# California Oilfield Underground Aquifer Injection Monitoring by Blockchain Technology

Jonathan Crawford
Department of Computer Science
*California State University Bakersfield*
Bakersfield, USA
jcrawford4@csub.edu

Andrew Folsom
Department of Computer Science
*California State University Bakersfield*
Bakersfield, USA
afolsom1@csub.edu

Vananh Vo
Department of Computer Science
*California State University Bakersfield*
Bakersfield, USA
vvo3@csub.edu

Angela Dana Tante
Department of Computer Science
*California State University Bakersfield*
Bakersfield, USA
atante@csub.edu

John P. Yu
Department of Geological Sciences
*California State University Bakersfield*
Bakersfield, USA
jyu2@csub.edu

Chengwei Lei
Department of Computer Science
*California State University Bakersfield*
Bakersfield, USA
clei@csub.edu

***Abstract*—Water contamination is a primary concern in a region where water and petroleum play such vital roles in the economy, and where both industry and regulatory agencies pay close attention to environmental quality. In this research, we built a Distributed Ledger Technology (DLT) based prototype using R3 Corda. Its purpose applies in the oil & gas underground injection control (UIC) operations for the underground aquifer protection. This DLT prototype is a permissioned network which allows oil & gas companies to create, disseminate and trace immutable records. This network allows oil & gas companies, state agencies, and all other participants to share secure records such as well information while maintaining data integrity, traceability, and security. With the underlying cryptographic technologies of DLT, any unauthorized changes to the information, ownership or its history will become infeasible.**

***Index Terms*—Blockchain, Oilfield injection, aquifer protection**

## I. Introduction

More than 60 years ago, California's oil and gas companies injected gas and water into hydrocarbon bearing formations for production enhancement and waste disposal. So far there are about 55,000 injection wells in California, mainly for secondary and tertiary recovery of oil and gas reservoirs [1]. At the present time, California produces 15 barrels of wastewater for every barrel of crude oil produced, and a total of 2/3 of the wastewater (60 million barrels per day) needs to be re-injected into the original reservoirs underground. About 300,000 barrels of wastewater are treated and mixed with other available water for agricultural irrigation and industrial use. About 2.7 million barrels of wastewater are injected into the 1,800 injection wells designated by the US Environmental Protection Agency (EPA). The EPA will designate some more exempted areas for wastewater injection to ensure safety of aquifers. California is short supply of fresh drinking water and there is a growing demand for underground water for human consumption. Therefore, authorities require a higher standard in underground water management. The EPA established injection exempt areas where wastewater injection can be injected as the water quality is not suitable for human consumption. The produced wastewater injection fluid is not allowed to be injected outside of the exempt area. The exempt area designation was completed in 1972 by the EPA [2]. In the last 50 years, there were many additional wells drilled and oil, gas, and water production depleted reservoir pressure. Most oil reservoirs need additional pressure enhancement to produce, such as gas or water injections. There is a need to re-designate or expand the exemption areas for further injection. Once the newly designated exempt area is approved, the gas and water will be allowed to conduct injection operations.

In 1977, US Clean Water Act (CWA) was approved by the US congress. The CWA is the primary federal law in the United States governing water pollution. Its objective is to restore and maintain the chemical, physical, and biological integrity of the nation's waters; recognizing the responsibilities of the states in addressing pollution and aiding states to do so. The underground injection control process is a part of the CWA which is regulated by the EPA. The EPA designated all the oil produced and injected fluids as Type II fluids and assigned the legislation right to California Geology & Energy Management (CalGEM, formerly known as Division of Oil, Gas & Geothermal Resources) by Primacy Agreement in 1974. During the production process of an oil field, there will be invasion of natural gas, underground water and the changes that occur during the process of secondary and tertiary oil recovery. Agricultural fields and oilfields are in the same area and agriculture water wells and oil wells are also in the same neighborhood. There is a possibility that oilfield fluids could contaminate water wells. The EPA established the UIC review and execution process to ensure the injection process complies of the CWA's laws and regulations. Therefore, the tracking of oilfield operation and contamination history data

Thanks to California Energy Research Center at California State University Bakersfield for the support to this research

are very important. In this case, future uncertainties of water contamination can be minimized.

