Different Approaches to in-house Identity Management
Justification of an Assumption

L. Fuchs, C. Broser, and G. Pernul

Abstract—**The use of roles in Identity Management Infrastructures (IdMI) has proven to be a solution for reorganising and securing access structures of employees. The definition of enterprise-wide roles is one of the most challenging and costly tasks during role development projects. It needs to be carried out on the basis of a predefined Role Development Methodology (RDM). In this paper we present existing methodologies and show their respective pros and cons. Lately some researchers have informally stated that hybrid role development is the most promising way to define roles, however, there hasn't been given a well-defined justification for this decision. The main contribution of this paper is hence the deduction of evaluation criteria based on information gathered from literature, practical experiences, and shortcomings of existing role development approaches. The evaluation criteria form the basis for a comparison framework verifying the assumption that hybrid RDMs are superior to Role Engineering and Role Mining methodologies.**

Index Terms— **Identity Management, Role Development, Role Engineering, Role Mining, Information Security**

I. Introduction

In today's increasingly open business environment companies provide access to resources to a greater number of users, and more heterogeneous types of users, than ever before. As a result of improper account management users accumulate a number of excessive rights over time, violating the principle of the least privilege [1]. This situation results in a so called identity chaos. Popular studies [2] show that major security problems arise because of employees gaining unauthorised access to resources as a result of manually handling user accounts. In-house Identity Management (IdM) has become a means to solve the aforementioned identity chaos. It is concerned with the storage, administration, and usage of digital identities during their lifecycle in the organisation. Roles acting as intermediary between employees and their access rights are an essential element of IdM. By defining business roles companies migrate from identity-based towards a role-based user and access management. This allows them to ease and secure provisioning processes, i.e. the allocation of digital and non-digital assets to employees, and access to resources in their IdM Infrastructure (IdMI) [3]. However, according to a NIST study [4] the most expensive challenge companies face before they achieve the benefits of role usage is the preliminary definition of valid roles. Some companies deal with this issue by installing resource-intensive procedures based on organisational and operational structures. Others create roles using data mining tools that analyse and cluster existing user permissions providing a high degree of automation. However, human support is needed in an extensive manner as the models fail to provide a detailed and structured guideline for discovering roles.

The main goal of this paper is to present and classify existing RDMs and compare them on basis of a well-defined set of evaluation criteria. The definition and classification of the evaluation criteria is based on a literature survey, several studies, and practical experiences. They provide the basis for a comparison framework and therefore allow for a classification of methodologies in order to show strengths and weaknesses of different Role Development Methodologies.

This paper is structured as follows. In section 2 we give insight into existing role development approaches. Section 3 classifies the existing methodologies and shows their respective shortcomings. Subsequently, section 4 defines our catalogue of evaluation criteria and compares existing RDMs. Final conclusions are given in section 5.

II. Related Work

The initial role definition is the central challenge enterprises face after having decided to implement roles in their IT infrastructure. Several approaches have been published to address this problem since the upcoming of the original RBAC model [5] in 1996. The importance of the definition of a valid role catalogue was first mentioned by Edward Coyne [6]. In the following years Thomsen et al [7] facilitated a layered approach to derive roles for access control purposes in their framework "Role Based Access Control Framework for Network Enterprises" (FNE). Detailed permission sets are grouped into related sets which may in turn be incorporated into still larger sets. Rights-bundles are in general aggregated

Manuscript received October 15, 2008.

Ludwig Fuchs is with the Department of Information Systems, University of Regensburg, Germany, (e-mail: Ludwig.Fuchs@wiwi.uni-regensburg.de).

Christian Broser is with the Department of Information Systems, University of Regensburg, Germany, (e-mail: Christian.Broser@wiwi.uni-regensburg.de).

Prof. Dr. Günther Pernul is head of the Department of Information Systems, University of Regensburg, Germany, (e-mail: Guenther.Pernul@wiwi.uni-regensburg.de).

and assigned to corresponding layers. In order to ease the role development Epstein and Sandhu [8] facilitated the usage of the UML language for modelling application- and enterprise key chains or constraints on basis of Thomsen's FNE. In the following Epstein and Sandhu extended the original RBAC model by integrating jobs, workpatterns, and tasks into the role development process [9]. Their approach supports the definition of the jobs of single roles and the consecutive decomposition into tasks and related permissions. It also allows the aggregation of permissions as building blocks into roles according to predefined workpatterns. Similar to Epstein, Neumann and Strembeck facilitate an aggregation approach [10]. However, their RDM is based on scenarios as the main input information for defining roles. Crook et al. [11] underline the need for integrating business structures into role development. They showed that using organisational structure to define roles has significant advantages by providing a clear focus for analysts and users eliciting requirements. Roeckle et al.'s approach [12] on the contrary integrates business processes into role development.

