11    }
12  },
13  "content": {},
14  "extractor": {
15    "id": "",
16    "signature": {
17      "message_digest": "",
18      "url": ""
19    }
20  },
21  "validators": [
22    {
23      "id": "",
24      "signature": {
25        "message_digest": "",
26        "url": ""
27      },
28      "signed_content": "",
29      "signed_signatures": []
30    }
31  ]
32 }
```

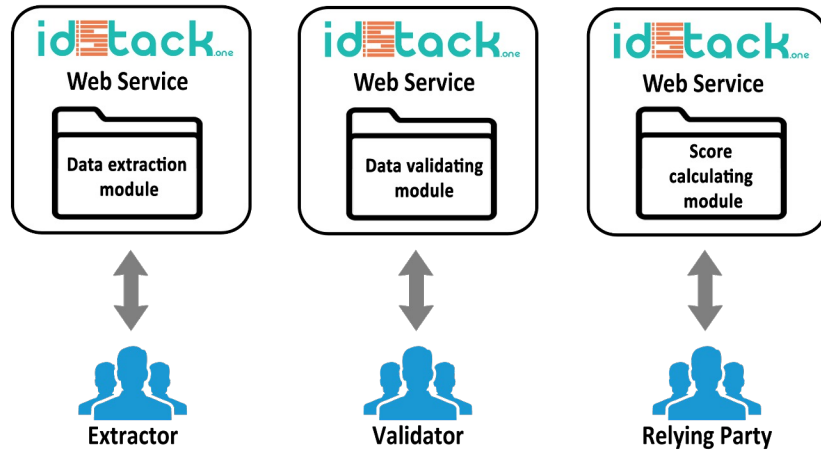


```
1 {
2   "meta_data": {
3     "name": "IDStack",
4     "version": "1.0",
5     "document_id": "a8heEgg",
6     "document_type": "passport",
7     "issuer": {
8       "organization": "department of immigration and emmigration",
9       "email": "controller@immigration.gov.lk",
10      "url": "http://www.immigration.gov.lk/usr/local/idstack/extractor/cert/pub/9e18a0b3-45cd-41e9-81a2-073e91e0c1b0.cer"
11    }
12  },
