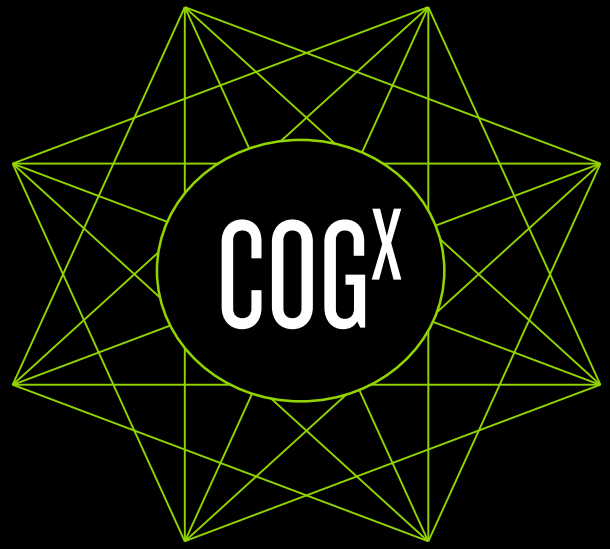
COG^X KICKOFF

Cognitive Xtraction

Wednesday, June 21, 2017



AGENDA

1	OBJECTIVES	• Provide background and discuss the vision of the COG^X Platform
3	PLATFORM ROADMAP & ARCHITECTURE	• Discuss high-level roadmap, and functional and technical architecture
4	OWNERSHIP	• Discuss roles and responsibilities across teams, and meeting cadence
5	NEXT STEPS	• Discuss tactical next steps

COG^{X} OBJECTIVES

OUR MISSION

COG^{X} IS A COGNITIVE PLATFORM THAT TRANSFORMS RAW INFORMATION INTO DIGITAL DATA FORMATS TO SUPPORT ANALYSIS AND WORKFLOW EXECUTION

Designed to replicate human thinking in order to learn and infer more accurate outputs, COG^{X} utilizes cognitive technology to automate and improve the process of ingesting and analyzing semi-structured and unstructured information

WHAT WE ARE GOING AFTER

The market lacks a product that can process a variety of information types (e.g., paper documents), with high accuracy, that can be tailored to solve specific use cases. To satisfy this unmet need, we will develop a cognitive platform that provides core information processing capabilities that can be configured to enable specific solutions

Cognitive Platform

A flexible and modular design to support client customizations and interoperability

- Cognitive Intake & Processing
- Case Management Workflow
- COG^{X} Platform
- Entity Creation & Relationship Tagging
- Metrics & Reporting