After various methodologies considering business structures have been proposed in literature, researchers identified the need for automatically integrating existing access right structures and identity information into the role development process. They argue that organisations cannot construct a role model from scratch but rather have to consider their existing IT and identity infrastructure. Kuhlmann et al. [13] were the first to link the role development process with data mining technologies. These approaches deal with the automatic extraction of patterns and other statistically significant structures from input data sets. Kuhlmann et al. present an iterative method for defining roles on basis of an existing database storing cross-platform access rights. Those rights are investigated and clustered into single roles using a clustering technique that is closely related to the k-means algorithm. In [14], Schlegelmilch et al. facilitate agglomerative hierarchical clustering and discover roles by merging permissions appropriately. They moreover present the ORCA tool (now "getROLE") as a prototypical implementation for the visualisation and analysis of permission cluster hierarchies. Recently Vaidya et al. proposed the Role Mining Problem as well as "RoleMiner", an algorithm for automatic role creation based on subset enumeration ([15], [16]). In contrast to Vaidja et al. Colantonio et al. integrated cost and performance decisions into Role Mining ([17], [18]).

III. Role Development Approaches

Role Development has been a vital research area at least since 1996. Many different methodologies have been published to define roles for different purposes. In order to avoid misunderstandings we introduce a classification scheme for RDMs in this section. Based on the used input data, the general approach, and various techniques a three-step differentiation of existing RDMs is given in Figure 1.

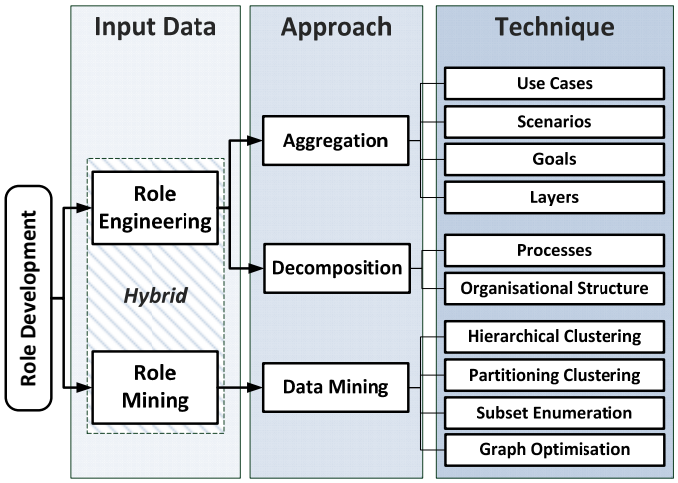

Figure 1: Classification of existing RDMs

In a first step RDMs can be categorised according to the input data they are based on: Role Engineering, Role Mining, and hybrid methodologies. Input data can be derived either from operational and organisational structures (OOS) or existing user account repositories within an organisation - or a combination of both. In the context of developing enterprise-wide roles Role Engineering is defined as the approach to create roles based on input information from the business perspective representing operational and organisational structures. Role Mining on the contrary is the tool-based approach discovering roles on basis of existing access rights and identity information using, e.g., data mining technology. Role Development is defined as the umbrella term for Role Engineering, Role Mining, and hybrid combinations of both.

Note that there has not been a standardised view on the terms *Top-Down*, *Bottom-Up*, *Role Engineering*, and *Role Mining* in literature, leading to various interpretations. Shin et al. [19] as well as Epstein and Sandhu [9] e.g. coin the terms Top-Down and Bottom-Up according to the used decomposition (Top-Down) or aggregation approach (Bottom-Up). In regards to system-independent roles and IdM, the term Bottom-Up is related to Role Mining, i.e. tool-based discovering of roles in existing access right structures [13].

A. Role Engineering

Role Engineering as defined above is considered as the theoretical way of developing roles. It can follow an aggregation or decomposition approach. The latter defines roles and decomposes them into permissions needed while aggregation works the opposite way [9]. Both approaches offer the chance to define a role catalogue that is closely aligned to the business perspective within an enterprise. Organisational and operational structures represent the basic input information sources for this approach.

1) Decomposition approaches

Role Engineering following the decomposition approach involves an in-depth analysis of business processes, functional structures and existing organisational charts in order to break down these elements to system-specific features needed to

fulfil certain tasks. Representatives of this class of approaches ([11], [12]) relate the definition and usage of roles to organisational theory and distinguish between organisational and functional roles. Decomposition approaches are facilitated for defining enterprise-wide roles that are used for automatic provisioning processes in IdM infrastructures. Using existing business structures as input information they are able to define so called structural roles [20]. Such system-independent roles do rather grant or deny access to an application in terms of a user account than dealing with the local access control within the target applications.

2) Aggregation approaches

Aggregation approaches are mostly adopted in the process of developing an application-specific role model. They are based on use case- or scenario descriptions, goals, or other input information. Members of this class of approaches like [7] or [10] use this information to define the way of interaction with an application and the bundles of permissions needed to fulfil certain tasks within this application. In order to streamline the mainly manual process, Strembeck et al. presented a tool-based approach for the definition of scenarios [21] and the automatic extraction of RBAC-models from BPEL4WS processes [22]. As it is impossible to maintain all processes, use cases, or scenarios defined within bigger companies at a certain point of time, the scope of aggregation approaches is limited, making them hardly applicable for the development of enterprise-wide roles in IdMIs [23].