The current methods used in aquifer exemption and UIC review impede their timely approval. Due to the number of organizations involved, the review processes can be very lengthy and generally delay the effective execution of the UIC projects. Each entity must receive information on the proposed well, record it into their individual data centers, and then evaluate the viability of the project. This distribution of data is not only redundant but lacks the transparency desired in distributed data storage. Furthermore, as the data required for approval changes hands, the chances of data tampering, either due to personal bias or human error, increases. This can lead to irregularities in the information each entity has, especially as new or updated information is slowly disseminated. In order to reduce these delays and facilitate continued drilling, a new method for submitting and evaluating these projects is required. This new method must ensure data integrity, uniformity, accessibility, and accelerate the review process.

## II. Methodology

### A. Injection Data Security and Integrity

Oilfield injection data security is a priority for any organization, and data integrity is especially crucial for those that use information from a variety of sources to make critical decisions. However, reliable, transparent, and secure data sources and data traceability remain problematic for oil & gas industries and regulatory agencies [3]. For example, Kern County, California has a population of 900,000 and has been ranked as one of the nation's most productive in both agriculture (#1 County in the nation [4]) and petroleum generation (#2 County in the nation [5]). Water contamination is a primary concern in a region where water and petroleum play such vital roles in the economy on a national scale, and where both industry and regulatory agencies pay close attention to environmental quality [6]. Monitoring and forecasting based on available data are imperative to mitigate complications [7]. The EPA and State-government agencies (CalGEM, State & Regional Water Boards) oversee both water and oil industries. These agencies have collected industries' daily operating data for over 100 years [8] and continue to add to these data sets, providing ideal databases for data analyses and security. However, data integrity is a primary issue of concern for those that monitor and analyze environmental data.

### B. BlockChain Technology

Blockchain is a method of implementing a secure, distributed record of information. The early concept of blockchain was originally conceived by Stuart Haber and W. Scott Scornetta in the 1990's when they attempted to develop a way of time-stamping documents that would prevent people from altering that data later [9]. It wouldn't be until 2008, when Satoshi Nakamoto introduced Bitcoin, that blockchain would reach notoriety.

The simplest description of a blockchain's structure is a linked list of blocks, distributed among peers, with each peer maintaining their own copy. Each block contains a set of records or transactions, a unique hash, and the hash of the block before it. To add a new block to the chain, peers on the network attempt to solve a computationally complex problem in a process called "proof of work." The first to solve the problem has their solution validated by the remaining peers in the network. If the solution is confirmed, each peer updates their copy of the chain. Otherwise, they do not update their chain and await the proper solution. The resource intensive nature of "proof of work" and validation process make the blockchain nearly impossible to tamper with. In order to successfully modify the exiting chain, a malicious user would require greater computational power than the entirety of the peer network.

For example, Bitcoin was envisioned as "[a] purely peer-to-peer version of electronic cash would allow online payments to be sent directly from one party to another without going through a financial institution." [10] Rather than depend on banks to validate transactions, members of the peer-to-peer blockchain network are able to validate the transactions themselves. This would allow digital purchases to overcome the need for dedicated middlemen and the cost of transaction fees. Peers validating the transactions can be trusted because of the blockchain's structure and how it validates blocks that are added to the chain. While financial applications may have been the initial focus of blockchain development, many industries requiring reliable ledgers are attempting to integrate it into their operations. The immutable, distributed nature of the blockchain make it an excellent solution in situations accuracy and consistency of shared records has been difficult to maintain.

## III. Implementation

### A. Data Sources

In this research, we use the public domain data from the State regulatory agencies to construct a blockchain database. Their data are obtained from different entitles such as oil & gas companies, agriculture companies and service companies. The companies or agencies have different departments as the nodes and they enter & retrieve data for their own needs. A distributed system with nodes spreading over multiple entities and the sources will serve the needs in the blockchain system. Therefore, this system will ensure the data ownership for all the entities as well as access control/security of the entities' propriety data. We also propose to change the current manual review and operational workflows into a 'smart contract' in the blockchain applications. In this case, the new review process and operational workflow can avoid dual tasks among State agencies. The secure and traceable data will then be publicly available with the objective of improving the efficiency of regulatory approval and operating processes to prompt economic development in the local community.