Enabling Solutions

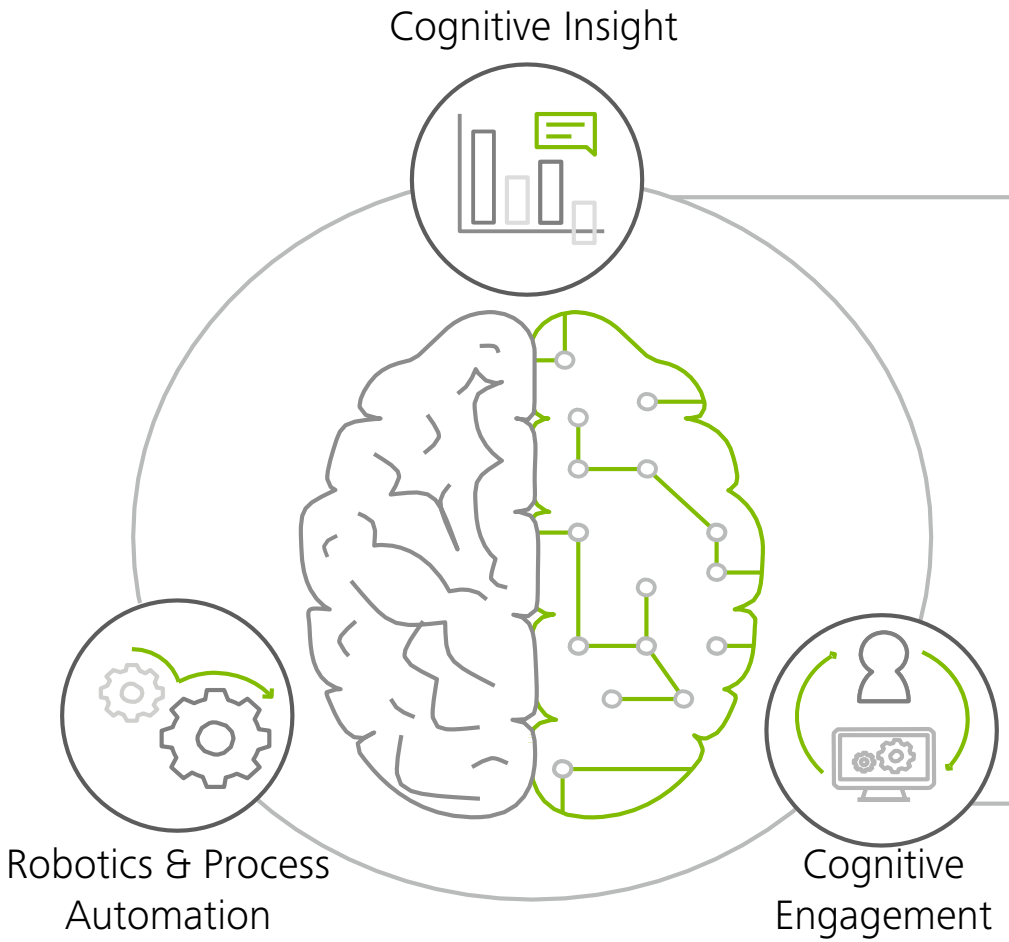
Provider/Claim Intake: IOCR on paper documents to automatically extract key fields required to update a provider's profile and adjudicate claims

Adverse Events: IOCR and NLP analysis on AE reports to classify patient and event information for reporting and investigation

Others: Fraud waste & abuse analysis, as well as financial document processing

BUILDING THE COGNITIVE ADVANTAGE

The COG^X platform is aligned to Deloitte's cognitive advantage portfolio, directed towards building our robotics and process automation capabilities



Robotics and process automation will serve as the primary emphasis of the COG^X platform aimed at accelerating existing client processes and augmenting resources with consolidated information

Cognitive engagement bots that leverage the information extracted by the platform can serve as a potential future extension of the platform

HOW WE WILL ENGAGE OUR CLIENTS

We will drive our client pipeline from initial conversation to operation through an iterative, three step approach

1 VALUE DEMONSTRATION

- Perform a demonstration (1-2 day turnaround) using a client's own data to showcase the cognitive platform's ability to ingest and process their information
- The demonstration will not include any solution/client specific configurations, but catalyze a conversation for long-term potential

2 RAPID CONFIGURATION

- Over 2-4 weeks the COG^{x} team will incorporate client specific ontologies, document templates, data sets, and other use case requirements
- The configured platform will be ready for client use and oriented for ongoing learning / improvement

3 IMPLEMENTATION & SUPPORT

- The system will be fully deployed into the client's ecosystem (cloud or on premise) and licenses will be established
- Support for the product will be continuously provided by Deloitte to the client per engagement contract

COG^{x} VALUE PROPOSITION

Faster (Speed)

- Adoption of the platform will automate and expedite existing client processes

Better (Accuracy/Satisfaction)

- Cognitive capabilities can provide outputs with comparable precision to humans

Cheaper (Cost)