Shortcomings

Role Engineering significantly depends on human factors and the amount and quality of input information available. Above all in settings where the quality of organisational charts, job descriptions, or position definitions is high, Role Engineering is a promising approach to find role candidates. However, on the other hand it is primarily a manual task involving extensive communication between stakeholders [14]. Only decomposition approaches are feasible for developing system-independent roles. Aggregating single elements like tasks comprehensively into roles is not applicable in an enterprise-wide project as most approaches are lacking any tool support. With dozens of business processes, thousands of users, and millions of authorisations, this is seemingly a difficult task. Besides the high complexity and costs ([12], [16]) the collection and preparation of input information are the main drawbacks [14]. Additionally, Role Engineering often creates a theoretical role catalogue as result of neglecting the actual access rights structures within an enterprise.

B. Role Mining

As a result of the presented shortcomings of Role Engineering, Role Mining has evolved as the pragmatic approach to rapidly define roles. It is gaining importance among the research community with specific focus on the usage of data mining technology to streamline role development. It focuses on the definition of system-independent roles that are, amongst others, used in IdMIs for user management, provisioning, and compliance purposes. Role Mining aims at automating role development steps by using tools to identify potential roles which can then be examined to determine whether they are appropriate given existing functions and business processes. In contrast to Role Engineering, Role Mining is based on the assumption that the actual roles already exist within the organisations' IT infrastructure [14]. Existing permission assignments are aggregated to define roles using data mining technologies like hierarchical or partitioning clustering algorithms, unsupervised learning methods, or graph-based approaches.

Shortcomings

Even though providing a high degree of automation, Role Mining has several serious unaddressed drawbacks: If the input quality is erroneous the role candidates discovered are also incorrect. All existing approaches assume that cleansing already took place before the role definition starts. We argue that this drawback needs to be addressed by introducing a mandatory customisable data cleansing and -preparation phase as shown in [24]. A performance analysis of several published Role Mining algorithms moreover revealed performance and quality issues arising from working with large input datasets that comprise several 100 permissions and users [25]. Analysing those datasets Role Mining algorithms tend to discover a large number of candidate roles and need to be parameterised with a minimum number of role members. However, the found role candidates only can be seen as preliminary results. Most approaches present algorithms for finding the optimal role set without taking into consideration that the business needs have to be involved in a role development project. Nearly all Role Mining publications explicitly mention the need of a combinatory use together with Role Engineering techniques.

C. Hybrid Approaches

The presentation of Role Mining approaches in the previous section has on the one hand underlined the need for tool-support and process automation in order to streamline role development. Nevertheless it has been pointed out that solely adopting Role Mining has several drawbacks. On the other hand consideration of business functions and organisational structure as carried out by Role Engineering is also mandatory for the definition of an initial role catalogue. A hybrid combination of the aforementioned methodologies offers the chance to define an improved role catalogue by minimising the constraints of the respective approaches. However, hybrid RDM needs to offer more than just the combination of Role Engineering and Mining techniques. It needs to define interfaces and structure the overall process of role creation integrating existing operational and organisational structures, access rights, and identity information. Various side effects like cleansed identity information additionally provides benefits for organisations applying a hybrid RDM.

While several authors superficially mentioned the need, Fuchs et al. recently proposed a detailed hybrid RDM called *HyDRo* (Hybrid Development of Roles) [25]. Their approach defines enterprise-wide roles used in IdMIs. The underlying philosophy is to perform joint Role Mining/Engineering and integrate OOS layer representatives (managers, executives, CIO) as frequently as necessary but as infrequently as possible.

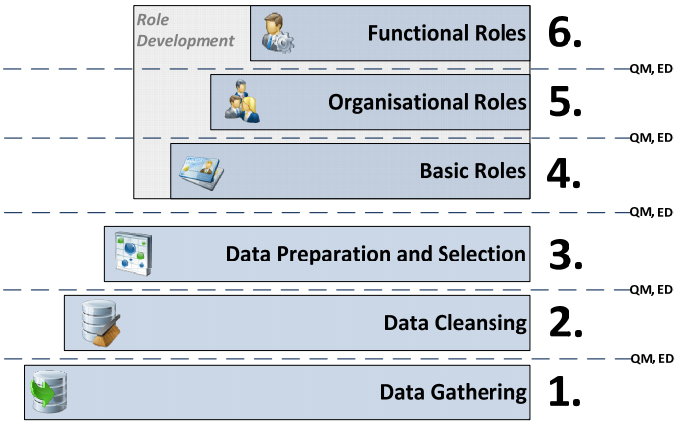

Figure 2: The HyDRo phases

Figure 2 gives an overview of the different phases of the role development process according to HyDRo: The methodology starts with the import of available input (Data Gathering) and consecutive Data Cleansing in order to improve the data quality. Defining suitable roles requires a classification and selection of the used input data in a separate phase (Data Preparation and Selection). The role development process itself is split into three phases, namely the definition of Basic-, Organisational-, and Functional Roles. The need for those role types within IdMIs has been shown in [23]. The integration of Role Engineering and Role Mining elements is carried out by hybrid information flows: Data mining technologies, adopted graph-based approaches, and clustering algorithms are used to reveal initial clusters of access rights and users according to a given organisational structure. Additionally existing identity information, job descriptions, etc. are integrated into the role definition process. The results are prepared, visualised, and sent for approval. After valid roles have been defined iteratively, a result backflow to IT systems like the HR repository ensures the propagation and usage of the defined roles within the IT infrastructure.