### B. Platform

Corda [11] is an open source blockchain platform. Developed by R3, it is designed for use in enterprise environments, where strong emphasis is placed on data privacy, security, and compliance. Corda shares data on a need-to-know basis. Corda does this by modelling records of data in a specific direction and sending dependencies of a record only when the record is relevant to that party. This approach makes Corda amenable to enterprise use cases, where parties may be competitors who want to keep business relationships and details secret from one another. Therefore, Corda uses a permissioned blockchain model. Participants must first obtain a digital certificate from the doorman before they can join the network as illustrated in the following flowchart. Once the oil & gas companies have the digital certificate, they can run a Corda node on their server and communicate with the Corda network. The node's IP address and digital certificate will be broadcast to all parties through the network map. The network map is a server that publishes the list of IP addresses and digital certificates inside the Corda network.

As a permissioned blockchain, Corda can avoid the large amount of computation resources required by a proof of work system. Rather, it uses notarization, in which every information is notarized. Notarization is where special nodes called notary nodes create digital signatures as proof that the transaction is valid, and it does not double-entry in any record. When a Corda node sees a record of data with this signature, they can be sure that the record is valid. The notary nodes collaborate among themselves using a consensus algorithm so that they make the same decision of whether to sign a particular record of data.

### C. Design

Figure 1 shows our overall architecture for the prototype model. The "Corda Infrastructure Shared Services" Corda node includes the notary nodes are being used for notarization of transactions, and the doorman & CA service which manages the membership of the permissioned blockchain and provides the network map. Then, each party joining the UIC prototype network operates a State Agencies Corda Site. The State Agencies Corda Node is the central component, connecting with the rest of the blockchain network. The Corda Node uses Relational Database Services to store data including records and transactions. The API Gateway provides the API (application programming interface) for interacting with the Corda Node, such as reading data and making changes on the blockchain. Finally, the Web Services provide the user interface for the end user. For example, State Agencies staff will login to the Web Services in their Corda Site. In between sites, a Corda Node can communicate with other Corda Nodes, the notary services, and the doorman & CA service. Through this communication, they can perform tasks such as signing, sending, and notarizing records, such as approval well history and wellbore diagrams etc.

There are two key reasons to design the system like this, privacy and extensibility. In terms of privacy, Corda protects the privacy between the parties by sharing data on a need-to-know basis. Oil and gas companies need not worry about leaking trade secrets while using the prototype model. Furthermore, oil and gas companies can run their own Corda node, by running the smart contracts under their control. In terms of extensibility, Corda allows installing CorDapps (Corda Distributed Applications), which are packages of smart contracts. The CorDapps can utilize various components provided by Corda, such as Messaging, Vault and Scheduler. These built-in components make it easy to implement and extend the functionalities of the prototype model. We will explain these modules in detail in the following paragraphs.