- Automation of manual tasks will significantly reduce client operational costs

DEPLOYMENT OPTIONS

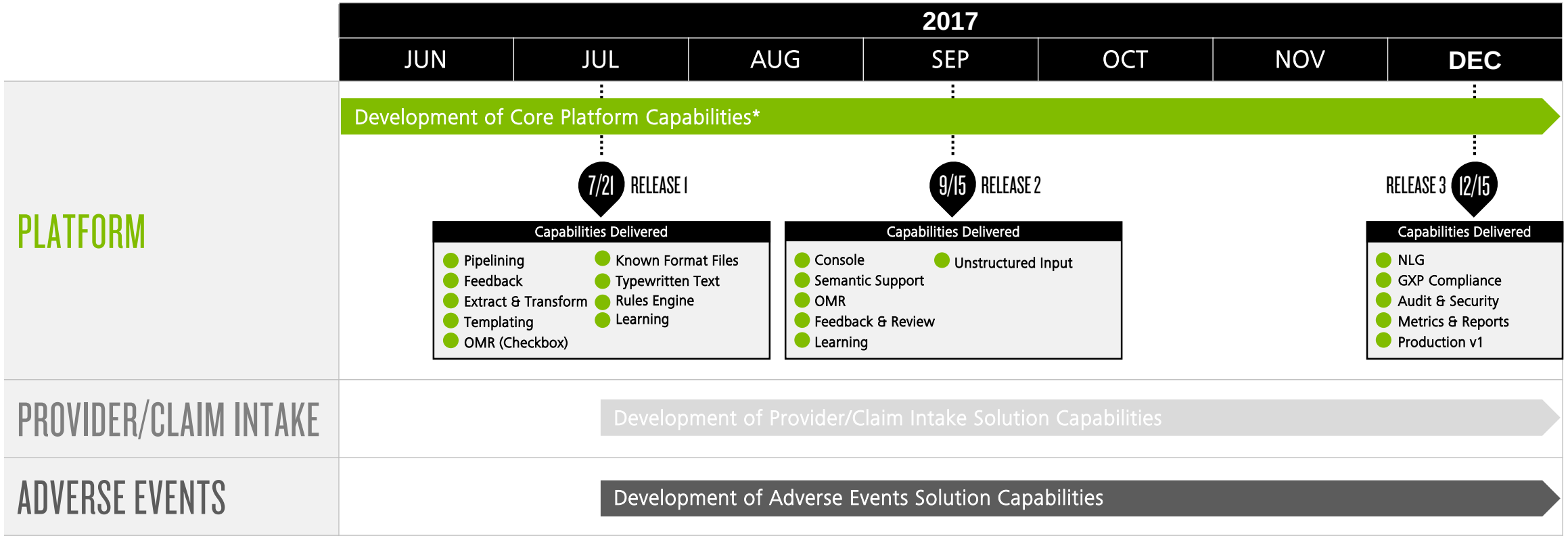
Deloitte will have a variety of options to implement and operate the COG^{X} platform for clients

	ENGAGEMENT ORIENTATION	INFRASTRUCTURE MODEL	INITIAL & ONGOING FEES
	**OOTB Configured Product** Deloitte configures the product based on the client specific solution requirements; however, the client takes on the responsibility of building out additional customizations (e.g., screens, integrations)	**On-premise** Deloitte works with the client to deploy the software **SaaS** Deloitte hosts the solution	• Flat product fee for configuration & deployment • Yearly licensing fee • Infra support fee (if SaaS) • Potential VBB arrangement
	**Customized Product** In addition to configuring the product, Deloitte will build customizations (e.g., screens, integrations) as needed for the client	**On-premise** Deloitte works with the client to deploy the software **SaaS** Deloitte hosts the solution	• Flat product fee for configuration & deployment • T&M implementation fee • Yearly licensing fee • Infra support fee (if SaaS) • Potential VBB arrangement