IV. Evaluation

Most authors from the Role Engineering- as well as Role Mining sector have agreed that hybrid role development offers the chance to minimise the failure risk. However, there hasn't been given a well-defined justification for this decision in terms of a comparison of existing RDMs on basis of a well-defined evaluation framework. In this section we thus deduce and classify a number of evaluation criteria based on information gathered from literature research, practical experiences, and shortcomings of existing models. Literature analysis on the one hand led us to the combination of positive (proven valuable approaches) as well as negative (shortcomings) research findings from the various publications examined. On the other hand practical needs gathered from industry partners as well as close relations to vendors of role development tools round off the criteria catalogue. By providing this evaluation framework for role development methodologies we allow for a comparison of existing approaches and the integration of future approaches concerning their focus, strengths, and weaknesses. We are aware that this framework is not exhaustive. However, the criteria given form the most important basic evaluation criteria RDMs have to meet. In a nutshell our research activities resulted in the definition of three groups of evaluation criteria:

- **Methodological**

Role projects should be carried out on the basis of an underlying methodology. The field of method engineering provides mandatory elements a methodology must encompany, like a detailed procedure model or an overview of the involved parties. These elements are used as the methodological evaluation criteria for a hybrid RDM. If a certain approach does not define the needed elements it cannot be regarded as a comprehensive RDM.

- **Domain-specific**

Domain-specific evaluation criteria represent mandatory steps and fundamental approaches supporting the role development process itself. Various researchers have addressed specific aspects of role development. The set of aspects together with related shortcomings presented in the previous section and experiences from practical projects like the need for cleansing input data before its usage rounds up this group of criteria.

- **Business-related**

Several failures of role development projects are due to missing adoption among the stakeholder [26]. Thus, companies' practical needs have to be considered during the process of role creation. By taking such business-related criteria into account during the construction process of a methodology the adoption within particular projects is fostered. Furthermore, risk mitigation by providing progress measurement or modularising of the role development process is expected additionally.

A. Methodological Evaluation Criteria

Method Engineering is the scientific basis of developing a methodology. In order to specify the structure of a method, Braun et al. and Gutzwiller ([27],[28]) analysed numerous contributions dealing with methods and method engineering. They describe a method by six fundamental elements, namely "activity", "specification document", "role", "technique", "tool" and "meta-model". Figure 3 shows the RDM method elements and their relations roughly corresponding to the findings of [27] and [28]. Some changes have been made due to the characteristics of the role development process. The

Procedure Model which can be divided into **Phases** and used as the central element of a RDM is introduced. Such a comprehensive guidance reduces complexity, helps to completely work off each necessary step, and enables reusing of already created results [24]. The **Results** entity represents the specification Documents which can be included into a comprehensive **Document Model**.

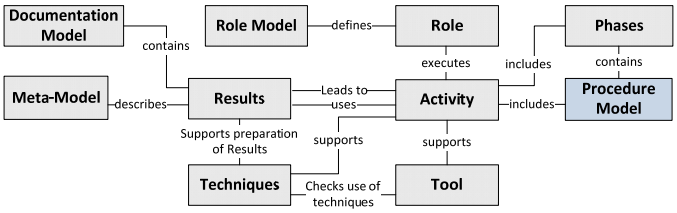

Figure 3: Method elements of a RDM

Procedure Models include each required **Activity** to develop roles, whereby the single **Activities** are brought in a specific order and grouped hierarchically in **Phases**. **Activities** may use output of preceding steps as input and generate defined **Results** which are recorded in predefined specification documents. The **Meta-Model** acts as the conceptual data model of the results, comprising the respective entities, relationships, and attributes. Several steps including different input data on the one hand as well as numerous stakeholders and respective tools on the other hand, increase the complexity of a RDM and affect its transparency in a negative way. Under these circumstances the users' acceptance suffers, leading to possible complications. In the run-up to the project all parties therefore have to be identified in order to assign them the respective **Roles** which carry out associated **Activities**. By means of the definition and application of **Techniques** in terms of detailed instructions which describe each process step extensively, involved stakeholders are supported to perform their tasks within the scope of their respective **Role**. **Tool** support throughout the entire process is indispensable in order to foster the performance of the project as well as the users' acceptance. These elements in conjunction with the **Procedure Model** were vital wishes of our business partners.

