Engaging Practitioners within Design Science Research:
A Natural Language Processing Case Study

Seán O'Riain, Edward Curry, Paul Buitelaar

Digital Enterprise Research Institute (DERI),
IDA Business Park, Lower Dangan, Galway, Ireland
{sean.oriain, ed.curry, paul.buitelaar}@deri.org

Abstract. Using domain knowledge to instruct domain specific Natural Language Processing (NLP) applications requires that knowledge intensive design challenges associated with developing extraction rules and conceptual models from that knowledge be addressed. Applying the nested problem-driven approach of Design Science Research (DSR) assists knowledge problem reduction to practical problems, delivering artifacts of utility. Where artifact design has to facilitate practitioner and research stakeholder expectations, dual leveraging of design process stages and conceived artifacts is required. This paper presents how an existing Information Systems (IS) framework, previously applied to enterprise architecture research, can be adapted to enhance stakeholder engagement within a collaborative informatics research project. The business motivation behind domain specific NLP is explained and design challenges encountered in framework application to use case development, discussed. Further contributions that outline artifact evolution using problem decomposition are made through integrating expert domain knowledge and design knowledge translation as part of the adapted research process.

Keywords: Natural Language Processing, NLP, Information Extraction, Design Science Research, Financial Analytics, Ontology Modeling, Knowledge Transformation

1   Introduction

For academics engaged in information systems (IS) industrial collaboration, balancing industrial expectation with academic research pursuits is a constant challenge. Disconnects between research partners and practitioners ensure that activities with either complimentary or common goals are often overlooked and opportunity to leverage output in respective fields is seldom realized to its full potential. Ensuring that research activity remain practicable and relevant while at the same time delivering a quantifiable research contribution to the knowledge base, demands artifacts that satisfies both research partner and industrial practitioner expectations. Overlap do exist; both Design Science and an application domain environment produce results of

knowledge and artifacts as models, constructs, methods and instantiations [1-3]. Introducing clarity between researchers on the one hand and practitioners on the other requires that adequate collaborative engagement takes place supported by rigorous procedures that satisfy expected solution artifact delivery.

To investigate this dynamic we apply Design Science Research to the IS subdomain of Natural Language Processing (NLP), specifically addressing the framework adopted, its differing deliverable types and how they can be assessed and measured by researchers and practitioners as stakeholders. Using a collaborative industrial - research informatics project[1] (described in Section 1.1) we outline how the framework facilitates practitioner business need through the application development of a qualitative decision support system on the one hand and researcher knowledge base contribution on the other, using complimentary artifacts design (cf. Section 4). Adhering to design-orientated information systems research process principles [4], and design-science research guidelines [1], our methodology in keeping with Meyer [3], centers on the intended application environment, artifact development and its changing nature throughout the research process. Key to our contribution and in similar vain to Wieringa [5], is the application of DSR's nested problems decomposition approach to research project structuring, allowing research methodology stages enrichment with researcher artifact output and its alignment with existing practitioner artifact output. Our research aims to reinforce existing IS design theory by providing guidance to collaborative research projects participants in terms of achieving engagement that delivers on these differing end goal expectations.

This paper is structured as follows: Section 1.1 first introduces the business motivation driver behind the research use case and the design goals required. In Section 2, NLP, positioned under the broader area of Information Extraction considers implications for Design Science stemming from the design goals. Adaption of Design Science frameworks and methodology to reconcile practitioner and researcher artifact and deliverable expectations are introduced in Section 3. Section 4 discussed research process application to our informatics use case and lessons learned are presented in Section 5. Section 6 concludes the paper.

1.1   Business Motivation and Value Proposition

Competitive analysis is used as an investigative tool by business analysts to deliver insight into critical business processes. Those processes include: a firms or competitor's operations and strategy; understanding market movement; identifying competitors; determining strengths and weaknesses; and prediction of the next strategic and tactical moves [6]. Competitive analysis monitors competition or environmental factors, captures essential measures of activity, and organizes those measures to help decision makers detect and respond to changes [7]. In particular, the free text management statements that comment on corporate performance and intangibles such as people, brands and patents are actively searched for key information and interpretation of company position. Manually locating and correlating key information from

---

[1] Concentrating on the areas of competitive intelligence that targets fact extraction from business reports

within the financial statements is recognized as presenting particular difficulty due to their textual nature, lack of structure and lack of common format [8]. The filings size and sheer volume, ensures that up to 75% of analyst resource availability is expended in information gathering to support analysis [9]. Previous investigations have noted that analysts conducting such manual information acquisition dedicated 12.5% of available time to searching the filings introduction sections and establishing where in the filing to look for relevant information and the remaining 87.5% analyzing the identified sections [10].