13  "content": {
14    "university": "University of Moratuwa",
15    "faculty": "Engineering",
16    "degree": "BSc Eng Hons",
17    "name": {
18      "surname": "Dangalla",
19      "initials": "DADJS"
20    },
21    "index_no": "130094R",
22    "address": {
23      "line1": "9A, De Mel Watta Road, Koswatta",
24      "line2": "Nawala"
25    },
26    "nic": "937741180V"
27  },
28  "extractor": {
29    "id": "b178dad3",
30    "signature": {
31      "message_digest":
32      "MIAGCSqGSIB3DQEHAQCAMIACAQEXDzANBglghkgBZQMEAgEFADCBABgkqhkiG9w0BB
33      wGggCSABIBN3sidW5pdmVyc2l0eSI611VuaXZlcmluNpdHkgb2YgTW9yYXRidZElL",
34      "url": "http://www.mrt.ac.lk/usr/local/idstack/extractor/cert/pub/9e18a0b3-45cd-41e9-81a2-073e91e0c1b0.cer"
35    }
36  },
37  "validators": [
38    {
39      "id": "lUed5eg",
40      "signature": {
41        "message_digest":
42        "MIIFRjCCBC6gAwIBAgIRAMxtHV1q5YH597HXLW01n14wDQXJKoZlIhvcNAQELBQA
43        wgZAXcCAJBgNVBAYTAkdCMRswGQYDVQQIEXJHcmVhdG9yIE1hbmNoZ",
44      "url": "http://idstack.one/gtzKCGE0xH87M9P"
45    }
46  ]
47 }
```

idstack.one, The Common Protocol



Architecture



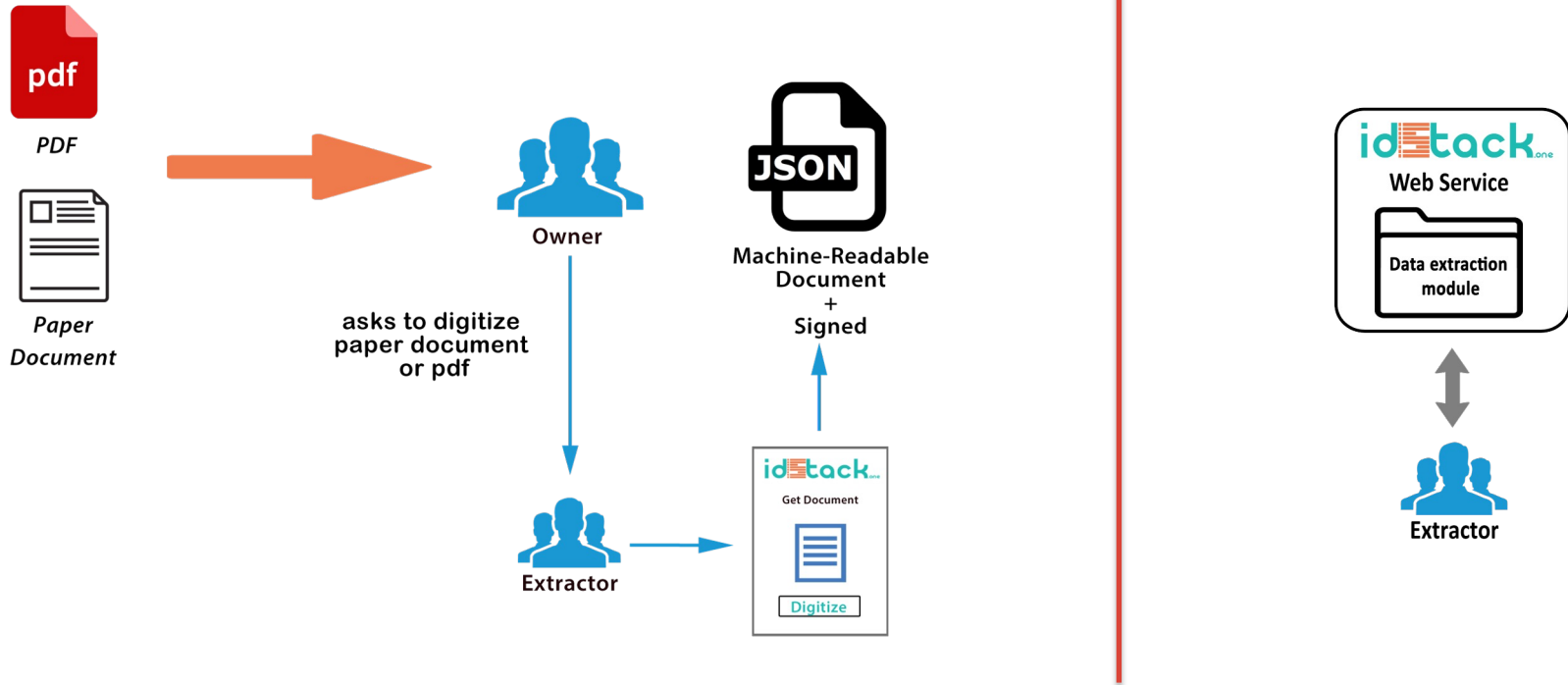
> Web Service

Spring Boot REST API with 3 modules mapped into 3 user roles

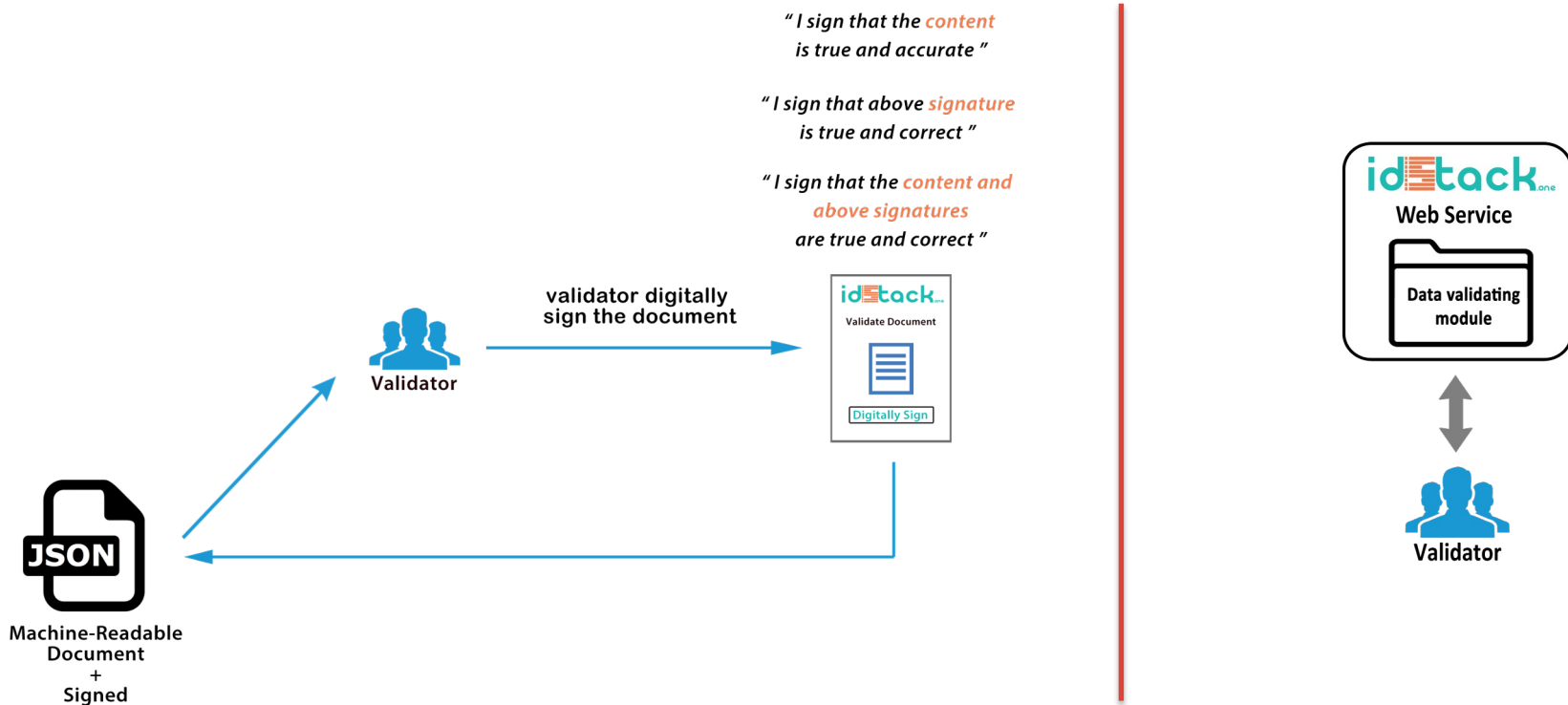
Data extraction module	Extractor
Data validating module	Validator
Score calculating module	Relying Party

> Client Applications

Desktop applications built using Electron API for 3 user roles



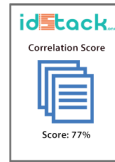
User Role: Validator | Module: Data Validating Module



User Role: Relying Party | Module: Score Calculating Module



IDStack calculates the confidence score for individual documents