B. *Domain-specific Evaluation Criteria*

Domain-Specific evaluation criteria cover the functional demands of a RDM. As already shown in [24], role development requires different functionalities like data collection, -cleansing, -preparation, and -selection. Additionally the underlying methodology must be able to define various role types in an iterative and incremental development process.

1) *Data Gathering*

After the decision to start a role project, an organisation has to identify required input information and its various sources. Needed input will be chosen dependant on the availability and quality of existing documents and user information as well as temporal, personnel, and financial aspects. Subsequently the data has to be imported and checked to ensure that it is of an appropriate quality. Compiling and pre-processing input information is a major step during the data gathering phase. [24] shows that the RDM itself might have to be adjusted if not enough input information is available. The result of the data gathering phase is a repository of input information which can then be used within the role development process.

2) *Data Cleansing*

The outcome of the RDM, i.e. the resulting roles, highly depends on the quality of the underlying input information. Unrecognised failures regarding syntax as well as semantics will be propagated throughout the whole process and will be reflected in the results [29]. Due to this fact, different steps and techniques in order to check and improve the data quality in an iterative refinement process are essential. For this purpose tool support is mandatory because the amount of underlying data records exceeds human capacities. However, a check by responsible stakeholders is still indispensable. Automated tools may solely highlight conspicuous data records which have to be rechecked manually.

3) *Data Preparation and Selection*

As mentioned beforehand, the quality of input information needs to be checked and improved using data cleansing techniques. However, high quality is by no means a guarantee for successful role definition. Input data must be prepared and thereby categorised according to their applicability for the development process by several measures like statistical clustering or classification. Based on this categorisation suitable input data must be selected. For instance, permissions which are only assigned to a very small number of employees might be unsuitable for role development.

4) *Different Role Types*

The complexity of the role development process is reduced by the introduction of various role types [23]. Depending on job positions, functional structures, and basic tasks, several role types need to be created. Thereby organisational basic rights representing general tasks, organisational roles corresponding with positions, as well as functional roles matching with task bundles of employees can be defined accordingly. The classification of different role types forms the foundation of an incremental role development approach. According to their focus, companies need to be flexibly able to decide which role types they want to implement and at which stage they want to abort the role development process.

5) *Iterative Approach*

Because of the complexity of the role definition process in large organisations it is impossible to achieve a satisfying role catalogue within one development cycle. Data needs to be gathered, cleansed, and pre-processed. Roles must be defined iteratively whereby different underlying data and documents are used and diverse stakeholders get involved. Thus a continuous refinement of the results is guaranteed. Obviously, the results of the iterations have to be measured in order to indicate the completion of one specific iteration step.

6) Incremental Role Development

As aforementioned the development of different role types needs to be implemented as an incremental process, i.e. the various types are derived successively. Companies start with straightforward identifiable roles which require only little effort and input data. Thereby they already gain a lot of experience which assists them during the following, more complex steps. In order to create more complex role types, increased involvement of responsible business managers as well as more input data and manual effort is necessary.

C. Business-Related Evaluation Criteria

Implementing roles carries big potential but also involves risks for organisations due to the complex, longsome, and costly process which may not lead to a satisfying result, i.e. to a usable role catalogue. In the course of discussions with our user partners some important business-related criteria of a successful RDM have been identified and are introduced in the following. We admit that the given list of criteria is not comprehensive and represents only the main demands of organisations in respect to role projects. Other business-related issues still need to be investigated and integrated in our evaluation framework.

1) Modularisation

Modularisation has its origins in Software Engineering and denotes the characteristic that a system is structured in single components representing, for instance, functions or subroutines. Thereby the complexity is controllable and the maintenance processes of the system are eased. Accordingly the RDM needs to be structured unitised in order to gain more flexibility and consequently more acceptance by organisations. Dividing the development process in phases and sub-phases could be a promising approach to achieve modularisation. Thus, additionally, the overall complexity will be reduced and the planning and execution of the role development process will be simplified.

2) Partial Results

Throughout discussions with business partners it came out that failure of role projects leads to a significant loss of money without any outcome. Therefore, a RDM must provide the possibility to achieve at least partial results in order to prevent organisations going away empty-handed. The aforementioned modularisation can be a promising approach to fulfil this demand. Partial results can be provided in the single modules. Thus, organisations will be enabled to abort the development process in case of likely failure. Nevertheless the already achieved results during the preceding process steps are kept. Appropriate measurement of the result quality assists the enterprise throughout this decision process. An example for a partial result that can be used in various other scenarios is cleansed identity information and access rights. In a situation where the user information is of insufficient data quality, the company might abort the role development process and initially resolve data quality issues.

3) Situational Adaptability

The application of a fixed methodology in different situations [30], without any adjusting options is nearly impossible. This general statement also refers to role development. Thus a modified version of the RDM has to be derived against the background of diverse preconditions in different organisations like different company structures and sizes. One can expect that the available documents and the quality of identity information as well as temporal and personnel preconditions will differ strongly. Several process parts or phases of the RDM might need to be restructured and carried out differently.