The use case company, due to changing business practices, tasked its Business Process Outsourcing (BPO) team with the identification of software vendors as potential sources of new business opportunity for the outsourcing of their software and supporting production services. Manual competitive analysis was their main analytical tool. An artifact that could reduce the level of manual effort while assisting in identifying key pieces of information to better support competitive intelligence insight was required. Design goals contributing to the emergence of such an application artifact would minimally have to support individual knowledge transformation processes of: i) linguistic modeling of filing narrative section based on analyst heuristics and insight; ii) automated linguistic analysis of filing using the conceptual linguistic model; and iii) an artifact that provides an environment to support analyst perform the competitive intelligence task.

2   Natural Language Processing Implications for IS Research

Information Extraction (IE) as a fundamental process involves taking unstructured text input and outputting formatted unambiguous data [11]. IE systems can be categorized as adhering to knowledge engineering (KE) or machine learning (ML) approaches. KE depends on domain expert knowledge to hand craft extraction rules that are used to automatically identify and extract information as part of a natural language processing activity. ML on the other hand utilizes algorithms to train on annotated text and automatically generate extraction rules[2]. IE applications in specialized information domains such as financial, business, medical or biology are hugely reliant on this domain knowledge to define the conceptually complex information sought[3]. Central to any KE successful design is the process used to formally represent that knowledge and its decomposition into manageable sub-processes.

Any DSR originating methodology applied to this environment must therefore take account of and allow for knowledge intensive processes involving domain expert practitioners at all levels of the research methodology. Knowledge intensive processes have however challenging information requirements. They require knowledge and expertise in its application. Markus et al. [12] researched design problems providing IT support to a class of user requirements termed emergent knowledge processes

---

[2] For discussion on general data extraction methods we refer the reader to [13] and for data mining techniques related to financial applications to [14].
[3] For example, in business the concept of 'market movement' is more difficult to define and identify in text than concepts such as person, location or lesser events.

(EKP). Product development and strategy-making are organizational EMP examples that bring a level of practitioner uncertainty. EKPs' are characterized by: i) emergent processes of deliberations with no best structure or sequence; ii) unpredictable actor sets with regard to job role or prior knowledge and; iii) knowledge requirements for general and specific distributed expertise [12]. Problem solving within Design Science has been classified as mutual nesting of knowledge problems and practical problems [5]. Design questions can contain knowledge questions which can in turn be decomposed into known practical sub-problems solvable with existing methodologies. Knowledge questions are answerable from the knowledge base or with further research such as conceptual analysis, empirical research using experiments, case studies or modelling. Practical problems are solved by matching problems and solution as part of a regulative cycle [5].

Characteristics of NLP are only partly shared with EKPs. The customizable ability of knowledge engineering based NLP ensures that a best structure, typically minimally represented by a semi-formal model, emerges. Both the practitioner and required knowledge must be identified during design for a rule set to emerge. The last characteristic of dependency on both general and specific environmental knowledge is shared with EKP. For specialized domain applications knowledge can be derived from that generally available in the knowledge base in the first instance (e.g. in finance: standards, best practices and regulatory rules) and practitioner heuristic knowledge processes (e.g. operational know how, experience). DSR application to the area of customizable NLP is therefore suitable as its frameworks support IS artifact design and development through its nested problem solving capability, allowing the reduction of knowledge problems into manageable practical problems. Understanding the evolving artifact state throughout the design process and the accommodation of practitioner and researcher engagement as part of that process, is key to ensuring that resulting artifact deliverables reflect stakeholder need. Accommodating this engagement, as will be discussed, proved to be critical for research success. Similarly, principled uses of DSR has seen its application to developing modeling techniques for service design [15] and both direct and indirect effects of environmental sustainability on artifact design [16].

3   Design Science Application to Natural Language Processing

The methodology used adheres to the design-orientated information systems research process guidelines and principles from [4], and design-science research guidelines from [1]. The overall methodology applied in our use case follows the four basic process phases of analysis, design, evaluation and diffusion [4]. Principles outlined in Table 1, consider each DSR principle from an NLP knowledge intensive view point. The considerations are reflected in the research framework (cf. Section 3.1), and abstracted in the research methodology (cf. Section 3.2). Detail discussion is left to the case study in Section 4.

