

Amazon Textract

Amazon Textract is a fully managed machine learning service that analyzes documents and automatically extracts text and structure.



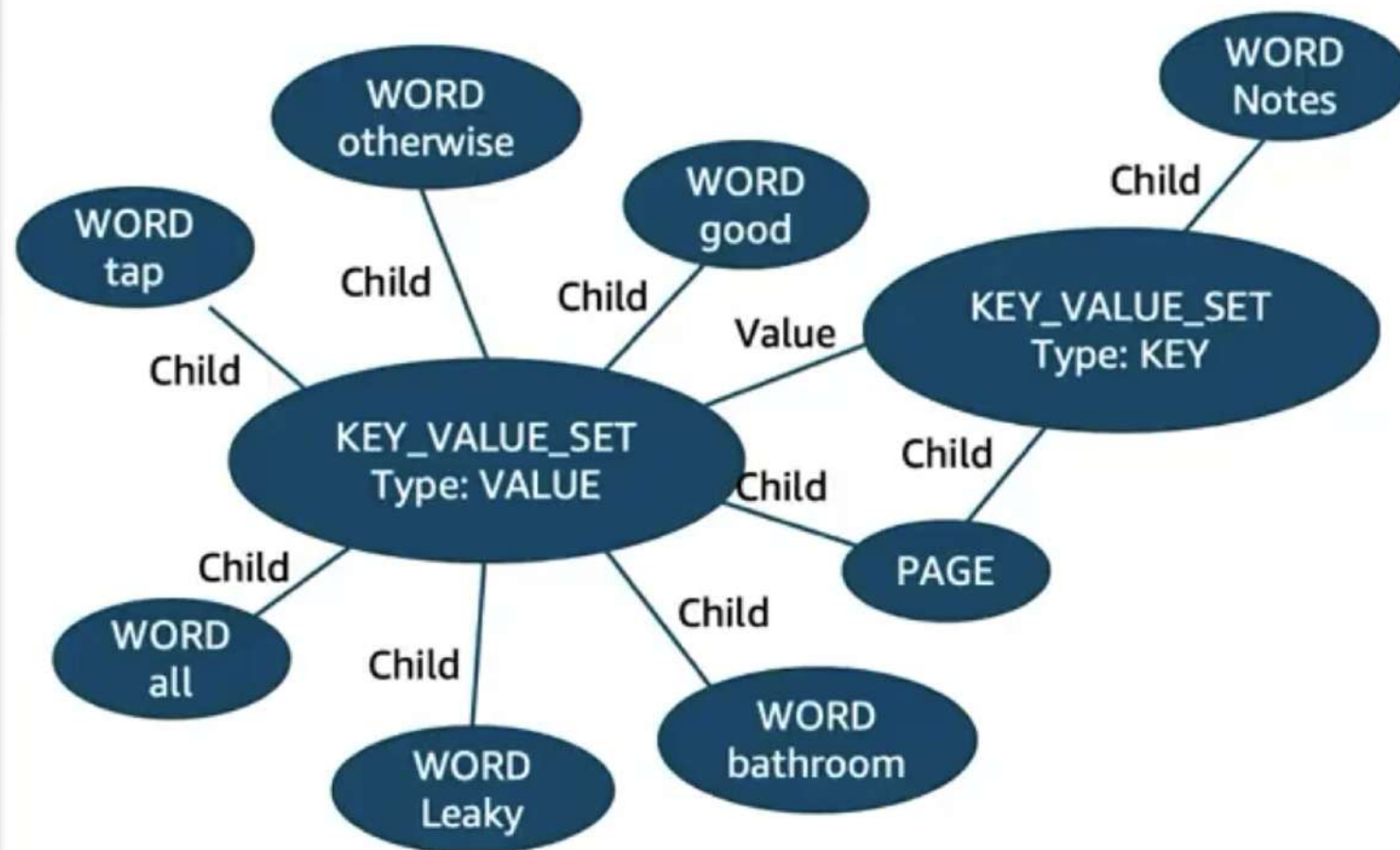
Amazon Textract

Amazon Textract API operations

- AnalyzeDocument / StartDocumentAnalysis / GetDocumentAnalysis
- AnalyzeExpense / StartExpenseAnalysis / GetExpenseAnalysis
- AnalyzeID
- DetectDocumentText / StartDocumentTextDetection / GetDocumentTextDetection

AnalyzeDocument: FeatureTypes=FORMS

House ID	279K
Address	123 Any Street Any Town, USA
Number of Bedrooms	4
Price Estimate	\$ 650,000
Notes	leaky bathroom tap otherwise all good



AnalyzeDocument: FeatureTypes=QUERIES

House ID	279K
Address	123 Any Street Any Town, USA
Number of Bedrooms	4
Price Estimate	\$ 650,000
Notes	leaky bathroom tap. otherwise all good

Request:

"Text": "What are the notes? "

"Alias": "Notes"

Response:

"BlockType": "QUERY_RESULT",

"Confidence": 95.0,

"Text": "Leaky bathroom tap. Otherwise all good.",

AnalyzeExpense



```
"SummaryFields": [  
  {  
    "Type": {  
      "Text": "VENDOR_NAME", "Confidence":  
97.87217712402344  
    },  
    "ValueDetection": {  
      "Text": "WHOLE FOODS\nMARKET",  
      "Geometry": ..,  
      "Confidence": 70.45414733886719  
    },  
  },  
  {  
    "Type": {  
      "Text": "OTHER",  
      "Confidence": 84  
    },  
    "LabelDetection": {  
      "Text": "Net Sales:",
```

© 2022, Amazon Web Services, Inc. or its affiliates. All rights reserved.

