openIDL - Known Issues

This page covers the known issues with openIDL.

#	Component	Issue		MVP MoSCoW (Scope may not include all future thinking)	Priority	Complexity
		Application cluster deployment should be done from gitops repo	The process of deploying a new node requires images to already exist. The images are created as part of the development process which happens in the openidl-main repostiory. The images that are published are available publicly. Any deployment of these images should happen from the openidl-gitops repo actions so that a carrier can get setup without requiring access to the main repo.		Medium	Medium
		Cost should be optimized for the technology	The use of kubernetes should be optimized for the type of underlying compute usage. We may want to consider serverless instead of ec2 underpinnings.\n\nThere are probably many other places where costs can be limited.		Medium /Low	Medium
x		No console for managing the hyperledger fabric network	There are two available options that can be considered - hyperledger explorer and hyperledger operations console. Must have at least visibility of what is in the blockchain.	Must	High /Medium	High
x		ChunkId and BatchId are antiquated and unnecessary.	The chunkid and batchid were notions introduced by IBM. They are not needed anymore, me thinks. (see data loading hash below)	Must	Low	Low
	Reference Implementation	Need a bootstrap script for adding users, data calls, extraction patterns and data to hds	Provide a script that can create enough elements to get started with the system immediately.\n\n- Add users (if necessary)\n- Add data into the HDS\n- Add data calls\n- Add extraction patterns		Low	Medium/Low
		Configure file should drive the pipelines			Medium /Low	High
		Automate initial account setup			Medium /Low	Medium
		Responsive UI	The User Interface is not responsive.		Low	Medium
		SSO / Identity Management	Consider using a universal identity management solution like that discussed by Chainyard.	Could	Medium /Low	High
x		Cognito Alternative	Something other than cognito or ibm appid	Must	?	Medium
x		Monitoring is missing	The IBM system did not implement monitoring. The current scope does not include monitoring.\n\nHere is an article on how to provide monitoring in Kubernetes using Prometheus: https://phoenixnap.com/kb/prometheus-kubernetes-monitoring\n\nThere is monitoring implemented in the reference implementation using AWS native services.	Must	High	Medium
x		Maintenance Strategy	The system is a distributed network. The nodes reside in foreign clouds that AAIS does not own. In order to keep the nodes up to date, they must be managed. GitOps is a practice that enables this. AAIS will establish these practices for ongoing maintenance of the distributed nodes.	Must	High	High/Medium
a		Data Architecture is not fully defined	The data architecture is not fully informed by the problem space. Having distributed nodes, that include quality assurance of the data and the extraction for analytics means some of the current architecture must be reimagined.	Must	TBD	TBD
a		Messaging Standard	The format for sharing data with the node is required. This may be best served by a messaging model that is light weight, well documented and performant. This is where the bulk data processing occurs in real time or through events.	Must	TBD	TBD

		HDS Format Standard	Once data has been ingested via the messaging format, it will be validated and cleansed to a high standard valid for use in extraction and reporting. This is not a "transaction" format, but a logical format that fits the needs of the extraction itself. For Data Calls a policy oriented model is more appropriate.	""	TBD	TBD
		Data Lifecycle / Time to Live / Auditing	Data has become a major part of the cost of cloud computing. Keeping data around forever, especially when it is derived from other data, is likely not the best choice. The lifetime of the data must be considered and optimized for the use case.	"	TBD	TBD
		Extraction Processing Implementation	The extraction pattern model currently uses a map reduce function in MongoDB. This locks us into MongoDB and uses a closed environment without access to the outside world we'll need for correlating other data like census. The extraction capability must be reimagined.	Wont	TBD	TBD
		Data Loading UI	The user interface for loading data and the ETL provided by IBM works only with stat plan data. This functionality is shifting to the member for implementation. AAIS can provide a reference implementation or none at all.	Could	Low	Medium
x		Data Loading Scalability	The current API inherited from IBM does not scale for large data sets. This component must be reimagined in a way that can handle very large volumes of data.	Must	High	High
x		Data Loading Hash	Since the data in the pipeline is derived from other data, it is likely to be transient. We need to to track that data has been provided, but if the data architecture changes, then we must rethink where we take snapshots and record them in the ledger.	Must (see above)	High	High/Medium
		Data Load Quality Assurance (Rules)	The rules that validate the submitted data are currently not part of the architecture. That which IBM provided were applied to the stat records, not the "HDS" format. Most submission of data will not follow the stat plans. The HDS is where we know the data will be the same and where the rules should be applied. This must be built into the node.	Wont	Medium	Medium
x		Consent Processing	The consensus process does not currently work correctly. It picks up the data from the harmonized data store upon consent, instead of when the data call expires. This must be fixed in alignment with any other data or application architecture changes resulting from previous items.	Must	High	High
		Multi Tenant Consent Processing	Allow individual carriers on the multi-tenant node ability to consent for their data	Could		Medium
		Stat Plan Mapping	The stat plan is the current form of data submission. Not all data will come in this form. Hopefully, it can be retired at some point. Until that time, any reporting via stat plan through the openIDL requires the stat plan mapping to exist. The IBM implementation is inadequate and must be replaced.		Medium	Medium
x		Incomplete Analytics Node	The analytics node, as inherited from IBM, is not fully functional even for the data call. The remaining functionality must be developed.	Must	High	High
x		Inadequate Unit Tests	IBM left us with minimal automated testing, including a minimum of unit tests and no CI/CD that automatically executes them.	Must	High /Medium	Medium
x		Scalability /Performance Tests	There are no performance/scalability tests.	Must	High	High/Medium
		Automated UI Tests	There are no automated UI tests.	Should	High /Medium	Medium/Low
		Automated API Tests	There are no automated API tests.	Should	High /Medium	Medium/Low
x		Penetration Testing	Enable penetration testing	Must		Medium
a		QA Plans	Plans for testing all levels of the application for deliveries	Must		
		Authentication	Not using the best practice of handing off to authentication provider. Should we do this or keep in control to make it more plug and play with different providers?	Could	Low	High
		Enable kubernetes dashboard	Add the dashboard to the setup of the cluster in IaC			

x		Bugs	There are several bugs in the current code that must be fixed. • UI Count of Data Calls • Adequate Filtering of Data Calls / Workflow / In-box concept • UI icons require code change • Validation of expired tokens in insurance data manager? - was validating token, but not expiration • The UI is not responsive. It does not scale to any resolution under 1080 nor to other devices. • use library charts to remove dups of helm chart templates • put utilities into kubernetes - for creating users • move icons out of code • stat-agent should not be able to like • block explorer • view organisations or carriers is not working (from ibm) • remove dependency on chunkId from data loading	Must	High	High/Medium
			Templating all the config files – Currently configuration files must be created manually whenever a new carrier needs to get on-boarded.		Medium	Medium
			Automate the process pushing the config files to vault.		Medium	Medium/High
			Automate the process of creating git actions secrets.		Medium	Medium/High
			Add a database authentication for reference implementation (this will remove the dependency with Cognito)		High	Medium
			Vault High Availability		Medium	Medium
			CA Password is hardcoded to orgname-pw in BAF open source implementation	Must	High	Medium
			Volume mount size is hardcoded to 50 GB in BAF open source implementation	Must	Medium	Medium/High
			Updating application helm charts with RBAC rules and service account creation	??	Medium /High	Medium
		Blockchain Hosting	Can we use AWS Templates for deploying HLF?			