D. Comparison

After having introduced the different criteria we are in the following retrospectively evaluating existing RDMs accordingly. Figure 4 provides an overview of the results ex post underlining the superiority of hybrid role development compared to its non-hybrid counterpart. To the best of our knowledge, no such in-depth justification has been given in literature. For appropriate visualisation we use a pie chart diagram where a completely filled out circle represents complete fulfilment and a three-quarter filled out circle is equivalent to an intensive treatment of the respective criteria. In contrast, a half filled one represents a partial performance and quarter filled out circles just a rudimentary conformance, e.g. an author only mentioning a certain aspect of role development. An empty circle expresses that the criteria has not been taken into account at all. In the course of our research we recognised that many authors published various papers for one role development approach. We decided to group those related contributions according to the main technique (see Figure 1) used. Note that only the main publications of the respective authors are cited.

It can be seen that non-hybrid RDMs fail to fulfil the established criteria framework to a great extent. Combinations of Role Engineering techniques used e.g. by Shin et al. [19] are superior to methodologies that facilitate just one single technique. However, note that [19] only refer to the original publications without presenting difficulties that arise during their combination. The assessment of Shin et al.'s approach is nevertheless influenced by the evaluation of [10] and [12].

Looking at Figure 4 it can be recognised that the business-related evaluation criteria apparently play a subordinate role in all non-hybrid RDMs. None of the approaches provides the possibility to create partial results or options to adapt to the respective situation within an organisation. In comparison, hybrid role development according to HyDRo puts special emphasis on meeting these aspects by modularising the process of role creation and explicitly defining partial results after each phase.

Besides those drawbacks our analysis reveals that meeting fundamental methodological criteria is not the main intention of the analysed non-hybrid approaches. They itemise single activities and techniques during the role development process without structuring and integrating them in a detailed and

Technique		Criteria:	Methodological						Domain-specific						Business		
			Document Model	Meta Model	Role Model	Techniques	Tools	Procedure Model	Data Gathering	Data Cleansing	Data Preparation / Selection	Different Role Types	Iterative Approach	Incremental Approach	Modularisation	Partial Results	Situational Adaptability
Role Engineering	Use Cases	Fernandez/Hawkins [32]	○	◔	○	◔	◔	◔	◔	○	○	○	○	○	○	○	○
		Poniszewska-Maranda [33]	○	◑	○	◔	◔	◔	○	○	○	○	○	○	○	○	○
	Scenarios	Neumann/Strembeck [10]	◑	◑	○	◕	◕	◕	◔	○	○	◔	◑	◑	◔	○	○
	Goals	He [34]	○	◑	○	◔	◔	◑	◑	○	○	○	◔	◔	◔	○	○
	Layers	Thomsen et al. [7]	○	◑	◔	◔	◑	◑	◔	○	○	○	○	○	◑	○	○
	Processes	Röckle et al. [12]	○	●	◑	◔	◑	◑	◔	○	○	◕	○	◔	◔	◔	○
		Chandramouli [35]	○	◔	○	◑	○	◑	◔	○	○	○	○	○	○	○	○
	Organisational Structure	Crook et al. [11]	○	◑	○	◔	○	◔	○	○	○	●	○	○	○	○	○
		Seufert [36]	○	●	○	◑	○	◑	◔	○	○	◑	○	○	○	○	○
	Org. structure / Processes	Kern et al. [37]	○	●	◔	◔	◑	◑	◑	○	○	◑	◕	◕	◑	○	○
	Org. structure / Processes / Scenarios [1]	Shin et al. [19]	◑	●	◔	◑	◔	◑	◔	○	○	◕	◑	◑	◔	○	○
Role Mining	Hierarchical Clustering	Schlegelmilch/Steffens [14]	○	◔	○	◑	◕	◔	◔	◔	○	○	◑	◔	○	○	○
	Graph Optimisation	Zhang et al. [38]	○	◔	○	◕	◔	◔	◑	○	○	○	◑	◔	○	○	○
	Subset Enumeration	Vaidya et al. [16]	○	◔	○	◕	◑	◔	◑	○	○	○	◑	◔	○	○	○
		Colantonio et al. [17]	○	◔	○	◕	◔	◔	◔	○	○	○	◑	◔	○	○	○
	Partitioning Clustering	Kuhlmann et al. [13]	○	●	○	◔	◑	◑	◑	◔	◔	◔	◕	◔	◔	○	○
Hybrid	Hybrid RDM	Fuchs/Pernul [25]	◑	●	◕	◕	◕	●	◑	◕	◕	●	●	●	◕	◕	◔

[1] Shin et al. facilitate a combination of Role Engineering based on Organisational Structures, Processes according to [12], and Scenarios according to [10].

Figure 4: Comparison of selected role development approaches against the criteria using a pie chart diagram evaluation

applicable procedure model [24]. They also lack a clearly defined documentation- and role model usable for standardising communication between the involved parties. HyDRo, in contrast, provides required method elements like a detailed procedure model divided into phases and activities. Similar to our argumentation, Wortmann [31] states that existing non-hybrid RDMs fail to give detailed procedure models for the integration of role-based authorisation. Another related drawback of many existing non-hybrid approaches is the lacking tool support. HyDRo overcomes this by being supported by the *contROLE* tool [25] throughout the complete role development process.

