Birkbeck College, University of London

Department of Computer Science and Information Systems

Proposal submitted in partial fulfilment of the requirement for the MSc in Data Analytics (Advance Computing Technologies)

"Assembling data: An application for understanding and analysing the block chain beyond cryptocurrency"

Author: Daniel Michael Ishmael Hassan

13/04/15

Number of words: 3123 excluding contents, figures, bibliography and appendices.

CONTENTS

TABLES

ABSTRACT

Block chain technologies have revolutionised virtual currency and cryptographic distributed data storage. At this time, block chain technology has been predominantly associated with Bitcoin, however, its uses beyond the economic realm are slowly being recognised; these uses have not yet been consolidated and are still largely uncharted. To address this gap in knowledge, this thesis asks: *to what extent has the block chain been used for more-than-monetary purposes, and how has this changed over time*. In order to answer this question, it will develop a web-application capable of analysing the Bitcoin block chain using *Scala* and *libbitcoin*, searching text data in the block chain in order to aid investigations of 'more-than-monetary' uses. In doing so, it will additionally demonstrate the application of Agile software development practices in the delivery of a well tested and stable application.

DEFINITION OF TERMS

Behaviour Driven Development: Is an evolution in software development process that aims to place into the lexicon of the project (tests, method names, classes etc) so that the intentions are clearly communicated. In particular this communication is geared toward valuable business use-cases - instead of focusing on the 'how' (Agile Alliance, n.d.)

Bitcoin: Bitcoin is described as the world's first decentralised cryptocurrency (Brito, 2013). In the words of the Bitcoin Foundation **"**Bitcoin is a consensus network that enables a new payment system and a completely digital money. It is the first decentralized peer-to-peer payment network that is powered by its users with no central authority or middlemen. From a user perspective, Bitcoin is pretty much like cash for the Internet" (bitcoin.org, n.d.)

Block chain: A shared, open and public documentation tool which contains all confirmed Bitcoin transactions, and which is thus essential to the Bitcoin network. It is secured, and its integrity is maintained, via cryptography (bitcoin.org, n.d.)

Cryptocurrency: A medium of value exchange which uses cryptography as the securing mechanism for transactions, and also for control in the creation of new units (Greenberg, 2011).

Consensus network: In relation to Bitcoin, this refers to the process by which the peer-to-peer network of computers which run the Bitcoin software stays in sync. Cryptographic protocols ensure that the distributed and replicated data-structure (the block chain) is agreed upon by all nodes in the network which maintains the integration of the transaction history of the network (bitcoin.org, n.d.)

Decentralised: This refers to the lack of central bank or authority by whom Bitcoin is developed and governed. It also refers to the entirely digital nature of Bitcoin, which has no physical manifestation as such (Sagona-Stophel, n.d.: p. 1)

libbitcoin: An extensive Bitcoin toolkit library written in *C++* which has an extendable, scalable and configurable architecture, making it much easier to interact with the block chain (libbitcoin, n.d.).

‘more-than-monetary’: For the purposes of this thesis, this refers to any use that is not immediately fiscal in nature. Examples of this may be a voting system (Scott, 2014) or an immutable storage of decisions made collaboratively (d-cent, n.d.).

A. INTRODUCTION

A.1 Topic

In 2009 block chain database technology was first released as part of Bitcoin. The block chain, a distributed database containing the entire history of all transactions on the network, was detailed as Bitcoin's "core innovation" (Nakamoto, 2008: p. 1). An open-source online payment system invented by Satoshi Nakamoto, Bitcoin was described as the first decentralised digital currency. As such, it quickly developed into a key technology widely adopted across both public and private sectors, with a top end market value cap of 11.8 billion USD (Coindesk, n.d.).