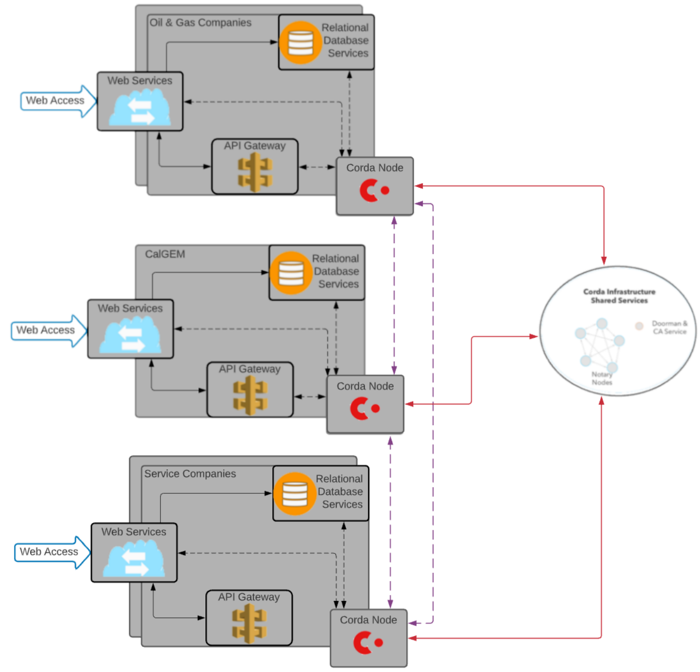


Fig. 1. UIC Prototype network

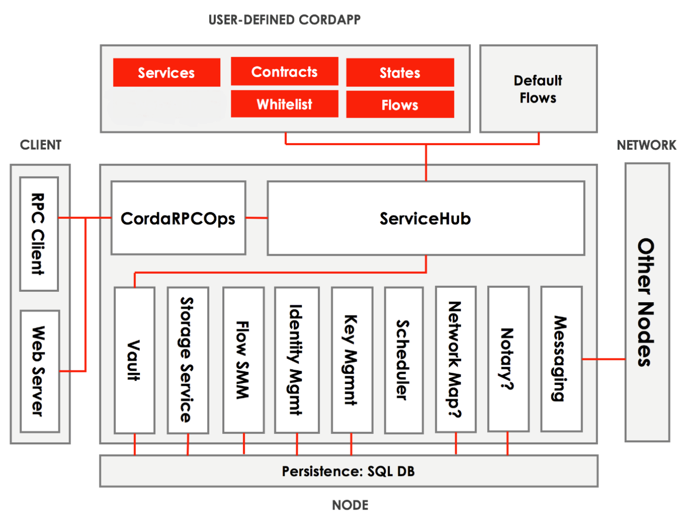


Fig. 2. Corda node architectural functionalities in the prototype model

### D. Structure

The Figure 2, from Corda Documentation 2020 [12], illustrates the architecture of a Corda node. The following explanations highlight the individual component as follows.

- The RPC Server and RPC Client allow other service companies and public users to connect to the Corda node. In the prototype model, the API Gateway connects to the Corda node using the RPC Client. Using this, the API Gateway can connect public users' actions on the user interface to access well information on the blockchain.
- The Messaging Service provides direct, peer-to-peer and secure communication beween Corda nodes. Within the network, all the oil & gas companies, CalGEM and services companies can communicate directly and securely.
- The Notary Service keeps track of whether each record is validated. This is used for ensuring the integrity and correctness of the entire network.
- The Network Map Service manages the membership of the network. When a new company, for example, an oil and gas company, is added to the network or an existing company is removed, the service will automatically update its records.
- The Key Management service manages the identities of the Corda node. This includes one or more private keys stored at the Vault service.
- The Identity Management service identifies other oil & gas companies by using the digital certificates downloaded from the network map. Validated records stored in the Vault may contain fields that refer to the companies; for example, an oil and gas well record may contain a field referring to its owner.
- The Storage Service manages the location where the permits, well information and testing results are stored. Corda supports common relational databases. CalGEM is using WellSTAR which is a relational database. Most of oil& gas companies can use their existing databases to manage all data storage in one place.
- The Flows are executed using the Flow SMM (state machine manager). It allows two Corda nodes to communicate and collaborate with each other to perform a regulatory process; for example, the oil & gas company apply to CalGEM for UIC permit. This can involve, for example, the exchange of information only held by oil & gas company initially, requesting CalGEM to sign the approved permit for the project, and even talking to service company to obtain testing results. In the prototype model, we use flows to create the oil and gas well records, and to create, approve, reject the UIC application. Each of the tasks involves creating a validated record, asking for Water Boards' signatures, and notarizing the approval. Flows can also help validate business level transaction logic. For example, if a service company were to send CalGEM something actually meant for an oil & gas compnay, CalGEM can check who the flow was meant for and respond apporiately.
- The User-Defined CorDapp is where compliance requirements can be applied with a "Smart Contract". A CorDapp defines records, testing information, logging, etc. A smart contract works with the program code and the Corda node to implement compliance requirements. Notably, this includes records that contain data to describe a certain wellbore diagram, well history, well location and more. These records are shared between the Corda nodes on a need-to-know basis. Further, constraints are captured in the smart contracts such as existence of well permits, compliant requirements for each data field, how the data fields can be changed in a workover job, and under what conditions the new record can be created. These defined in the CorDapp. Finally, the creation and modification of these well workover jobs or tests are defined in a well history, which specifies steps for CalGEM to communicate, review compliance, verifies history & test results and approves them.