Table 1. Aligning DSR guidelines to Information Extraction

Guideline	Description	Consideration for IE
Design as an artifact	DSR must produce a viable artifact in the form of a construct, model method or instantiation	We designed a number of artifacts that contribute towards a main construct that supports a business analyst: an ontology to specify the information requirement; a linguistic analytics pipeline that provides that information and; a qualitative DSS, that combines these artifacts, supporting an analyst find and retrieve information
Problem Relevance	Objective of DSR is to develop technology based solutions to important and relevant business problems	Established by business stakeholders. In general consideration of any enhancement in the provision of contextualized information has business utility and value
Design Evaluation	Utility, quality and efficacy of a design artifact must be rigorously demonstrated with well executed evaluation methods	Artifact was implemented and evaluated as part of an industrial use case and demonstrator, using domain experts
Research Contribution	Effective DSR must provide clear and verifiable contributions in the areas of the design artifact, design foundations and/or design methodologies	The research contribution is centered on the design artifacts that concentrate on the ability to specify information sought and its provision to the relevant audience
Research Rigor	DSR relies on the application of rigorous methods in both the construction and evaluation of design artifacts	For both design and evaluation methods we utilised methods from computer science, information retrieval and information science to deliver research rigor
Design as a search process	Search for an effective artifact requires utilising available means to reach desired ends while satisfying laws in the problem environment	Each stage of the design process was performed iteratively and in collaboration with the business stakeholders. Stakeholder engagement ensured that environmental needs and rules are satisfied, and that a rich knowledge base was also available
Communication of research	DSR must be presented effectively both to technology-orientated as well as management-orientated audiences	Our research, multidisciplinary in nature, required both business and technical knowledge. Dissemination activities have targeted the management-orientated Business Intelligence and technology-orientated semantic applications audiences

3.1 Research Framework

As illustrated in Figure 1, the IS framework employed is based on Henvers [2] original three cycle view of DSR and recent model modification to cater for its application to business value assessment for Enterprise Architectures [3]. The *Environment* consists of the company's application domain, organizational strategy, processes and technology and applicable stakeholders. The environment informs on relevant business problems, need and application context, allowing artifact requirement definition. Acceptance criteria is also included as part of the relevance cycle to ensure on-going

artifact alignment with business need and research design goals. Artifact design was explored using the companies *Knowledge Base* for exploitable frameworks or technology offerings. Where recommendations were not forth coming, the academic knowledge base was also interrogated for foundational technologies such as linguistic analysis [10, 14, 17] and complementary methodologies from information systems to assist with knowledge representation (i.e. ontology development [18, 19]) then can be adopted as necessary. *IS Research* leverages design-orientated information systems research process guidelines and principles [4] specifically to assist with artifact identification and evolution. The research guidelines are expanded with a research methodology (cf. Section 3.2) adapted to enhance practitioner and researcher engagement and meet their different artifact expectations. Artifact design rigor was enhanced using the principles of [3, 4], namely: abstraction, ensuring that artifacts solved a class of problems; originality, catering for artifact contribution to both knowledge bases; validation, artifacts must be justified and; benefit, artifacts must deliver business value to stakeholders.
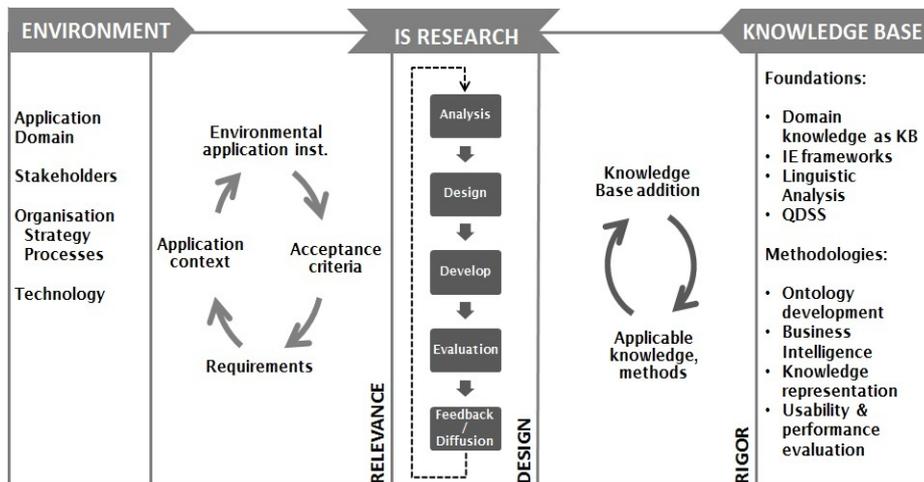