Until recently the block chain formed the basis for predominantly finance-based tech-initiatives, including cryptocurrencies. However, the block chain's properties of being *decentralised, distributed, replicated, publicly verifiable* and *an immutable store of data* means that its potential uses extend far beyond finance based initiatives. A growing number of companies and publicly funded organisations are starting to take advantage of these possibilities. One of these is d-cent, an EU funded project creating the means for groups of citizens to store records on the block chain, aiding in large-scale collaboration and decision-making processes. For such groups, the block chain offers a way to store information in a resilient, verifiable and secure way, to aid transparency and auditability.

Given the limited knowledge of the block chain's wider possibilities, and the background of these possibilities, this thesis asks: *to what extent has the block chain been used for more-than-monetary purposes, and how has this changed over time*. The extent to which the block chain is being implemented for 'more-than-monetary' uses is significantly understudied. Drawing inspiration from the d-cent project, this dissertation will address this knowledge gap by building a web application to enable the study of the migration in block chain technology from economic to more-than-monetary purposes. This web application will innovatively explore and analyse the full historical transaction data contained within the block chain to measure these migrations. Emphasis will be placed on Agile software development techniques resulting in a comprehensively-tested application, with a well defined test suite for all libraries, modules and API's. The web-application will be implemented in *Scala* and will employ a modular design utilising available APIs such as the *libbitcoin C++* library for interfacing with the block chain. The final application will be a proof-of-concept that creates basic data-analytic tools for investigating the various uses of the block chain. Given the focus of the proposed project, it fits well within the *MSc in Data Analytics* stream.

A.2 The Field

The usefulness of the block chain for economic purposes has been widely recognised since the inception of decentralised cryptocurrencies such as Bitcoin (Grinberg 2011; Meiklejoh et al 2013; Mittal 2012). In a 2013 article in New Scientist, Hodson states the block chain "has the potential to transform commerce". Along the same lines, Rosenfeld argues that: "Bitcoin has revolutionized currency by creating a medium of exchange that can be stored and transferred digitally without reliance on any single third party, with overreaching implications for efficiency, security and counterparty risk" (2012: p. 1).

A recognition of the usefulness of the block chain beyond the economic realm has developed slowly over the years following the introduction of Bitcoin. In *The Second Bitcoin Whitepaper* (2012), J R Willet proposes MasterCoin as a supplementary protocol, which can be layered on top of the Bitcoin block chain to allow for many of the same purposes for which the block chain can be used, ranging from "smart property, company stock, deterministic contracts, bonds, demand deposits, emergent currencies, decentralised digital representation of physical assets" (Rosenfeld ibid). Willet's proposal formed the inspiration for what has been termed the 'second wave of block chain' - the general name given to a growing number of companies developing block chain technologies for a range of uses including Ethereum (who are developing a block chain with a built-in Turing-complete programming language, allowing anyone to write smart contracts and decentralized applications where they can create their own arbitrary rules for ownership, transaction formats and state transition functions); Blockstream (which aims to allow more block chains to be connected to the main Bitcoin block chain); and Factom (which maintains a permanent, time-stamped record of data in the block chain for purposes of more transparent auditing).

Close attention has also been paid to the further purposes of the block chain by Mainelli and von Gunten, who, in *Chain of a Lifetime: How Block Chain Technology Might Transform Personal Insurance*, state that "block chain's main innovation is a public transaction record of integrity without central authority" (2014: p. 4). The focus on integrity and ethics in further uses of the block chain has also been articulated by Scott (2014), who contends that "outside the media hype [around Bitcoin], a deeper movement is developing. It focuses…on the more general potential for *decentralised blockchains* to disrupt other types of centralised information intermediaries". Key within this literature is an attendance to the principles of decentralisation, distribution, replication, public verifiability and the immutable storage of data that underpin the functionality of the block chain. This

significantly extends the reach of what the block chain can be deployed towards, opening up to more outward facing, sociological and ethically oriented applications. Such considerations have underscored a very recent turn in the field of debate from the economic to the more-than-monetary (Al Kawasmi et al 2015; Bentov and Kumaresan 2014; Kostakis and Giotitsas 2014). However, very little scholarly work has as yet reviewed and explored the extent to which the block chain has been used to such ends.