### E. Graphical User Interface

The currently proposed Graphical User Interface (GUI) of the prototype model consists of two separate interfaces depending on the end user node. One GUI is intended for Oil and Gas Companies that will create instances of wells and projects for submittal to the DLT. The other is for the Regulatory Agency to review those submittals. Since Corda is a DLT permissioned database, a sign-in procedure is necessary to be verified and gain access to the network. This allows for the user to be verified by Corda's Identity Management service, allows for verification of records by party, and ensures data privacy by only sharing records with affected parties.

*1) Well Input Illustrations:* As shown in Figure 3, the main activities in the dashboard of the O&G/Service Company GUI are creating a well and submitting a well. Separating well creation from submittal allows O&G and service companies the option to populate data from separate internal company processes and review that data before submitting them to the regulatory agency by recording the data to the DLT.

Creation of an initial well instance in the well operator GUI requires entering a well name, well type, lease name and well location both in Public Land Survey System of Section, Township, and Range and NAD 83 coordinates. Once an instance of the well is created in the GUI they can be found under the dashboard and filtered by any of the entry attributes in addition to wells that have been created but unsubmitted, wells written to the DLT and pending approval from the regulatory agency, and wells reviewed and approved by the agency. Wells can be updated with supporting documents such as Notice of Intent, Area of Review, and wellbore diagrams can be attached to a well instance in the GUI.

*2) Regulatory Agency Input Illustrations:* In this research work, Edison Field steam flood/cycle case is selected for demonstration. The UIC project is a continuously steam flood/cyclic for enhanced oil recovery from Upper and basal Chanac formations at the Hathaway Cohn/Reddy Leases (approximately 308 acres) in Sec. 28, T29S & R29E. This project