Figure 1. DSR Framework adapted from [2, 3]

3.2   Research Methodology

The IS Research process used focused on the four process phases of evaluation, design, evaluation and diffusion [4]. Each phase is broken down into sub-stages (as indicated by the process boxes) that prompt design questions and decomposition of knowledge problems into practical problems [5]. Each process box has associated artefact output represented by document boxes. Practitioner interest artefacts are listed under processes boxes as the left most document box and researcher as the right. Discussion on artifact evolution and the engagement necessary between stakeholders to achieve these documents as part of the design process is presented in Section 4.
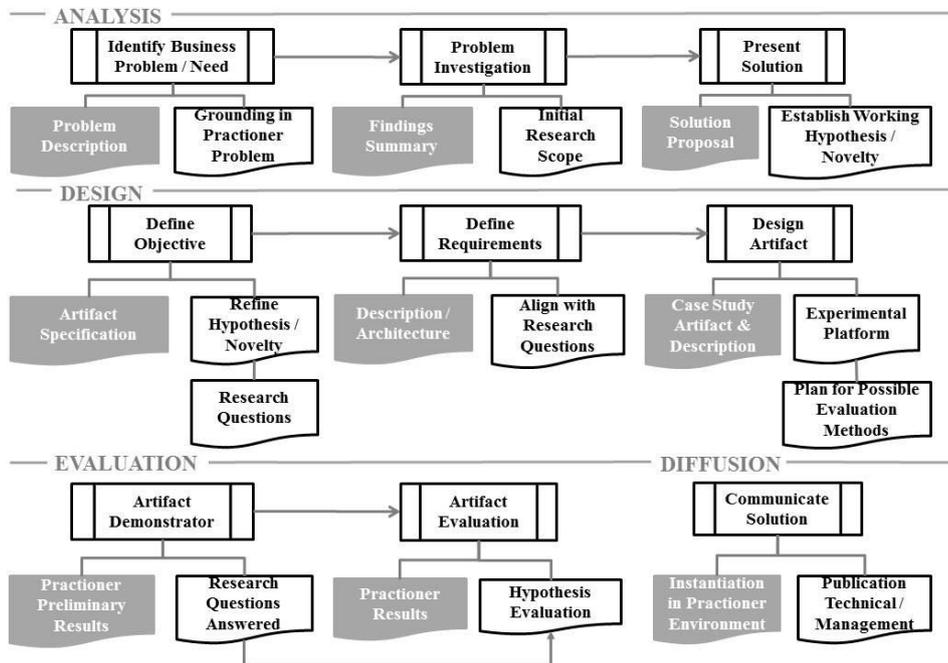
Figure 2. Research Methodology including deliverables adapted from [3]

4   Case Study: DSR Application to Business Informatics

Our use case involved the validation and evaluation of our approach (defined in Section 3.2) as part of a collaborative academic-industrial informatics research project. Practitioners were driven by the development of a qualitative decision support system to support business analysts perform a competitive analysis, while researchers had as goal, novel knowledge base artifact contribution (e.g. domain specific term lists, concept maps, ontologies), driven by a series of research questions, evaluated against a working hypothesis. The application of each research process phase to enhance stakeholder engagement and progress the use case is next discussed.

Analysis. The analysis phase established the business relevance, performed initial problem investigation and proposed a possible solution. The analysis was divided into three steps (cf. Figure 2). The first step concentrated on establishing the business need. Using the business motivation (Section 1.1) as background, a series of practitioner consultative interview sessions were undertaken with their business analyst team. The interviews investigated current knowledge base usage by practitioners reinforced with demonstrations to clearly define the environmental business problem. For practitioners this provided an initial frame of reference for the type of design artifact required and for researchers a context and level of transparency for the type of design innovation expected. Early engagement with stakeholders ensured that each had an understanding of the others interest areas and expectations from the project outset.

Defining a class of environmental business problems was used to progress problem investigation. Generalized recommendations regarding the development of new or customizations of existing IE/NLP systems [13] were first considered. Table 2 lists for each problem class design question summaries that outline influences on potential artifact design. An accompanying priority rating indicates the importance of addressing the problem class to the practitioner.