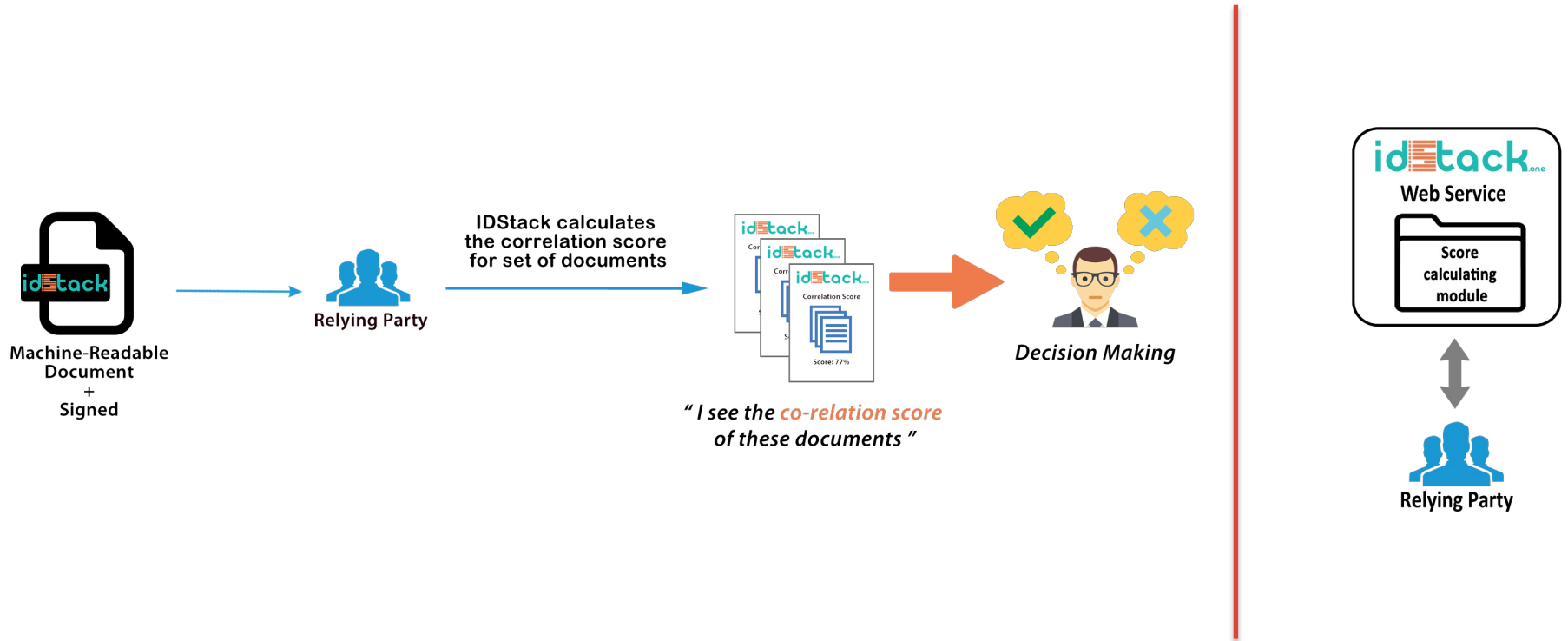
"I see the **confidence score** of the document"



Decision Making



User Role: Relying Party | Module: Score Calculating Module



Need of idstack_{.one}

- On average, **3.1 days** are added to most processes in order to collect **physical signatures**
- eSignatures **reduce document turnaround time by 80%**

Source: AIIM White paper study
<http://www.aiim.org/pdfdocuments/MIWP-DigitalSignatures-2013.pdf>

Need of

The Electronic Transactions Act (ETA) No. 19 of 2006 (Section 7) gives e-signatures the same legal weight as traditional hand-written signatures.

Source: Electronic Transaction Act, No.19 of 2006 (Section 7)

In budgets for the years 2016 and 2017, the government allocated LKR 15 billion to implement policy of digitalizing the economy.