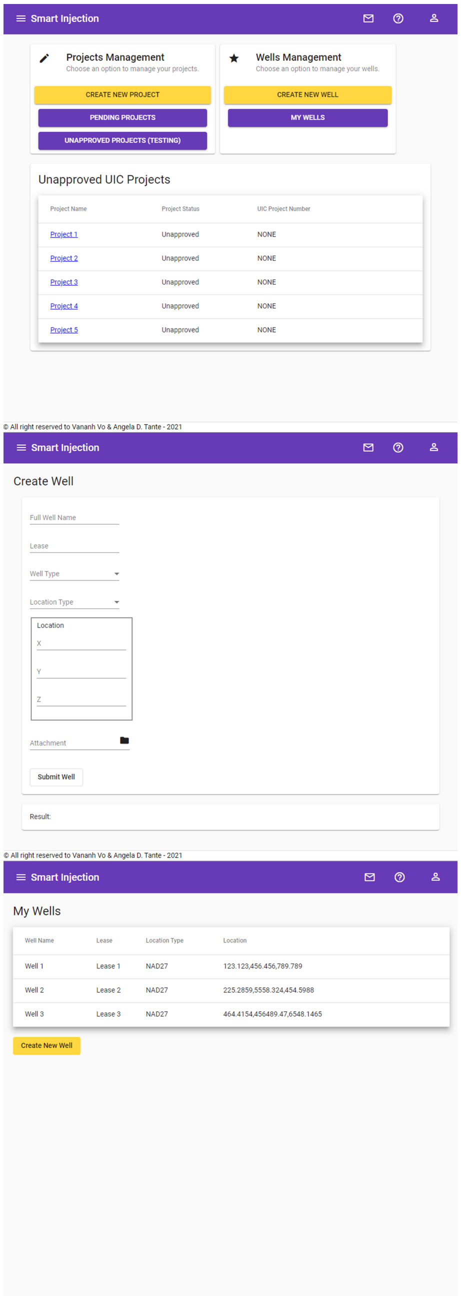


Fig. 3. Well Operator Interface

is in the Edison Groves Area of Edison Field. About 306 producing wells and 117 steam injectors will be drilled at a pattern spacing ranging from 2.3-7 acres. Nine horizontal producers have been drilled in the Basal Chanac sand. The area of review (AOR) for the project is located within the approved aquifer exempted boundary in the Chanac Formation. The UIC project was approved as UIC Project # 22203017 [13].

Figure 4 is the main dashboard of the regulatory GUI, which features a table of wells in the DLT that can be searched and filtered by multiple attributes such as API number, lease, well name, owner/operator, and approval status. Regulators have the

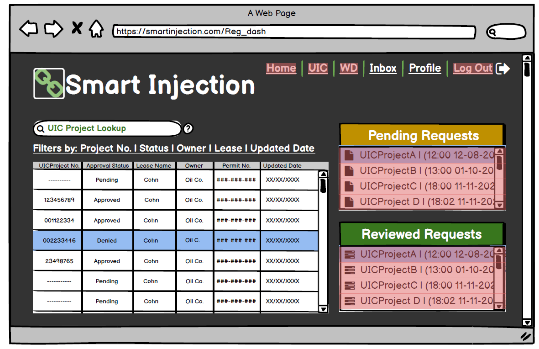


Fig. 4. Regulator Dashboard GUI

option of searching for individual wells for review or searching by UIC project. When searching by UIC project, a list of UIC requests pending review are listed separate from requests that have already been reviewed.

As shown in Figure 5, opening a UIC project in the GUI presents the reviewer with information such as Owner/Operator, lease and/or project name, the number of wells in the project, and links to other data included with the projects such as permits, area of review, and aquifer exemptions. A link to each well in the DLT can be clicked in the GUI to allow the regulator to review the wells in the submitted package. Once reviewed, the regulator can approve the UIC which will update the DLT .

For additional demos and downloads, please visit our project website ***https://cs.csubak.edu/˜atante/senior_expo/***.

## Conclusion and Future Work

In this research work, we have successfully built a prototype Corda model for oilfield thermal recovery process. The model proves the improvements in well data submission and accessing processes. The DLT provides a unified platform for storing and submitting records, allowing individual companies to have their own databases in house and they can submit and access well data at their fingertips. There is no central authority needed to control the data availability. Regulatory Agencies, oil & gas companies and service companies operate their own databases at the equal footing. The best of all, there is no downtime problem, and any incident withholds information from public.

The prototype Corda model demonstrates the data flow, immutability, and security. The cryptographic nature of the Corda ledger ensures data integrity and prevents unauthorized tampering. As each entity maintains it's own vault of shared states, Corda allows for a unified, transparent way for participants in the approval process to store and access proposals while maintaining privacy from outside interference. However, the model does not automate the review processes, such as wellbore integrity and area of review (AOR) analyses, which would greatly increase the speed of the approval process. Wellbore integrity and AOR analyses will be proposed in the future research studies, such as using data analytics and artificial intelligence etc. The current prototype model still