Table 2. Problem classes general to IE

Problem Class	Design Question Summary	Priority
Adequate specification of IE requirement	Is IE/NLP a suitable approach for the information task? Does the implementation cost outweigh expected results and does the enterprise have in-house technical competency? Would system predictability be better than random or might information retrieval results prove comparable or better? What is the resource overhead associated with rule maintenance, lexical resources and preparation of training data?	High
Availability of Knowledge resources	Is there availability of sufficient quantity and quality training data? If not, what is the cost of training data creation, accessibility to domain experts and availability of knowledge engineers for activities such as rule construction? Are there linguistic resources (e.g. dictionaries, taxonomies, ontologies or business specifications) that can be exploited?	High
Dealing with multiple text types:	What are the relevant source formats and text types that have to be gathered, processed and indexed? Will new domain information, language identification features, text genres and multilingual capability be required?	High for English
Adaptivity / reusability:	Will any created dictionaries, term lists or training data be adaptable to different business tasks?	Medium
Scalability	What type of processing response is required: Real time, off-line processing, or parallel?	Low

The generalized problem classes were found to be at too high a level to assist practitioners with emergent requirements or researchers with research scope and required further decomposed into specialized problem classes. Using the problem description and information gathered from the academic knowledge base, the financial IE problem class set (Table 3) complete with major design activities areas were defined. Influenced by the problem description, the sub-domain problem classes allow findings summary generation for the practitioner and based on knowledge base investigation, a research scope based on an initial set of working assumptions. Defining the problem class set ensured close engagement and problem understanding between stakeholders.

Table 3. Problem classes specific to financial IE

Problem Class	Design Question	Solution / IT Artifact
Extraction task supported	What is the financial/business sub-domain activity that the information extracted will directly support?	Competitive analysis with the Analyst Work Bench
Text type	What filing type to support text extraction from?	Construct to support extraction from XBRL / XML filings
Solution approach	Use the knowledge engineering or machine learning approach for extraction?	Construct Linguistic analytics pipeline using KE approach
Language model	Use a conceptual model to define domain linguistic and drive actual extraction?	Competitive Analysis Ontology
Lexical resource engineering	Development or use of existing domain lexical resources to support the solution approach?	Domain knowledge extraction as: i) process and ii) model

Step 3 saw the solution proposal presented to stakeholders for review. Conditional approval from multiple practioner roles was required, namely: management for access to domain expertise and environmental knowledge base, business for access to relevant sources, and technologists for assessment of solution offerings and technology fit within the application environment to business need. For researchers, identifying gaps in the knowledge base when considering solution design provided insight as to where innovative design could be applied, novelty contributed and a working hypothesis established. Using problem investigation and solution proposal to reconcile stakeholder understanding and achieve broad consensus on solution and research direction was instrumental in both establishing and consolidating stakeholder engagement.

Design. The design phase defines objectives, requirements and artifact design steps motivated by the business need. Clear and attainable objectives that aligned with artifact delivery from the problem statements (Table 3) and that also represents our research problem analysis are presented in Table 4. Although the artifact objectives are practitioner biased, the artifacts themselves evolved with researcher engagement through assessing their suitability to answer research questions and contribute to hypothesis evaluation at the evaluation stage. Based on the artifact objectives and agreed solution proposal (from the analysis phase), a detailed functional specification was produced along with an opportunity for researchers to re-visit their working hypothesis and define research questions that enable hypothesis validation. Stakeholder engagement during requirements definition allowed practitioners to introduce new or modify existing artifact functional specification and researchers an opportunity to validate whether developed artifacts would be apt to answer research questions.

Table 4. Design Solution Artifact Objectives

Artifact	Objectives
Competitive analysis with the Analyst Work Bench	Application that assists an analyst find relevant information as part of competitive analysis task performance
Construct to support extraction from XBRL[4] / XML filings	Improve ability to deal with regulatory reporting formatting standards
Construct Linguistic analytics pipeline using KE approach	Improve identification and extraction of relevant information through semantic mark up of business filings
Domain knowledge extraction as: i) process and ii) model	Capture tacit knowledge relating to competitive analysis task performance, information provision and information associations
Competitive Analysis Ontology	Formally represent knowledge requirement relating to competitive analysis task performance