Looking at the domain-specific criteria one can see that none of the Role Engineering approaches considers the input data quality and consequently no data cleansing measures are supported. Only selected Role Mining publications touch this issue on the brink. The same goes for data preparation and –selection in role development projects. Existing non-hybrid RDMs provide none or hardly any guidance in this area, even though cleansing or preparing input data are essential elements to develop an optimal role catalogue.

These couple of examples underline just a number of the shortcomings given in Figure 4. However, we admit some limitations of this evaluation: RDMs with more than one publication (e.g. Neumann et al. [10]) outscored others as a result of more available information. This complicates an objective comparative evaluation. Additionally, the goal of several approaches is to address a specific aspect of role development. Hence, they do not put emphasis on any business-related or methodological aspects. Moreover some surveyed RDMs are provided by consultants (like [12] or [13]) whose focus doesn't necessarily correspond with a researcher's one. As we have to rely exclusively on the given information in the academic publications, we assume that solutions to some identified shortcomings are offered directly by the consultants but are not accessible to the public in form of research publications.

V. Conclusion

In this paper we have motivated the importance of roles within Identity Management Infrastructures of modern organisations. We presented and classified existing methodologies for role development showing their respective pros and cons. The main contribution of this paper is the deduction of an evaluation criteria framework for role development methodologies and the consecutive comparison of existing RDMs on basis of the framework. Several researchers have informally stated that the hybrid combination of Role

Engineering and Role Mining is the most promising way for creating enterprise-wide roles. However, they have not given any justification for this assumption. In this paper, to the best of our knowledge, this justification is given for the first time on basis of a well-defined evaluation framework. The findings have shown that hybrid role development on the one hand is superior to non-hybrid role development and on the other hand offers more than just a combination of two already existing approaches for role discovery.

REFERENCES

[1] D. F. Ferraiolo, R. D. Kuhn, and R. Chandramouli, Role-Based Access Control. Boston, MA: Artech House, 2007.

[2] G. Dhillon, "Violation of Safeguards by Trusted Personnel and Understanding Related Information Security Concerns", Computers & Security, vol. 20, no. 2, 2001, pp. 165-172.

[3] L. Fuchs and G. Pernul, "Supporting Compliant and Secure User Handling – a Structured Approach for In-house Identity Management," in Proc. 2nd International Conference on Availability, Reliability and Security (ARES'07), Washington, 2007, pp. 374-384.

[4] M. P. Gallaher, A. C. O'Connor, and B. Kropp, "The economic impact of role-based access control", 2002. Available: http://www.nist.gov/director/prog-ofc/report02-1.pdf

[5] R. S. Sandhu, E. J. Coyne, H. L. Feinstein, and C. E. Youman, "Role-Based Access Control Models," IEEE Computer, vol. 29, no. 2, 1996, pp. 39-47.

[6] E. J. Coyne, "Role Engineering," in Proc. 1st ACM Workshop on Role-based access control, article no. 4, 1996.

[7] D. Thomsen, D. O'Brien, and J. Bogle, "Role Based Access Control Framework for Network Enterprises," in Proc. 14th Annual Computer Security Applications Conference, 1998, pp. 50-58.

[8] P. Epstein and R. Sandhu, "Towards a UML based approach to role engineering," in Proc. 4th ACM workshop on Role-based access control, 1999, pp. 135-143.

[9] P. Epstein and R. Sandhu, "Engineering of Role/Permission Assignments," in Proc. of the 17th Annual Computer Security Applications Conference, 2001, pp. 127-136.

[10] G. Neumann and M. Strembeck, "A scenario-driven role engineering process for functional RBAC roles," in Proc. 7th ACM symposium on Access control models and technologies, 2002, pp. 33-42.

[11] R. Crook, D. Ince, and B. Nuseibeh, "Towards an Analytical Role Modelling Framework for Security Requirements," 2002. Available: http://mcs.open.ac.uk/ban25/papers/refsq02.pdf

[12] H. Roeckle, G. Schimpf, and R. Weidinger, "Process-oriented approach for role-finding to implement role-based security administration in a large industrial organization," in Proc. 5th ACM workshop on Role-based access control, 2000, pp. 103-110.

[13] M. Kuhlmann, D. Shohat, and G. Schimpf, "Role mining - revealing business roles for security administration using data mining technology," in Proc. 8th ACM Symposium on Access Control Models and Technologies, 2003, pp. 179-186.

[14] J. Schlegelmilch and U. Steffens, "Role mining with ORCA," in Proc. 10th ACM Symposium on Access Control Models and Technologies, 2005, pp. 168-176.

[15] J. Vaidya, V. Atluri, and Q. Guo, "The role mining problem: finding a minimal descriptive set of roles," in Proc. 12th ACM Symposium on Access Control Models and Technologies, 2007, pp. 175-184.