COG^{X} ROADMAP & ARCHITECTURE

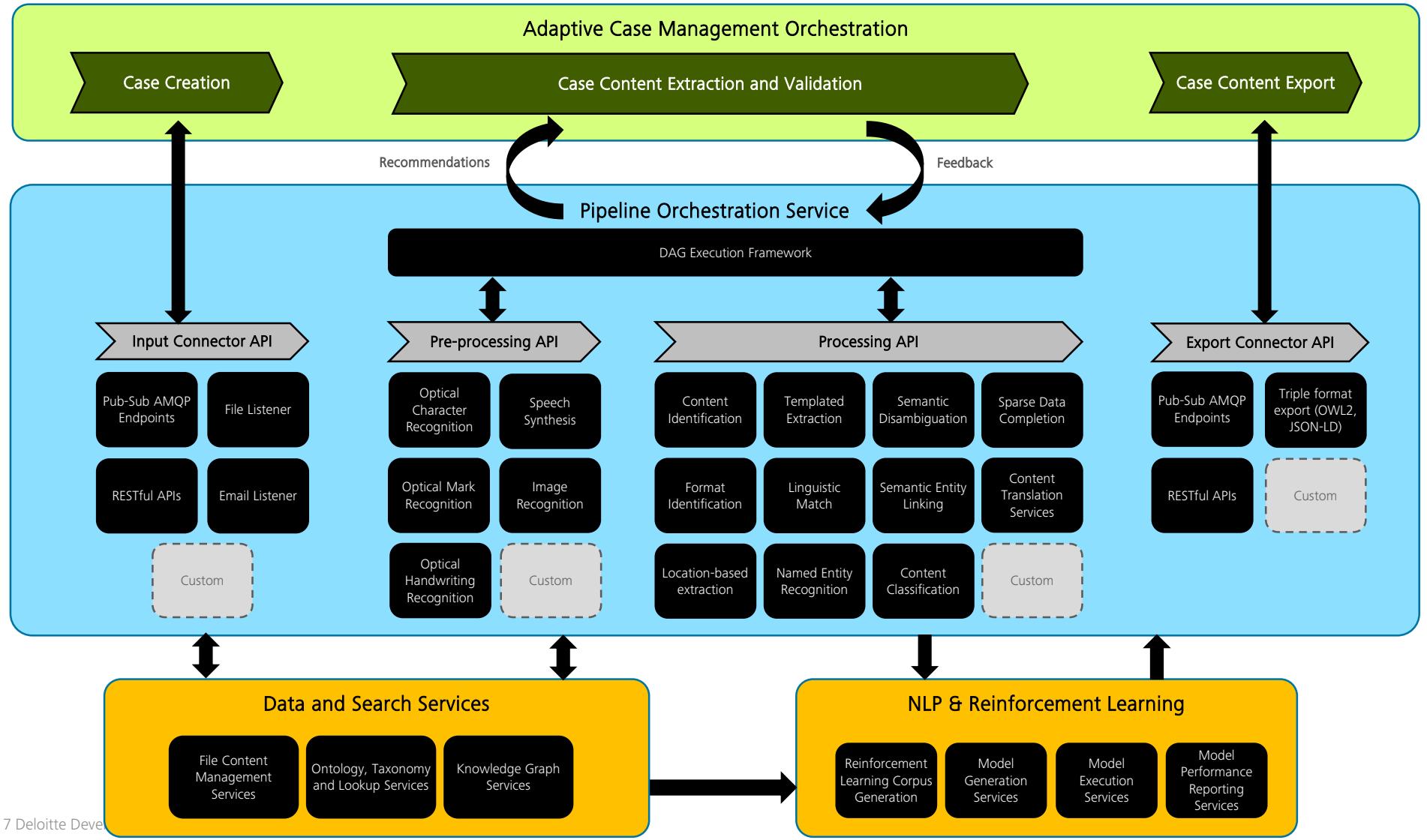
NEAR TERM ROADMAP

Initially, while the platform is being developed, solutions will be developed in parallel in order to be able to meet client commitments while minimizing duplication where possible. Once the platform has attained significant maturity, new clients will use solutions configured and implemented on the platform while existing clients may be migrated