B. BACKGROUND

B.1 Problem Domain

As of April 2015 there are no web applications publicly available which allow easy searching of text within the Bitcoin block chain. This is problematic because an analysis of the uses of the block chain beyond financial transactions requires an ability to search those fields connected with the transfer of message information within the block chain.

B.1.1 Limitations of Currently Available Tools

Given block chain technology is in its infancy, existing applications are relatively homogenous in the scope of their intended use (see for instance blockchain.info, blockexplorer.com, coindesk.com). Their features predominantly address commercial aspects of Bitcoin such as network activity, transaction volume, number of unique addresses used, market valuation of transactions and so forth. However none of the underlying technologies have been composed to form a fully extensive text search tool, despite the fact that the building blocks for such an application exists.

B.1.2 Block Chain Meta-Data

The text field is the site that must be employed by users to link a reference to the 'more-than-monetary' uses of the block chain, an example of which might be a hash to a sensitive data dump made available by Wikileaks (Assange, 2014: p.101). It is of key importance that such a tool be developed as a first step in being able to analyse the extent to which the block chain can be, and is being, used for 'more-than-monetary' ends.

B.1.3 Proposed Application

The proposed application will enable users to search through the block chain text fields assigned to each transaction within the full transaction history, thus predicting further analysis of such data.

B.2 Application Research

The assemblage and consolidation of various open-source components, frameworks, libraries and modules in novel ways will be necessary to produce an application capable of aiding a scholarly analysis of the data within the block chain. In the following section some of the benefits and drawbacks of the various components will be considered.

B.2.1 Block Chain Meta-Data Parsing

Directly interfacing with the block chain requires knowledge of *C++*. Although there are libraries written in other languages such as *Python* or *Java* these do not tend to scale with complex queries as they are much slower. *libbitcoin* is an extensive Bitcoin toolkit library written in *C++*. It is an extendable, scalable and configurable architecture making it much easier to interact with the block chain and is considered to be one of the most advanced alternative implementations of the Bitcoin protocol (Buterin, 2013). It uses *ZeroMQ* for which there is a *Scala* binding - allowing fast access to the data contained within the block chain.

B.2.2 Text-Data Parsing

libbitcoin returns *JSON* objects containing the requested data from the block chain. *Scala* is very well suited to simple *JSON* object manipulation, so no further tools will be required for this task.

B.2.3 Software Development Tools

There are a wealth of technologies which can enable the quick development of modern web applications. Development will be done in IntelliJ, with the approach of Behaviour Driven Development. This is seen as most suitable as the approach as writing clear tests for each feature prior to the code for the feature being written will ensure well documented and robust software; *ScalaTest* will be used to meet these ends. Compared to other testing libraries in *Scala, ScalaTest* is agnostic and is compatible with other testing frameworks such as JUnit.

C. PROJECT IDENTIFICATION

The main aims of the development portion of this dissertation will be to produce an application which can search the block chain text data field to investigate the changing uses of block chain technology since its invention. This project proposes to achieve this by developing a Scala application combining various open-source and non-proprietary programmes. *Scala* is a good choice due to it's static typing, so is quicker than other options such as *Ruby* or *Python* - but also involves less boilerplate code compared to other statically typed languages such as *Java.*

C.1 Core System

Table 0 presents a simplified proposition of basic inputs and outputs to achieve the project's aims:

Table 0 Application Inputs and Outputs

INPUTS	OUTPUTS
User query	Visual representation of user query
Block chain data	Associated metadata for the relevant transactions
	libbitcoin JSON object

This project will circumnavigate the limitations of current tools used to search the block chain, by recomposing the granular tools made available as part of the *libbitcoin* toolkit. In this way it can be considered a novel and timely solution to the questions we are seeking to answer.

C.2 Use Case and Requirements

C.2.1 Use Cases

There are two distinct use cases that this project will seek to address:

Firstly, for a user who wants to search for transactions in the block chain which contain some text in the corresponding field. The application should present the return results in a clear and concise way.