Source: Budget Speech 2017, Ministry of Finance

Need of **idstack**.one



Processing time

12 hours



15 minutes

Error Rates

40%



5%

Source: Accepting E-Documents with E-Signatures, VERITE Research
(February 2017)



Source: The Global Enabling Trade Report 2016,
World Economic Forum

TABLE 1: TIME TAKEN (IN HOURS) TO COMPLY WITH DOCUMENTATION REQUIREMENTS FOR INTERNATIONAL TRADE

	To Export	To Import
Singapore	4	1
UAE	6	37
Malaysia	10	10
Oman	31	24
India	61	67
Pakistan	62	153
Sri Lanka	76	58

Source: World Bank, Doing Business Index 2016

ICTA implement e-Document Attesting System at Ministry of Foreign Affairs



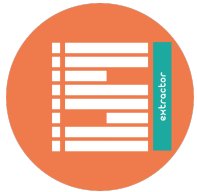
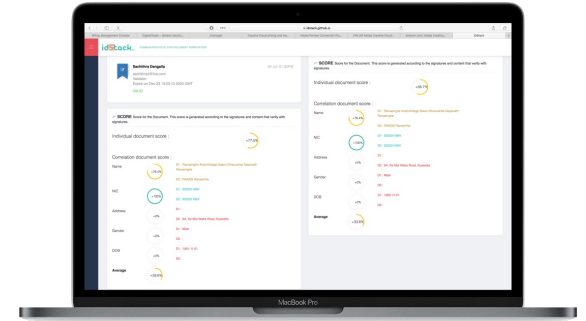
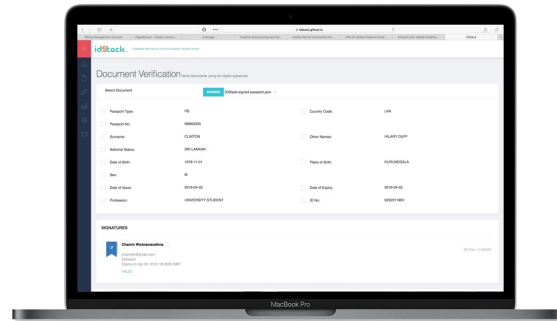
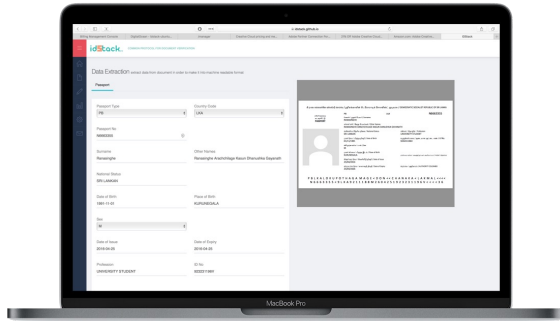
The Information and Communication Technology Agency (ICTA) has taken another step towards creating a digitally-empowered nation by implementing an electronic Document Attesting System (eDAS) at the Ministry of Foreign Affairs (MFA).

Source: <http://www.ft.lk/article/596399/ICTA-implements-e-Document-Attesting-System-at-Ministry-of-Foreign-Affairs>
(February 2017)

idstack.one