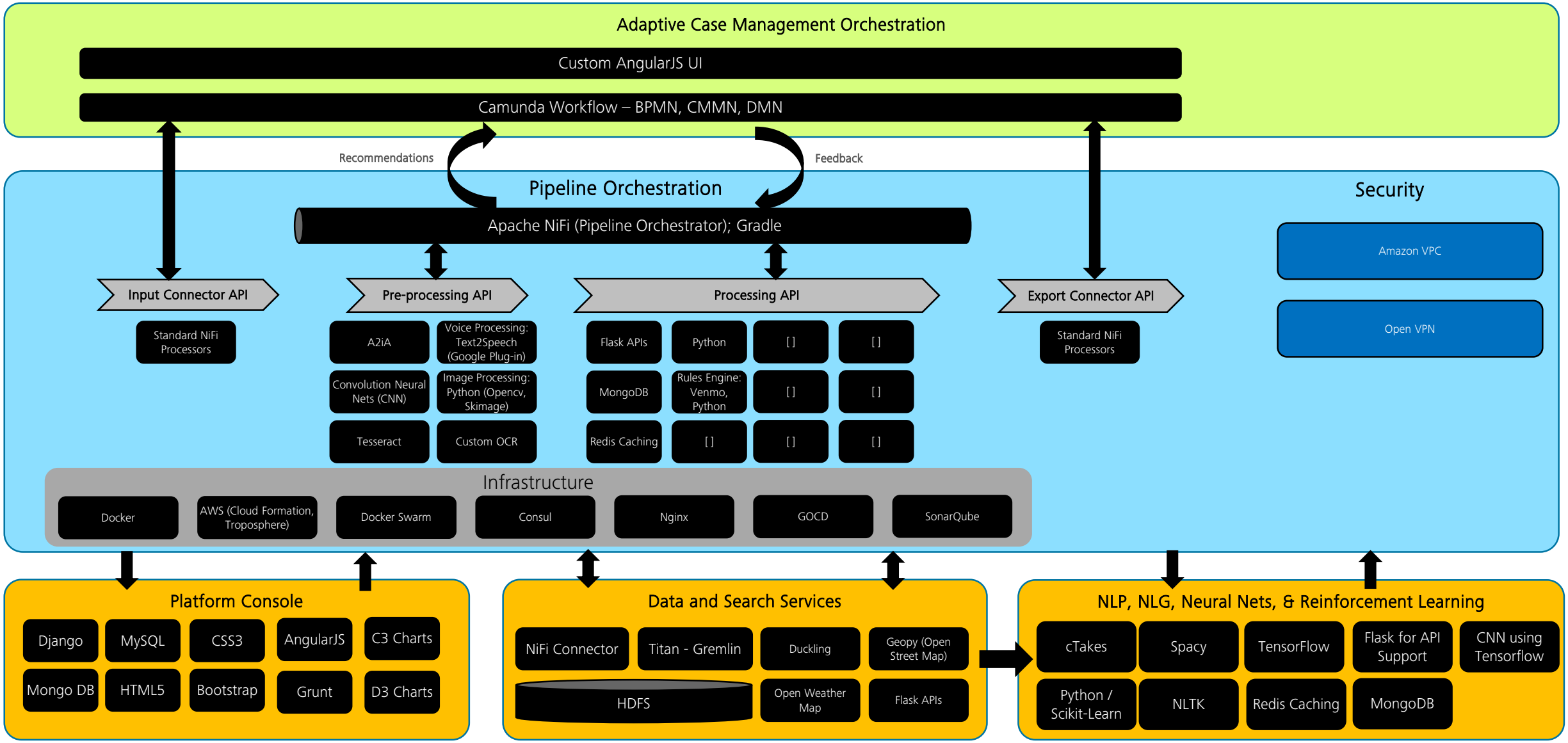


Following each major platform release, capabilities developed for the platform will be integrated into each applicable solution

FUNCTIONAL ARCHITECTURE



WIP TECHNICAL ARCHITECTURE



COG^{X} OWNERSHIP

DEVELOPMENT RESPONSIBILITIES

High level development responsibilities are as follows

DELOITTE US

Requirements

Translate markets needs into business requirements for platform and solutions

Sales & Marketing

Lead client opportunity identification and go to market sales strategy

Learning Framework Design

Design reinforcement learning techniques

DELOITTE USI

Delivery Management

Own the delivery execution across development groups

OCR Integration & Semi-Structured Extraction

Develop OCR integration and extraction capabilities for semi-structured information sources