Secondly, as an extension allowing analysis of transactions in the block chain which may be 'more-than-monetary' - this will allow for a whole secondary research production of datasets supporting further scholarly research into alternate social uses of the block chain.

C.3 Determining Requirements

Requirements are laid out below in Table 1. They are the specifications of the requirements that must be satisfied to fulfil the use cases.

Table 1 Functional and Nonfunctional Requirements

Functional Requirements	Nonfunctional Requirements
User can modify queries that can be made to the block chain	The application should not depend on any proprietary software
User can see data analysis of the composition of results from the block chain from within date ranges of available data	The web application should be useable from all modern web browsers
User can make refining queries from within returned results to search for specific text	The return of results to the user should not take more than 5 seconds once input has been selected
	The application should warn about invalid queries as soon as possible

C.4 Assessing Suitability

Assessing various *Scala* compatible libraries and tools will be an important part of this project, and it is critical to establish the criteria by which different libraries, modules and frameworks will be assessed for achieving the project's aims. Below is a suggested methodology for assessment:

Table 2 Determinants of Component Software Suitability

Area	Topic	Concern
Querying block chain	Difficulty level for developer	Interface between *Scala* and *C++* library proves to be unstable requiring a new route being built
	Computational difficulty	The block chain grows by 1GB per month so server memory may become more of an issue over time
	Stability of tools	Queries will not be able to be made in the event *libbitcoin* becomes unavailable
Text-parsing	Text field nonconforming requiring complex parsing	Will this field yield data that can be used towards a statistical analysis.
Web application	Difficulty level for developer	Focus of this project is not solely on building a web-application - concern is that there be too much boilerplate coding required
	Stability of framework	Need to ensure that framework relies only on stable modules and libraries to ensure

		final application is as stable as possible.

D METHODOLOGY

D.1 Software Development Planning

The principles of BDD (Behaviour Driven Development) will guide the building of the application. This approach is typified by the practice of breaking down the desired functionality into 'user-stories' (Agile Alliance, n.d.). The 'business-needs' drive which 'user-stories' are first selected to ensure a MVP (Minimum-Viable-Product) is established in the shortest development time possible.

The application will be developed using five iterative process sprints, each taking two weeks. The goal will be to ensure that at the end of each sprint the application is working with any additional functionality added, being well tested and stable. All code will be version controlled using Git to ensure ease of auditability and protection from data-loss.

D.2 Project management

At the genesis of the project, Unified Modelling Language (UML) will be mapped out to explore and identify key deliverables in terms of required models (International Organization for Standardization / International Electrotechnical Commission, 2012). These will be reviewed at the beginning of each two-week sprint.

A daily log will be kept during the application development to list all of the design decisions made. Issues will be noted in an accompanying wiki in order to maintain full documentation as the project progresses. The process of encountering and solving problems will be immediately documented so as to avoid unnecessary overhead in trying to do this at the report writing stage.

D.3 Limitations

Given the proposed innovative composition of libraries, modules and frameworks, it is crucial to detail what this project is **not** trying to achieve. This is laid out in the *Appendix: F.1. Limitations*, and as such this application should not be considered as industry complete.

D.4 Stages

D.4.1 Planning and Exploration

D.4.1.1 Phase Zero

This initial phase will involve trialing the possible components (libraries, modules, APIs and their respective *Scala* implementations/interfaces). The tools will be experimented within minor examples to gauge their suitability and viability for the final application. The format of the *JSON* object will be researched to allow creation of a mock object for use in the development stages of the application.

D.4.2 Implementation stages

D.4.2.1 Phase One

During this first sprint phase the base web-application will be built using *Scalatra*. RESTful methods will be produced to query the Bitcoin block chain. A mock of the *libbitcoin JSON* object will be used to isolate the regions of the application under development at any one time. By the end of this phase it is expected that a rudimentary application for making a query call to the block chain will be completed.