The Common Protocol for Document Verification built on Digital Signatures

End User Application



Extractor
App



Validator
App



Relying Party
App

End User Application



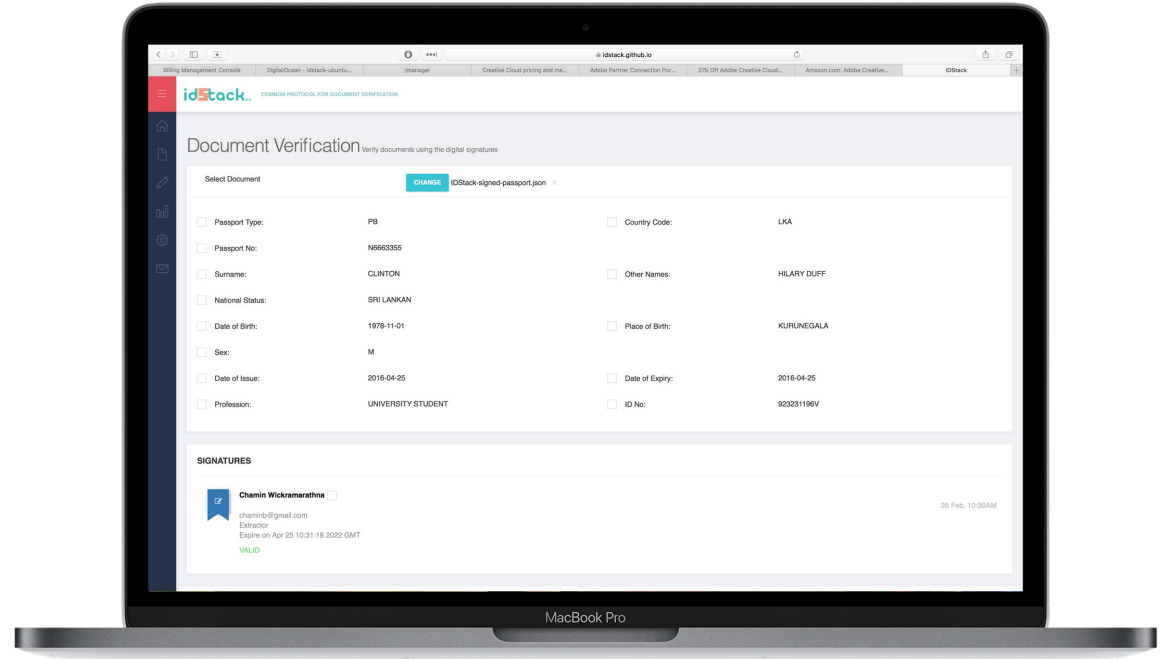
Extractor
App



Validator
App



Relying Party
App



End User Application



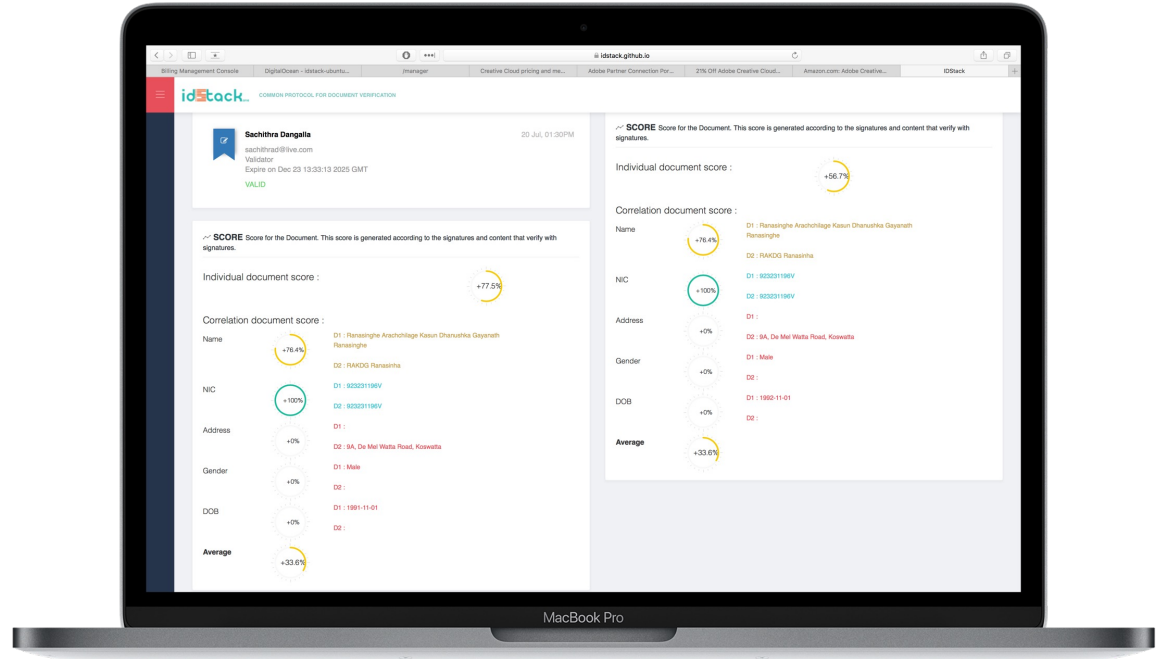
Extractor
App



Validator
App

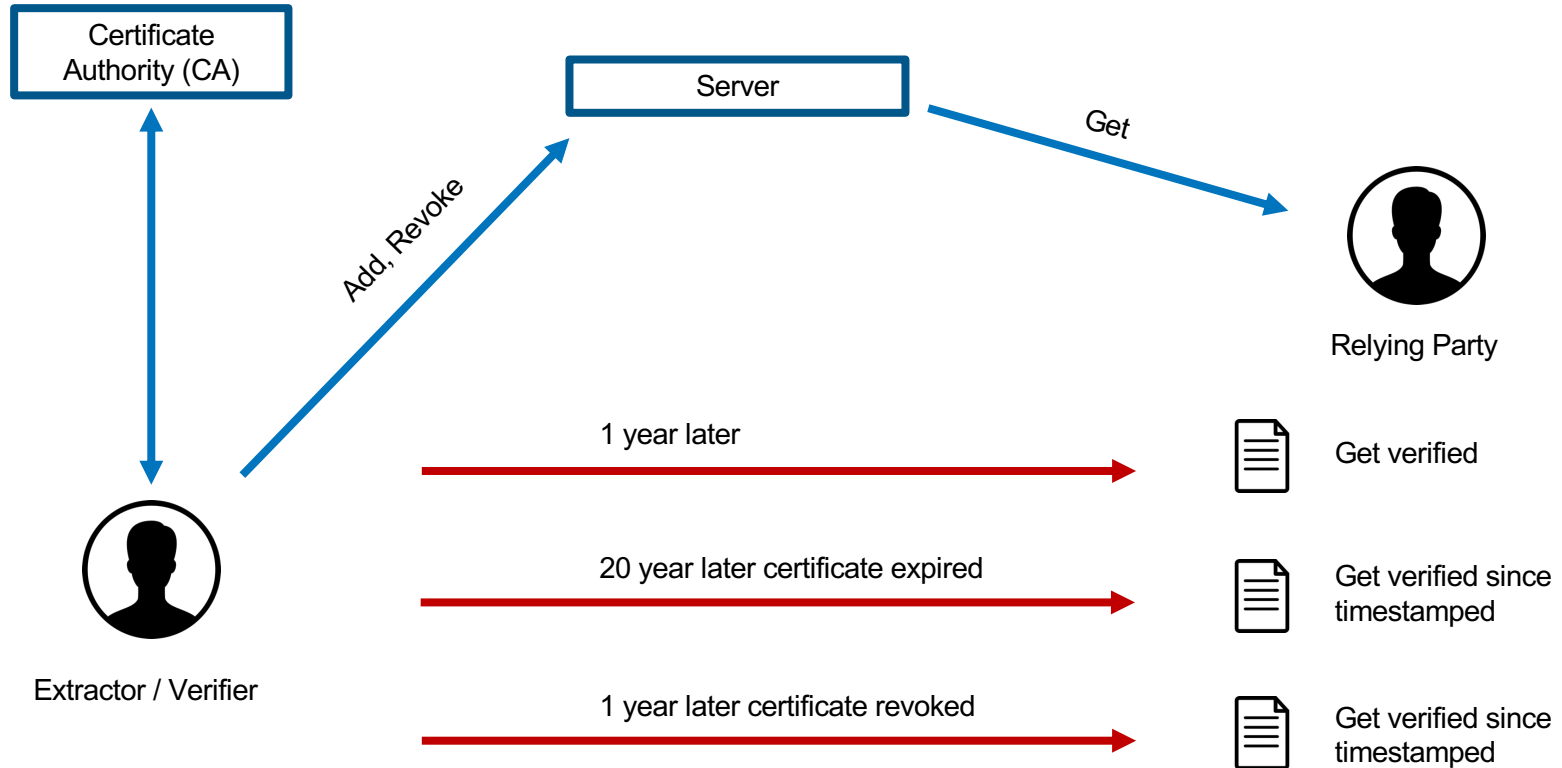


Relying Party
App



Thank You !

Appendix 2



Appendix 3

Correlation Score

Super Attributes:

1. Name
2. Address
3. DOB
4. Gender
5. NIC

Other information (with identical keys) are compared using String similarity measures.