Design knowledge frameworks that capture design knowledge across a multi-stage process that progresses individual design knowledge (implicit), to explicit design knowledge and finally computational design knowledge, have been proposed for ubiquitous information systems development projects [20]. The design process culminates with design knowledge representation as a knowledge object. The development process is perceived as a knowledge transformation system that does not distinguish between product and information. Generally suitable for our purposes we further define steps to take account of our knowledge intensive requirements for specialised application of NLP to financial IE. Knowledge problems identified from artifact objectives and requirements, relating to knowledge capture and conceptualization, triggered decomposition into practical sub-problems and existing methodologies were then applied to resolve them [5]. Table 5 shows the problem decomposition necessary for knowledge capture through to conceptualization and ontology generation based on the DOGMA[5] ontology development methodology [19]. Knowledge problems are first aligned with stages of DOGMA's domain conceptualization. Methods employed to decompose the problem into a practical problem are given and the resulting artifacts listed. For example knowledge elicitation and breakdown used contextual enquiry to perform discourse analysis and generate discourse term lists, concept maps[6] and their organization into statements of propositions. A similar activity that proposes the translation of process narratives (description of usage situations in natural language) into diagrammatic conceptual models termed pre-artifacts is found in Maass and Janzens' knowledge framework [20].

---

[4] eXtensible Business Reporting Language, a global standard for expressing the semantic meaning of information in business reports and its exchange, refer to http:/www.xbrl.org/.
[5] Developing Ontology Grounded Methods and Applications
[6] Semi-formal design knowledge based on a graph model with binary relationships between node (concept) types

Table 5. Knowledge problems aligned with Ontology Modeling Methodology based on [19]

Knowledge Problem	Method	Generated Artifact
Knowledge elicitation	Contextual enquiry	Discourse term list
Knowledge breakdown	Discourse analysis	Concept maps
		Concept proposition templates
		Semantic Paths
		Verbalized elementary sentences
Verbalized elementary sentences	Concept proposition templates	Extraction rules
		Ontology concept primitives
Knowledge negotiation	Semantic Paths	Competitive Analysis Ontology
Knowledge discovery	Lexon Engineering	

Using problem decomposition as the vehicle for design knowledge transformation introduced progressive artifact state change that had a direct bearing on stakeholder engagement and expectation management. Practitioner interest concentrated on progression towards final artifact design. Researchers on the other hand sought to assess:

i) Alignment of individual artifacts to inform on individual research questions
ii) Whether there were any usage or setup issues that may impact experimental validation performance and
iii) Fitness-for-purpose of the final artifact as an experimental platform to validate hypothesis against

Problem decomposition demanded selection of appropriate methodologies to address emergent practical problems. An observation regarding formal conceptual models was the lack of reliance on their use as a design tool for IS designers [20] but use instead as computational design knowledge. Our engagement model found this to be the case with the additional observation that semi-formal models (e.g. concept mappings) attracted greater interaction and usage from practitioners and formal models (e.g. ontology representation) attracted interest from researchers. Having respective stakeholders aware of artefact state but only taking ownership for those that matched their end goals, ensured that artefact design and development met expectations.

Evaluation. Having established design artifacts of interest during the design phase, stakeholders next progressed to perform instantiation and preliminary assessment of those artifacts relating to analysis goals, problem categories and design objectives. Specifically for researchers artifacts generated from addressing practical knowledge problems (Table 5) allowed research question evaluation. With feedback incorporated, the research process moved to the rigor phase and performing comprehensive evaluation of the main artifact based on design objectives and goals. Earlier stakeholder design phase discussions identified a number of areas where experimental design, setup and execution required further close engagement, owing to:

i) Artifact contribution validation through competitive analysis task information provision and effect on the qualitative aspects of task performance
ii) Knowledge complexity associated with verifying objectives and goals, demanded that evaluation take place within an organizational context and be performed using domain experts

iii) The alignments of design phase case study document descriptions with objectives, identifying that performance and usability evaluations were required

Criteria adopted for verifying artifact performance and usability are outlined in Table 6. The academic knowledge base provided the performance methods from information retrieval and usability measures from information science. Performance delivers a quantitative set of results and usability a qualitative set. Worthy of mention is that evaluation criteria for knowledge depends more on its truth value with respect to the subject domain of the knowledge rather than stakeholder goals [5], reinforcing the need for domain expert assessors. Post evaluation there is opportunity to re-visit the artifact analysis with corrections or modification to the original solution approach and design.