D.4.2.2 Phase Two

The second sprint phase will focus on ensuring that the web application can make a query to the Bitcoin block chain via the *libbitcoin* toolkit library through the ZeroMQ Scala binding. This will comprise the bulk of time for this sprint, to ensure the return *JSON* object from *libbitcoin* matches the expected input from Phase One.

D.4.2.3 Phase Three

Analysis of how the uses of the block chain have changed over time requires a deeper analysis of the metadata and text data which can be extracted from the returned JSON object upon querying the block chain. In this third sprint phase, decisions will be made about how best to achieve a statistical analysis of these changes. Questions on how to present the returned information will be engaged.

D.4.2.4 Phase Four

The fourth sprint phase will explore additional functionality. This might include further analysis of the *types* of text available from the text field - plain text or a hash reference. It may also be possible to follow any outlinks for further analysis. Refinement of queries made to the block chain will be developed as more nuanced understandings of which sites of information yield the most statistically

significant findings are advanced. During this sprint, it is also possible for more creative features to be added to the application. This may include inbuilt statistical analysis using R to figure out any underlying patterns in the distributions of the returned data-set. It may also be advantageous to develop the clear-text search engine functionality of the block chain query engine: functionality of this sort could be useful to those employing the block chain as a immutable distributed data-store.

D.4.2.5 Phase Five

Sprint five presents the opportunity for a final testing phase and opportunity to finalise any unfinished functionality for previous sprints. It will also be during this period that the accompanying report will be collated, as well as additional documentation for the code-base.

D.5 Timeframe

A proposed outline for key deliverables and project outline is presented in Table 3 below.

Table 3 Timeline for Project

Start	End	Activity
-	22nd June	Planning and Experimentation (Phase 0)
23rd June	6th July	Phase 1
7th July	20th July	Phase 2
21st July	2nd August	Phase 3
3rd August	19th August	Phase 4
20th August	13th September	Phase 5
14th September	-	Latest Submission

D.6 Risk

Outlined below in Table 4 are possible threats and opportunities to the project (Office of Government Commerce, 2009).

Table 4 Table of Risks and Opportunities

Risk	Issue	Response
Threat	Run out of time	Revert back to MVP from previous sprint. Assess this during Phase 3/ Phase 4
Threat	Text field data from block chain does not yield statistically significant information	Investigate alternate aspect of *JSON* object returned which could be used to flag non-commercial uses of the block chain. Assess in Phase 2
Threat	Selected tools become too time consuming	Either choose more basic route to MVP or switch languages / frameworks to simpler tools
Opportunity	Completes aims early	Develop further statistical analysis measurements for the returned datasets to develop a deeper analytical model for uses of the block chain
Opportunity	Some Phases completed earlier than planned	Either add a new feature, or do preparation for the next Phase

BIBLIOGRAPHY

Agile Alliance (n.d) *BDD Definition* [Online] http://guide.agilealliance.org/guide/bdd.html [Accessed 15 March 2015]

Al Kawasmi, E., Arnautovic, E. and Svetinovic, D. (2015) Bitcoin-based decentralized carbon emissions trading infrastructure model. *Systems Engineering* 18(2): 115–130

Assange, J. (2014) *When Google Met Wikileaks*. New York: OR Books.

Back, A., Corallo, M., Dashjr, L., Freidenbach, M., Maxwell, G., Miller, A., Poelstra, A., Timón, J. and Wuille, P. (2014) Enabling blockchain innovations with pegged sidechains. *Blockstream* [Online] http://www.blockstream.com/sidechains.pdf [Accessed 15 March 2015]

Bentov, I. and Kumaresan, R (2014) How to use Bitcoin to design fair protocols. *Advances in Cryptology – CRYPTO 2014 Lecture Notes in Computer Science* 8617: 421-439

Bitcoin (n.d) How it works? [Online] https://bitcoin.org/en/how-it-works [Accessed 12 March 2015]