[16] J. Vaidya, V. Atluri, and J. Warner, "RoleMiner: mining roles using subset enumeration," in Proc. 13th ACM Conference on Computer and Communications Security, 2006, pp. 144-153.

[17] A. Colantonio, R. Di Pietro, and A. Ocello, "Leveraging Lattices to Improve Role Mining," in Proc. 23rd International Information Security Conference, 2008, pp. 333-347.

[18] A. Colantonio, R. Di Pietro, and A. Ocello, "A cost-driven approach to role engineering," in Proc. 2008 ACM Symposium on Applied Computing, 2008, pp. 2129-2136.

[19] D. Shin, G. Ahn, S. Cho, and S. Jin, "On modeling system-centric information for role engineering," in Proc. 8th ACM Symposium on Access Control Models and Technologies, 2003, pp. 169-178.

[20] E. Coyne and J. Davis *Role Engineering for Enterprise Security Management*. Boston, MA: Artech House, 2007.

[21] M. Strembeck, "A Role Engineering Tool for Role-Based Access Control," 2005. Available: http://wi.wu-wien.ac.at/home/mark/publications/sreis05.pdf

[22] J. Mendling, M. Strembeck, G. Stermsek, and G. Neumann, "An Approach to Extract RBAC Models from BPEL4WS Processes," in Proc. 13th IEEE International Workshops on Enabling Technologies: Infrastructures for Collaborative Enterprises (WETICE), 2004, pp. 81-86.

[23] L. Fuchs and A. Preis, "BusiROLE: A Model for Integrating Business Roles into Identity Management," in Proc. of the 5th International Conference on Trust, Privacy, and Security in Digital Business (TrustBus), Torino, Italy, 2008.

[24] L. Fuchs and G. Pernul, "proROLE: A Process-oriented Lifecycle Model for Role Systems", in Proc 16th European Conference on Information Systems, 2008.

[25] L. Fuchs and G. Pernul, "HyDro - Hybrid Development of Roles," in Proc. 4th International Conference on Information Systems Security (ICISS 2008), Hyderabad, India, 2008, to be published.

[26] L. A. Kappelman, R. McKeeman, and L. Zhang, "Early Warning signs of IT Project Failure: The Dominant Dozen," EDPACS, vol. 35 no. 1, 2007, pp. 1-10.

[27] C. Braun, F. Wortmann, M. Hafner, and R. Winter, "Method Construction – A Core Approach to Organizational Engineering," in Proc. 2005 ACM Symposium on Applied Computing, 2005, pp. 1295-1299.

[28] T. Gutzwiller, „Das CC RIM-Referenzmodell für den Entwurf von betrieblichen, transaktionsorientierten Informationssystemen." Heidelberg, Physica-Verlag, 1994.

[29] M. L. Lee, H. Lu, T. W. Ling, and Y. T. Ko, "Cleansing Data for Mining and Warehousing," in DEXA '99. LNCS, vol. 1677, 1999, pp. 751-760.

[30] J. Becker, C. Janiesch, and D. Pfeiffer, "Reuse Mechanisms in Situational Method Engineering," in Situational Method Engineering – Fundamentals and Experiences, Proc. IFIP WG 8.1 Working Conference, 2007, pp. 79-93.

[31] F. Wortmann, „Vorgehensmodelle für die rollenbasierte Autorisierung in heterogenen Systemlandschaften," Wirtschaftsinformatik, vol. 49, no. 6, 2007, pp. 439-447.

[32] E. B. Fernandez and J. C. Hawkins, "Determining role rights from use cases," in Proc. 2nd ACM workshop on Role-based access control, 1997, pp. 121-125.

[33] A. Poniszewska-Maranda, "Role engineering of information system using extended RBAC model," in Proc. 14th IEEE International Workshops on Enabling Technologies: Infrastructure for Collaborative Enterprise, 2005, pp. 154 – 159.

[34] Q. He, "A structured Role Engineering Process for Privacy-Aware RBAC Systems," 2003. Available: http://www.ihp-ffo.de/systems/lv/ws0506/Schutz_Privatsp%E4re/Role%20Based%20Access%20Models/Privacy_aware_RBAC_Modeling.pdf

[35] R. Chandramouli, "Business Process Driven Framework for defining an Access Control Service based on Roles and Rules," 2003. Available: http://csrc.nist.gov/nissc/2000/proceedings/papers/047.pdf

[36] S. E. Seufert, „ Der Entwurf strukturierter rollenbasierter Zugriffskontrollmodelle," Informatik-Forschung, vol. 17, no. 1, 2002, pp. 1-11.

[37] A. Kern, M. Kuhlmann, A. Schaad, and J. Moffett, "Observations on the role life-cycle in the context of enterprise security management," in Proc. 7th ACM Symposium on Access Control Models and Technologies, 2002, pp. 43-51.

[38] D. Zhang, K. Ramamohanarao, and T. Ebringer, "Role Engineering using graph optimization," in Proc. 12th ACM symposium on Access control models and technologies, 2007, pp. 139-144.