Table 6. Evaluation Criteria

Criteria	Performance	Usability
Artifact usable as basis for experiment	AWB QDSS, as a provider of relevant information	AWB, as an interactive QDSS
Criteria representing system objectives	Relevance of information provided	Usefulness/usability of information provided and DSS environment
Measuring instrument	Relevance judgment expresses as a binary weighting	Success determination using Likert scale
Measures	Precision, recall	Weighted average
Methodology for measurement, evaluating performance	Based on the competitive analysis task	Questionnaire survey of participant usage of artefact

Diffusion. Managing artifact state change increased the frequency of artifact availability and opportunity for dissemination exploitation across the broader research process rather than post phase completion. For researchers this represented business information system [23] and domain technology-orientated [24] publication audiences [1]. Practitioners on the other hand looked to use case demonstrator instantiation within their environment.

5   Lessons Learned

Based on the application of our research methodology to a collaborative research informatics project we have identified the following four lessons that can be applied by practitioners and researchers in other projects. We believe these lessons are equally applicable within general knowledge engineering dependent business analytics activities and collaborative research projects.

i) **Align Business and Research Objectives**

A key requirement is to align the business problem with researcher design objectives. Alignment success was attributed to researcher grounding in the practioner problem, early establishment of environmental problem classes and their use to direct knowledge base searches. Knowledge base searches were instrumental in identifying research novelty, solution design and project scope. Adhering to this approach we

were able to identify solution novelty relative to stakeholder design objectives and goals, establish research questions and hypothesis and align project direction with stakeholder research interests from the outset.

ii) Manage Stakeholder Design Expectation

It is important that competing stakeholder deliverable expectations be actively managed. While systemic use of knowledge problem decomposition is a useful tool to progress research design artifacts it is understanding of the changing state of emergent artifacts to expedite shareholder exploitation that is the more significant challenge. Combining both is mandatory for stakeholder design expectation management. Process problem decomposition incrementally drives delivery of solution component artifacts but state change allows researchers recognize those suitable for research question answering and practioners, the opportunity for milestones assessment toward solution delivery. Within the project, problem class definition drove design requirements which in turn aided knowledge problem identified and addressing knowledge problems accommodated the introduction of researcher design objectives. Subsequent stakeholder agreement and management of the design artifact delivery ensured expectation was met.

iii) Early Consideration of Rigor

Contemplation of rigor should not be restricted to an obvious evaluation phase only. Any stage of the research process that advocates some solution or design activity should routinely be considered in terms of influence on, or suitability for, stakeholder success determination. Planned dual use of component artifacts delivers the environmental conditions for research question appraisal. Overall success determination is more complex, requiring the assembly of rigor criteria that is cognisant of environmental business objectives and research goals. Criteria shapes respective stakeholder evaluation and use of the practioner DSS artifact as the experimental platform caters for hypothesis evaluation. Adhering to this scheme ensured that the research process was successful in delivering the expected outcomes of both stakeholders.

iv) Stakeholder Engagement

Resolute stakeholder engagement from project inception is a prerequisite for project success. Close collaboration between practioners and researchers along the entire research process can be used to reconcile deliverable and therefore stakeholder expectation, provided that the business problem is clearly established and understood and business objectives are aligned with research goals. In particular stakeholder immersion in the design phase stages is an excellent opportunity for deliverable reconciliation. Active engagement was singularly the largest contributory factor to overall project planning and success through its formation of project technical direction and deliverable generation, based on solution design and knowledge base gaps.

6   Conclusions

This article describes methodology enhancements made to a practical DSR application [3] framework to actively support the level of practitioner and research engagement necessary to accommodate differing artifact expectations that arise in collabora-

tive research projects. Steps used in the overall design research methodology are discussed using a collaborative industrial-academic research project. The project focuses on knowledge translation and modeling intensive design activities that accompany natural language processing introduction and deployment. While further verification of the framework design and evaluation stages is necessary, we believe that the framework as construed, accompanying discussion on application and lessons learned can be applied to other projects seeking to enhance delivery of meaningful research output to industry and attracting academics to participate in industrial led projects.