Bitcoin (n.d) *What is Bitcoin?* [Online] https://bitcoin.org/en/faq#what-is-bitcoin [Accessed 10 March 2015]

Blockchain Info (n.d) *Blockchain Charts* [Online] https://blockchain.info/charts [Accessed 15 March 2015]

Block Explorer (n.d) [Online] https://blockexplorer.com/ [Accessed 15 March 2015]

Brito, J. and Castillo, A. (2013) *Bitcoin: A primer for policymakers*. The Mercatus Center, George Mason University [Online] http://mercatus.org/sites/default/files/Brito_BitcoinPrimer.pdf [Accessed 10 March 2015]

Buterin, V. (n.d) A next-generation smart contract and decentralised application platform. *Github* [Online] https://www.ethereum.org/pdfs/EthereumWhitePaper.pdf [Accessed 15 March 2015]

Buterin, V. (2013) What libbitcoin and SX are and why they matter [Online] https://bitcoinmagazine.com/6234/what-libbitcoin-and-sx-are-and-why-they-matter/ [Acessed 15 March 2015]

Coindesk (n.d) *Bitcoin market capitalisation* [Online] http://www.coindesk.com/data/bitcoin-market-capitalization/ [Accessed 15 March 2015]

Coindesk (n.d) *Bitcoin network data* [Online] http://www.coindesk.com/data/bitcoin/ [Accessed 15 March 2015]

D-cent Project (n.d) [Online] http://dcentproject.eu/ [Accessed 15 March 2015]

Greenberg, A. (2011) Crypto currency. *Forbes* [Online]

http://www.forbes.com/forbes/2011/0509/technology-psilocybin-bitcoins-gavin-andresen-crypto-currency.html [Accessed 15 March 2015]

Grinberg, R. (2011) Bitcoin: An innovative alternative digital currency. *Hastings Science and Technology Law Journa*l 160 [Online] http://hstlj.org/articles/bitcoin-an-innovative-alternative-digital-currency/ [Accessed 15 March 2015]

Hodson, H. (2013) Bitcoin moves beyond mere money. *New Scientist* [Online] http://www.newscientist.com/article/dn24620-bitcoin-moves-beyond-mere-money.html#.VSYSNxfwN3M [Accessed 15 March 2015]

International Organization for Standardization / International Electrotechnical Commission (2012) *Formal/12-05-07: Unified Modelling Language (UML) - Superstructure, v2.4.1.* [Online] http://www.omg.org/cgi-bin/doc?formal/12-05-07.pdf [Accessed 10 March 2015].

libbitcoin (n.d.) *Welcome* [Online] http://libbitcoin.github.io/ [Accessed 15 March 2015]

Mainelli, M. and von Gunten, C. (2014) *Chain of a lifetime: How blockchain technology might transform personal insurance*. Z/Yen Group [Online] http://www.longfinance.net/publications.html?id=903 [Accessed 15 March 2015]

Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko., McCoy, D., Voelker, G. and Savage, S. (2013) A fistful of Bitcoins: Characterizing payments among men with no names. *Proceedings of the 2013 Internet Measurement Conference* [Online] http://cs.gmu.edu/~mccoy/papers/imc13.pdf [Accessed 15 March 2015]

Mittal, S. (2012) Is Bitcoin money? Bitcoin and alternate theories of money. *Social Science Research Network* [Online] http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2434194 [Accessed 15 March 2015]

Nakamoto, S. (2008) *Bitcoin: A peer-to-peer electronic cash system* [Online] https://bitcoin.org/bitcoin.pdf [Accessed 15 March 2015]

Office of Government Commerce (2009) *Managing Successful Projects with PRINCE2*. Norwich: The Stationary Office.