XPMS

NLP

Develop ontologies and NLP models

Pipeline Processing

Develop processing pipeline

Capability Integration

Integrate capabilities developed across groups into the platform

INNOMINDS

User Interface

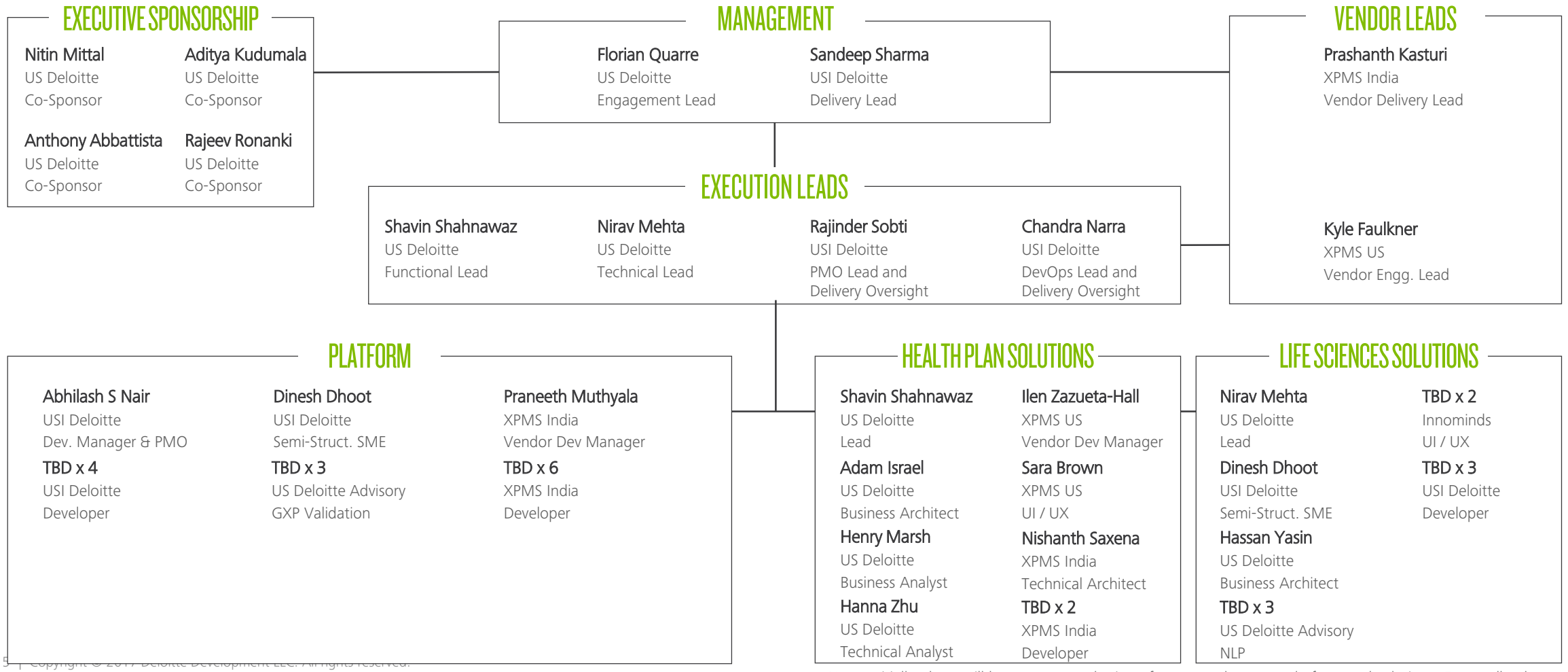
Develop front-end screens

Workflow and Case Management

Develop back-end workflow and case management

TEAM ORG CHART

Coordinated under a centralized leadership team, platform resources will develop the core COG^X capabilities while solution teams will develop use case specific capabilities enabled by the platform



Note: Initially, there will be some cross-sharing of resources between platform and solutions. Eventually, there will be a greater shift of resources to platform than solution

MEETING CADENCE

The following recurring meetings with be conducted to ensure alignment across teams throughout the development of the platform

NAME	FREQUENCY	ATTENDEES	FOCUS
Workthread Scrum	• Daily	• Individual platform/solution workthreads	• Discuss – as individual threads – work done since last scrum, work to be done before next scrum, and roadblocks
Scrum of Scrums	• Every Tuesday/Thursday	• Key representatives from each workthread	• Discuss – as a consolidated group – work done since last scrum, work to be done before next scrum, and roadblocks
Execution Lead Meeting	• Every Tuesday morning	• Execution leads • Management leads	• Align as a core group • Follow-up on key outstanding action items • Discussion questions regarding plan, budget, resources, risks, infra, and ops
Sponsor Update	• Every 2 weeks	• Executive sponsors • Management leads • Execution leads	• Update to leadership on progress to date • Key decisions if any needed
Demo, Retrospective, and Sprint Planning	• Every 2nd Friday	• Workthreads • Management leads • Execution leads	• Demo of sprint working software • Review of what to start doing, stop doing, and continue doing • Review stories planned for next sprint
Release Planning	• Start of each release	• Workthreads • Management leads • Execution leads	• Review of latest roadmap • Review of stories / functionality completed in previous release • Review of stories / functionality for upcoming release