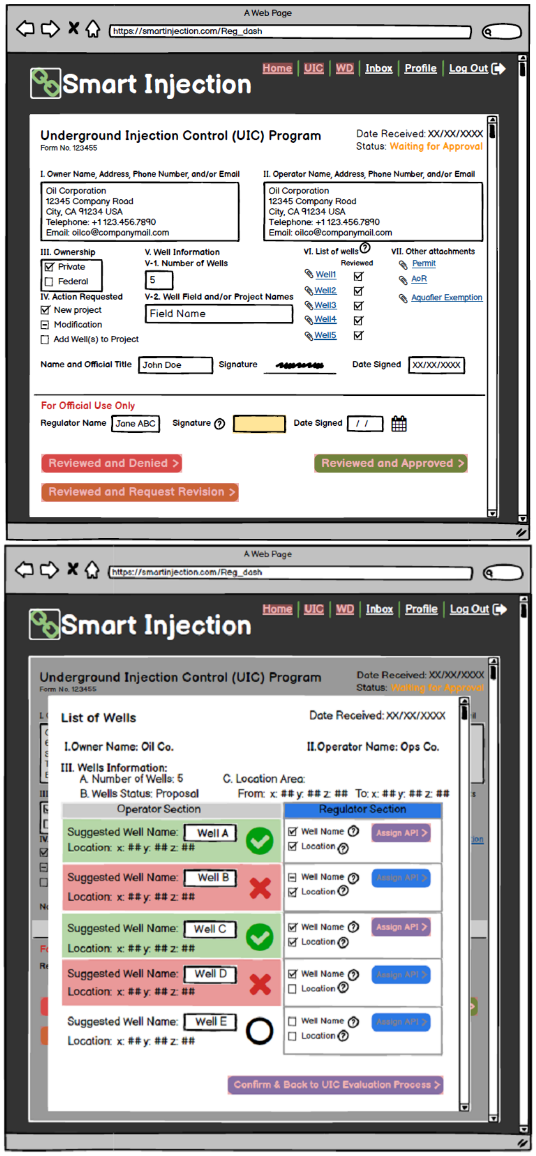


Fig. 5. Regulator UIC Review Interface

conducts a manually review of wellbore integrity and AOR. Then, the results are input into the prototype model to approve the UIC project. There are Corda built-in functionalities which enhance the UIC project's review and operational periods, such as Smart contract, Vault, Identity Management, Scheduler and Notary Services etc. Many of the functionalities automate the data processing within the DLT databases.

## References


[1] WSPA Report on Injection Wells in California, January 2016

[2] California Oil and Gas Fields, Volume 1 – Central California, DOGGR, Department of Conservation, 1998

[3] Cotton, Todd, Karen M. Gardner, Robert Konitzer, and Mary Jane Vansant. "Enterprise framework and applications supporting meta-data and data traceability requirements." U.S. Patent 7,016,919 issued March 21, 2006.

[4] Department of Agriculture and Measurement Standards, "2016 Kern County Agricultural Crop Report", Bakersfield, CA.

[5] Kern Economic Development Corp 2018 Report "Kern County Market Overview & Member Directory".

[6] Zebarth, B. J., J. W. Paul, and R. Van Kleeck. "The effect of nitrogen management in agricultural production on water and air quality: evaluation on a regional scale." Agriculture, ecosystems & environment 72, no. 1 (1999): 35-52.

[7] Clark, James S., Steven R. Carpenter, Mary Barber, Scott Collins, Andy Dobson, Jonathan A. Foley, David M. Lodge et al. "Ecological forecasts: an emerging imperative." science 293, no. 5530 (2001): 657-660.

[8] Hesson, P. E., and H. Bruce. "California Department of Conservation's Division of Oil, Gas, And Geothermal Resources: Orphan Well Program." In SPE Western Regional & AAPG Pacific Section Meeting 2013 Joint Technical Conference. Society of Petroleum Engineers, 2013.

[9] Haber S. , Scornetta W. S. , "How To Time-Stamp a Digital Document", Journal of Cryptography, 3, No 2, 99-111, 1991..

[10] Nakamoto S., Bitcoin: A Peer-to-Peer Electronic Cash System, Bitcoin Web Site, https://bitcoin.org/bitcoin.pdf, Updated July 3, 2018, Accessed September 17, 2019.

[11] Hearn, Mike and Brown, Richard G. "Corda: A distributed ledger." 2019, https://www.r3.com/reports/corda-technical-whitepaper/.

[12] "Corda Documentation", https://docs.corda.net/docs/corda-os/4.6/key-concepts-node.html.

[13] CalGEM, "UIC Project Approval Letter for Cyclic Steam Project #22203012 merged with Steamflood Project #22203017 as new Project #22203017), April 5, 2019.