Rosenfeld, M. (2012) *Overview of colored coins* [Online] https://bitcoil.co.il/BitcoinX.pdf [Accessed 15 March 2015]

Sagona-Stophel, K. (n.d) Bitcoin 101. *Thompson Reuters* [Online] http://site.thomsonreuters.com/business-unit/legal/digital-economy/bitcoin-101.pdf [Access 17 March 2015]

Scott, B. (2014) Visions of a techno-leviathan: The politics of the Bitcoin blockchain. *E-International*

Relations [Online] http://www.e-ir.info/2014/06/01/visions-of-a-techno-leviathan-the-politics-of-the-bitcoin-blockchain/ [Accessed 15 March 2015]

Snow, P., Deery, B., Lu, J., Johnston, D. and Kirby, K (2014) Factom: Business processes secured by immutable audit trails on the blockchain. *Github* [Online] https://github.com/FactomProject/FactomDocs/blob/master/Factom_Whitepaper.pdf?raw=true [Accessed 15 March 2015]

Swarm.Fund (2014) The second wave of blockchain innovation. *Medium* [Online] https://medium.com/@Swarm/the-second-wave-of-blockchain-innovation-270e6daff3f5 [Accessed 15 March 2015]

Vasilis, K. and Giotitsas, C. (2014) The (a)political economy of Bitcoin. *Journal for A Global Sustainable Information Society* [Online] http://triplec.at/index.php/tripleC/article/view/606 [Accessed 15 March 2015]

Wiles, J. (n.d.) *The second bitcoin paper draft* [Online] https://sites.google.com/site/2ndbtcwpaper/2ndBitcoinWhitepaper.pdf?attredirects=0 [Accessed 15 March 2015]

E. APPENDIX

E.1 Possible Frameworks, Modules and Libraries

E.1.1 Web Application

E.1.1.1 *Scalatra*

A rapid development web framework such as *Scalatra* will be used to create a MVC (model-view-controller) web-application which will make calls to the *libbitcoin* API to query the Bitcoin block chain.

E.1.1.2 *Angular*

On the front end a *JavaScript* library such as *Angular* or *KnockoutJS* will be used to present the information to the user in a presentable way and interactive way.

E.1.1.3 *PostgreSQL*

For the purposes of this project a Relational database will be a suitable data structure from which to query the returned data from *libbitcoin*.

E.1.2 Bitcoin Block Chain Exploration

E.1.2.1 *libbitcoin*

An extensive Bitcoin toolkit library written in *C++* which has an extendable, scalable and configurable architecture, making it much easier to interact with the block chain. Unique in its modularity, the toolkit has the most advanced tools for interacting with the block chain.

E.1.2.2 *Scala binding for ZeroMQ*

This will allow API calls using *Scala* from *Scalatra* to be sent through *ZeroMQ* and converted to *C++* for communication with the *libbitcoin* API. *libbitcoin* returns *JSON* objects, the data from which will then be passed back to the *Scalatra* web application

E.1.3 Scala Application Development Tools

E.1.3.1 *IntelliJIDEA*

There is an excellent *Scala* plugin in for IntelliJ which is a robust editor to develop the web application and associated tests.

E.1.3.2 *ScalaTest*

This will be the test suite of choice given the agnostic approach *ScalaTes*t takes to the specific test approach used. It is also backwards compatible with *Java* test libraries such as *JUnit*, which allows maximum flexibility in testing. Although writing tests first adds some overhead to development time, it is industry best-practice and better ensures that the application is working at each stage.

F. APPENDIX

F.1 Limitations

The following is a non definitive list of features that will not be attempted within the scope of this project:

- *Advanced statistical analysis of the composition of block chain data;*
- *Analysis of block chain data in real time. This would require 'cloud computing' as the current size of the block chain is 32GB (https://blockchain.info/charts/blocks-size);*
- *Figuring out relationships between connected messages on the block chain;*
- *A graph database look at the block chain to map the paths through which transactional messages have been inscribed on the block chain;*
- *A comparative analysis between wider network statistics such as aspects of the changing nature of the Bitcoin network and the changing nature of the uses of the block chain;*
- *A complete user management interface for the web application;*
- *Functionality to allow saving or exporting of search results;*
- *In general the look and feel of the web application is not of primary concern in this proof-of-concept.*