References

1. Hevner, A.R., Salvatore T., March, S.T., Jinsoo Park, J., Sudha Ram, S., *Design science in information systems research.* MIS Q., 2004. 28(1): p. 75-105.
2. Hevner, A.R., *A Three Cycle View of Design Science Research.* Scandinavian Journal of Information Systems, 2007. 19(2): p. 87-94.
3. Meyer, M., Helfert, M., Donnellan, B., Kenneally, J., *Applying design science research for enterprise architecture business value assessments*, in *Proceedings of the 7th international conference on Design Science Research in Information Systems: advances in theory and practice.* 2012, Springer-Verlag: Las Vegas, NV. p. 108-121.
4. Osterle, H., Becker, J., Frank, U.,Hess, T., Karagiannis, D.,Krcmar, H., Loos, P., Mertens, P., Obrweis, A.,Sinz, Elmar J., *Memorandum on design-oriented information systems research.* Eur J Inf Syst, 2011. 20(1): p. 7-10.
5. Wieringa, R., *Design science as nested problem solving*, in *Proceedings of the 4th International Conference on Design Science Research in Information Systems and Technology.* 2009, ACM: Philadelphia, Pennsylvania.
6. Sheng, Y.P.M., P.P. and Litecky, C.R., *Competitor Analysis and Its Defences in the EMarketplace.* . Communications of the ACM, 2005. 48(8): p. 107-112.
7. Sauter, V.L., Free, D. *Competitive intelligence systems: qualitative DSS for strategic decision making," in SIGMIS Database.* 2005. NY, USA.: ACM.
8. Grant, G.H.a.C., S.J., *EDGAR Extraction system: An automated Approach to Analyze Employee Stock Option Disclosures Journal of Information Systems*, 2006. 20(2): p. 119-142.
9. Zahra, S.A., Chaples, S.S., *Blind Spots in Competitive Analysis.* The Academy of Management Executive (1993-2005) 1993. 7(2): p. 7-28
10. O'Riain, S., Spyns, P. *Enhancing the Business Analysis Function with Semantics*, *in Proceedings of On the Move to Meaningful Internet Systems*

(OTM) 2006: CoopIS, DOA, GADA, and ODBASE. 2006, Montpellier, France, p 818-835, Springer-Verlag.

11. Cunningham, H., *Information Extraction, Automatic.* Encyclopedia of Language and Linguistics, 2006: p. 665-677.

12. Markus, M.L., Majchrzak, A., Gasser, L., *A design theory for systems that support emergent knowledge processes.* MIS Q., 2002. 26(3): p. 179-212.

13. Appelt, D.E., Israel, D. J.,. *Introduction to Information Extraction Technology.* in *International Joint Conference on Artificial Intelligence.* 1999.

14. Zhang, D., Zhou, L., *Discovering golden nuggets: data mining in financial application.* Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on, 2004. 34(4): p. 513-522.

15. Peffers, K., Rothenberger, M., Kuechler, B., Lessard, L., Yu, Eric, *Using Design Science Research to Develop a Modeling Technique for Service Design*, in *Design Science Research in Information Systems. Advances in Theory and Practice.* 2012, Springer Berlin Heidelberg. p. 66-77.

16. Brocke, v.J., Seidel, S., *Environmental sustainability in design science research: direct and indirect effects of design artifacts*, in *Proceedings of the 7th international conference on Design Science Research in Information Systems: advances in theory and practice.* 2012, Springer-Verlag.

17. Li-Yen, S., Ching-Wen, C., Shiue, W., *The development of an ontology-based expert system for corporate financial rating.* Journal of Expert Systems with Applications, 2009. 36(2): p. 2130-2142.

18. Jarrar, M., Meersman, R. *Formal Ontology Engineering in the DOGMA Approach.* in *On the Move to Meaningful Internet Systems 2002: Confederated International Conferences CoopIS, DOA, and ODBASE Proceedings.* 2002: Springer.

19. Spyns, P., Tang, Y., Meersman, R., *A model theory inspired collaborative ontology engineering methodology.* Journal of Applied Ontology, 2007. 4(1).

20. Maass, W., Janzen, S., *Towards design engineering of ubiquitous information systems*, in *Proceedings of the 7th international conference on Design Science Research in Information Systems: advances in theory and practice*, K. Peffers, M. Rothenberger, and B. Kuechler, Editors. 2012, Springer-Verlag: (Eds.). Las Vegas, NV.

21. Baskerville, R., *What design science is not.* European Journal of Information Systems, 2008. 17(5): p. 441-443.

22. Livari, J., Venable, J., *Action research and design science research-seemingly similar but decisively dissimilar*, in *17th European Conference on Information Systems.* 2009.

23. O'Riain, S., Curry, E., Pinsker, R. *Competitive Analysis of Business Filings Using Ontologies and Linguistic Analysis*, In *Business Intelligence Congress 3: Driving Innovation through Big Data Analytics,* Orlando, Florida, 2012.

24. O'Riain, S., Curry, E., Harth., A., *XBRL and Open Data for Global Financial Ecosystems: A Linked Data Approach.* In *International Journal of Accounting Information Systems*, 2012, 13(2), p. 141-162