COG^X NEAR TERM ACTIVITIES

NEXT STEPS

1. Complete USI and XPMS cross-team knowledge sharing and capability harvesting. Build out rationalized capability map, tech stack, and user stories for execution against roadmap
2. Formalize PMO activities such as budget tracking, meetings rationalization, team lists, PTO calendars, etc.
3. Consolidate process, tools, and environments for development among platform teams with focus on GXP compliance readiness
4. Initiate IOQ process for IP protection and go-to-market for platform-based industry solutions

APPENDIX

PLATFORM CAPABILITIES (1 OF 2)

The following core capabilities will be developed for the platform:

CAPABILITY	DESCRIPTION
Pipeline Processing	• Back-end processing to support data ingestion, conversion, processing, and analysis
Workflow Setup	• Configurable sequence of activities performed to translate text
User Interface	• Display digitized form and reporting and metrics measuring OCR accuracy
Template Definition	• Ability to define and configure mapping between standard input and output templates
Entity Definition	• Ability to create entity graphs to define content of specific form fields to provide context for text translation
Source Data Definition	• Ability to specify input data type
Document Ingestion	• Ability to ingest document and convert document to image
Source Location Definition	• Ability to define source locations from which to ingest input files
Image Pre-Processing	• Prepare document for text translation
Optical Character Recognition (OCR)	• Text translation conducted by OCR engine
Machine Learning	• Augments accuracy of OCR engine output and incorporate feedback to learn over time
Text Disambiguation	• Employ methods (e.g. disambiguation, regex) to improve accuracy of text translation
Rules Definition	• Capability to define business rules to support text disambiguation

PLATFORM CAPABILITIES (2 OF 2)

The following core capabilities will be developed for the platform:

CAPABILITY	DESCRIPTION
Corpus / Entity Relationships	• Provide context for text translation
Metric Reporting	• Ability to generate and display accuracy metrics (i.e. confidence scores)
Natural Language Processing	• Supports natural language analysis and querying of paper documents processed
Data Storage	• Stores source data and generated outputs
Security	• Protects PHI/PII housed in data files
Audit and Logging	• Logs system activities
Case Management	• Work queue of cases can be assigned to users for review and input

CLAIM INTAKE SOLUTION CAPABILITIES

The following core capabilities will be required for initial release of the claim intake solution

CAPABILITY	DESCRIPTION
User Interface	• Display the document form and resulting field level confidence scores
Template Definition	• CMS 1500 form's data elements are defined and configured for system processing
Entity Definition	• Entities can be defined to enable solution to recognize content of various fields on claim forms
Document Ingestion	• Ingest the claim form and convert to an image
Image Pre-Processing	• Ability to recognize individual segments of claim and crop accordingly to prepare document for processing
Optical Character Recognition	• Recognize and translate text from paper document into digital form
Learn From Feedback	• Users are able to provide feedback, and the solution is able to consume this feedback to improve its accuracy
Text Disambiguation	• Employ methods (e.g. disambiguation, regex) to improve accuracy of text translation
Rules Definition	• Applicable business rules can be defined to support text translation
Metric Reporting	• Tracks and reports confidence scores and accuracy metrics
Data Storage	• Stores source data and generated outputs
Security	• Protect PHI/PII found on the claim forms
Auditing/Logging	• Log text processing steps and all other system activities so that these can be reviewed by users
Case Management	• Create and manage work queues

PROVIDER SOLUTION CAPABILITIES

The following core capabilities will be required for initial release of the provider intake solution