Next up, we can see the results



AnalyzeExpense



```
"SummaryFields": [  
  {  
    "Type": {  
      "Text": "VENDOR_NAME", "Confidence":  
97.87217712402344  
    },  
    "ValueDetection": {  
      "Text": "WHOLE FOODS\\nMARKET",  
      "Geometry": ..,  
      "Confidence": 70.45414733886719  
    },  
  },  
  {  
    "Type": {  
      "Text": "OTHER",  
      "Confidence": 84  
    },  
    "LabelDetection": {  
      "Text": "Net Sales:",
```



AnalyzeExpense: Items and values

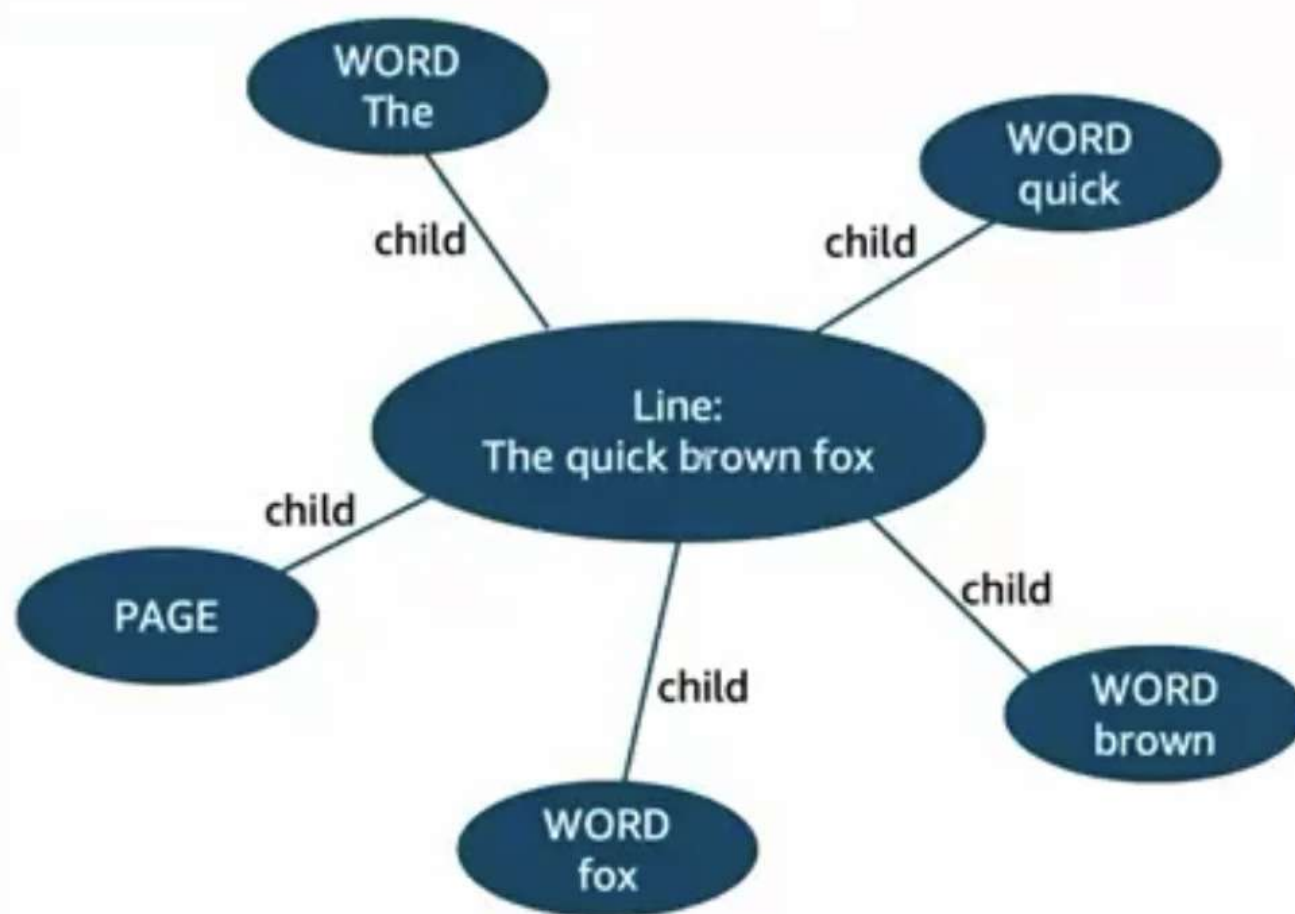


```
"LineItemGroups": [  
  {  
    "LineItemGroupIndex": 1,  
    "LineItems": [  
      {  
        "LineItemExpenseFields": [  
          {  
            "Type": {  
              "Text": "ITEM",  
              "Confidence": 95.61370086669922  
            },  
            "ValueDetection": {  
              "Text": "STMPT CHOC COLD BRW",  
              "Confidence": 95.59813690185547  
            },  
          },  
        ],  
      },  
    ],  
  },  
]
```



GetDocumentTextDetection

The quick brown fox



AnalyzeID



```
"IdentityDocumentFields": [  
  {  
    "Type": { "Text": "FIRST_NAME" },  
    "ValueDetection": { "Text": "RUSSELL", "Confidence":  
99.3445053100586 }  
  },  
  {  
    "Type": { "Text": "LAST_NAME" },  
    "ValueDetection": { "Text": "SAYERS", "Confidence":  
98.8835678100586 }  
  },  
  {  
    "Type": { "Text": "MIDDLE_NAME" },  
    ...  
  },  
  {  
    "Type": { "Text": "SUFFIX" },  
    ...  
  },  
]
```