Appendix 3

Correlation Score

Person-Identity based attributes can be divided into 4 categories [1]				
	Space-related attributes	Time-related attributes:	Classifying attributes:	Identification codes:
Example	Address	Date of birth	Gender	Passport No, Social Security No
Possibility to be updated	High	Very Low	Low	Very Low
For FYP scope covering Sri Lankan context we selected 4 attributes with weights	Address 0.4	Date of Birth 0.99	Gender 0.99	NIC 0.99

[1] B. Lisback, Linguistic Identity Matching

Appendix 3

Correlation Score

Problems:

1. Sri Lankan names are long and repetitive: existing name-similarity measuring methodologies can give conflicting results
2. Mostly in Sinhala or Tamil: translation to English can have variations
3. Different documents represent the name with different attributes and attribute order
4. Non-linguistic typing mistakes
5. People can change their names

Solution:

An algorithm that calculates:

- phonetic similarity
- order of name segments
- string similarity

Appendix 3

Correlation Score

	Document1	Document2
Example	Ranasinghe Arachchilage Kasun Dhanushka Gayanath Ranasinghe	Kevin Gayanath Ranasinha
Phonetic representation	R525 A622 K420 D522 G530 R525	K870 G530 R525
Identify overlapping name segments	R525 A622 K420 D522 G530 R525	K870 G530 R525
String similarity	Ranasinghe Arachchilage Kasun Dhanushka Gayanath Ranasinghe	Kevin Gayanath Ranasinha

Oder score between doc1 & doc2 = N-gram similarity
= 2/7
= 0.287

Appendix 3

Correlation Score

Unique phonetic IDs	Document 1	Document 2	String similarity (Normalized Levenshtein)
R525	Ranasinghe	Ranasinha	80%
A622	Arachchilage	-	
K420	Kasun	-	
D522	Dhanushka	-	
G530	Gayanath	Gayanath	100%
K870	-	Kevin	

String similarity score between doc1 & doc2 = $(0.8 + 1) / 6 = 0.3$

Total score between doc1 & doc2 = $(\text{order score} + \text{string score}) / 2$
= $0.287 + 0.3 / 2$
= 0.2935

Appendix 3

Correlation Score

	D1	D2	D3	D4
D1	-	a	b	d
D2	a	-	c	e
D3	b	c	-	f
D4	d	e	f	-

a = Score between D1 & D2

b = Score between D1 & D3

c = Score between D2 & D3 etc.

Correlation score of D1 = $(a+b+d) / 3$

Appendix 4

$$\text{Single Document Score} = \frac{\sum W_a S_a}{\sum W_a}$$

Single document score calculates a weighted average based on the uniqueness weight and determinacy score.

Uniqueness Weights (W_a)		
Super Attributes	NIC	0.99
	DOB	0.99
	Gender	0.99
	Name	0.7
	Address	0.4
Minor Attributes	Others	0.1

Determinacy Scores (S_a)	
Existence of super attribute	0.7
Existence of signed super attribute	0.9
Existence of signed minor attribute	0.2

Appendix 5

API - <http://docs.idstack.apiary.io>

Sample JSON

Student ID - <http://www.jsoneditoronline.org/?id=dd533c47497c94e42e7aabc40e7ab547>

Passport - <http://www.jsoneditoronline.org/?id=f85dbe5a204a0fa2f9cfb50bfaf10de1>

Appendix 6

FYP Scope

Input

- Digital PDF & Paper document without OCR

Language

- English

Template

- Use predefined set of templates
(*Student ID, Transcript, Passport, Driving Licence*)

Future Work

Input

- Physical Documents with OCR

Language

- Sinhala & Tamil

Appendix 7

A reviewed concept

Blockchain Technology - Bitcoin [1], Ethereum [2]

A concept we use

Digital Signature

Similar products

DocuSign [3], Stampery [4], ShoCard [5], BlockSign [6]

[some use cases]

A product we use

DocParser [7]

[1] S. Nakamoto, "Bitcoin A Peer-to-Peer Electronic Cash System", 2017. [Online]. Available: <https://bitcoin.org/bitcoin.pdf>

[2] "Ethereum Project", Ethereum.org, 2017. [Online]. Available: <https://www.ethereum.org/>.

[3] "Don't settle for less. Choose the #1 eSignature brand.", DocuSign, 2017. [Online]. Available: <https://www.docusign.com>

[4] L. I. C. G. Adán Sánchez de Pedro Crespo, "Stampery Blockchain Timestamping Architecture," [Online]. Available:

<https://s3.amazonaws.com/stampery-cdn/docs/Stampery-BTA-v5-whitepaper.pdf>.

[5] "ShoCard | Identity for a Mobile World", Shocard.com, 2017. [Online]. Available: <https://shocard.com/>

[6] "BlockSign", Blocksign.com, 2017. [Online]. Available: <https://blocksign.com/>

[7] "Docparser - Extract Data From PDF to Excel, JSON and Webhooks", Docparser.com, 2017. [Online]. Available: <https://docparser.com>