CAPABILITY	DESCRIPTION
User Interface	• Display the document form and resulting field level confidence scores
Template Definition	• Data elements can be defined for contract, correspondence, and delegate forms and configured for system processing. System can recognize the particular type of form being used.
Entity Definition	• Entities can be defined to enable solution to recognize content of various fields on forms
Image Pre-Processing	• Ability to recognize individual segments of form and crop accordingly to prepare document for processing
Optical Character Recognition	• Recognize and translate text from paper document into digital form
Learn From Feedback	• Users are able to provide feedback, and the solution is able to consume this feedback to improve its accuracy
Text Disambiguation	• Employ methods (e.g. disambiguation, regex) to improve accuracy of text translation
Natural Language Processing	• Supports natural language analysis and content extraction of unstructured text data
Rules Definition	• Business rules for text translation, validating input on form, and making database updates can be configured
Form Content Validation	• Able to leverage business rules defined to validate that all content required for updates is complete on the form
Update Databases	• Perform updates on provider information within databases after processing form content
Metric Reporting	• Tracks and reports confidence scores and accuracy metrics
Data Storage	• Stores source data and generated outputs
Security	• Protect PHI/PII found on the claim forms
Auditing/Logging	• Log text processing steps and all other system activities so that these can be reviewed by users
Case Management	• Create and manage work queues for different business areas and document types

ADVERSE EVENTS SOLUTION CAPABILITIES

The following core capabilities will be required for initial release of the adverse events solution

CAPABILITY	DESCRIPTION
Source Data Definition	• Ability to specify input data type of Adverse Event Report (E2B R2, R3, HL7, Unstructured, or Scanned Documents)
Data Storage	• Stores source data in a database agnostic repository and generated output in ICSR standard formats
Natural Language Processing	• Supports natural language analysis and content extraction of unstructured text data
Security	• Protects PHI/PII housed in data files
Audit and Logging	• Logs both system and user activities
Case Management	• Work queue of incoming Adverse Event cases that can be assigned to users for processing and validation
Optical Character Recognition (OCR)	• Text extraction of semi-structured and un-structured scanned documents conducted by OCR engine
Reinforcement Learning	• Augments OCR engine output by incorporating user feedback
Rules Definition	• Capability to define business rules to support text disambiguation, validation, and amplified medical review
Pipeline Processing	• Back-end processing to support data ingestion, conversion, processing, and analysis
Workflow Setup	• Configurable sequence of activities performed to ensure accuracy and consistency in case processing and validation
User Interface	• Screens used to guide OCR extraction and a customized digital version of the MedWatch form used in the workflow
Template Definition	• Ability to define and configure mapping between standard input source types and output digital forms
Entity Definition	• Ability to customize the data object used in the extraction and processing & validation workflows

ARCHITECTURAL CONSIDERATIONS

Architectural design and technology selection will consider the following

TOPIC	CONSIDERATIONS
SaaS vs. On-Premise	• Product should be cloud-ready but able to be deployed on premise with post production manageability
Fuzzy Matching & Feedback	• Using fuzzy matching could make it difficult to construct a feedback loop
Ontology Definition	• When defining the object type, we need to fundamentally define the ontology (e.g., for adverse impact vs. claim)
Work Queue On/Off	• If a solution requires only automation, we should be able to turn the work queue off or set accuracy thresholds to initiate cases
Field-Level Pipelining	• Should individual fields be cropped and processed or is this approach too granular?
Global/Local Knowledge Graph	• Should we use local knowledge graphs for each solution to build a global knowledge graph?
APIs	• Must be API enabled to allow us to be enterprise ready out of the box • All API orientated orchestration with each capability exposed as a processor
Data Schema	• Stored data must be schema based
Pipeline Processing	• Use Nifi or "Nameko processing' for pipelining with an interface that takes in information at each stage
NLP Model Type	• Need to agree on the model types (Stanford NLP, Tensorflow, etc.) and possibly plug in different models to generate different results
Templates	• Need to consider template versioning
Entity Definition	• Model the claim (ER diagram) and save it as an ontology; Support ability to extract definitions into OWL
Workflow	• Workflow is via 'Camunda'
DevOps Choices	• Jenkins vs GoCD
OCR Vendor	• Determine which OCR vendor should be used for platform (A2IA, Tesseract) • Discuss idea of swapping out A2IA for a cheaper vendor depending on the solution

