Abstract
Digitalization leads to increasing complexity and tight coupling of business processes. This increases the risk of "natural accidents". An answer to this proneness is a holistic mindfulness culture. The mindfulness dimension requires a holistic approach because it includes multiple aspects. This mindfulness should be reflected in the self-assessment of a company's employees. It is attempted to clarify whether social media allows an insight into how self-proclaimed qualifications are distributed in current enterprises. Social networks are particularly suitable because the members try to reflect the perceived demands of the companies while the characteristics presented are "quality assured" by connected members. More than 3500 profiles of senior IT managers were examined on the social network XING. COBIT has been used for categorization as a proven holistic approach for IT governance and structuring tool. The study shows that executives can be primarily categorized as generalists or specialists, both embracing a holistic perception, and therefore social media is suitable for general studies of qualifications in a business environment.

Keywords: Qualification Distribution; COBIT; Senior IT Management; Social Media.

1. Introduction
The influence of employee and company qualification on the control of business risks has been studied starting with the Natural Accident Theory [1] and building on "high reliability organizations" (HRO) [2-4] and mindfulness [5]. Empirical studies confirm these theories [6]. The practical significance from the perspective of IT governance can also be seen in the fact that in the widely accepted IT governance framework COBIT, "employee qualification" has its own classification dimension (People, Skills, and Competencies). It seems that companies' focus is on soft skills to handle uncertainty and problem-solving skills [7]. On the other hand, interest in formal and technical qualifications or systematic analysis seems low and is often driven by compliance. The distribution of qualifications seems to depend on factors such as the size of the company [8] and the innovativeness of the companies [9]. Scientific research indicates that general management skills are gaining in importance [10] and that specialised executives have little bargaining power in the laboratory market [11], and on the other hand, general qualifications facilitate job changes [12]. All that leads to an increase in interest in general qualifications on the employee side [13]. However, these previous surveys are either not performed in repetition due to effort, are only based on relatively small samples (n = 446 [7]; n = 318 [13]) or concentrate only on top-level executives [11]. The hypothesis of this work is that social media can give detailed insight into the distribution of qualifications at different organisational levels with an...
increased number of samples. It is expected that previous findings on qualification levels can be confirmed, thereby showing the relevance of this analysis method.

For identification of qualification profiles there are numerous approaches. In our analysis, we tried to build on an already common and time proven approach, and choose dimensions established there as a classification scheme. The COBIT framework was finally selected due to its worldwide use and holistic approach. Because of this it has been often applied by auditing firms [14, 15] and recommended from official sides [16, 17]. Other frameworks like IT Infrastructure Library framework (ITIL) which is more focused on IT operations does not offer classification schemes. On the other hand the framework of Committee of Sponsoring Organizations of the Treadway Commission (COSO) is too strategic and therefore not applicable for surveys on the micro (aka individual) level. By having links to ISO 9000, COBIT is even influenced by non-IT governance issues (Figure 1).

![Figure 1. COBIT's scope and coverage within the generic ecosystem of frameworks and standards](image)

COBIT defines seven areas of action (called “enabler”) which should be looked at if you want to have a holistic approach [18]. Their importance has been confirmed by security surveys, where IT security experts identified critical security factors which can be related to COBIT enablers [19].

1. Principles, Policies and Frameworks:
   - Concerned with compliance with external and internal regulations and standards or governance specifications i.e. laws and ethical standards.

2. Processes:
   - Concerned with structured and standardised approaches for accomplish targets.

3. Organisational Structures:
   - Concerned with roles, working organisation and principles, extend and manner of authority.

4. Culture, Ethics, Behaviour:
   - Concerned with individual and collective behaviour aka company culture.

5. Information:
   - Concerned with intrinsic quality, context sensitive quality, security and availability (security targets) of produced and used information.

6. Services, Infrastructure and Applications:
   - Concerned with technical resources and services used to accomplish the IT and business target.
7. People, Skills and Competencies:

Concerned with Skills and Competencies like education and qualification levels, technical skills, experience, knowledge and behaviour [18].

It can be seen that the COBIT standard does not only comprise technological aspects (as shown in item 6), but also considers so called “soft areas”, such as cultural factors (item 4) or qualification of employees (item 7) since they can play a vital role in solving acute crisis situations. By this means, the standard defines the requirements for a holistic view. The paper at hand investigates if standard IT classification schemas can be used to categorise qualification and attitudes of senior IT management from a holistic perspective taking into account the seven enablers of the COBIT standard. This is done by determining whether the self-proclaimed skills and competencies of IT executives actually reflect the entire spectrum of the COBIT specification. Therefore we conducted an analysis on the social media platform XING, in which the seven COBIT enablers are mapped to the user profiles of IT personnel having major management positions in German companies. We hypothesize that the usage of this data might give valuable insights into the perceptions of IT managers of what is important and expected by business partners and current or potential employers in terms of business continuity and crises mitigation.

An investigation on individual level doesn’t help for generalisations, therefore a cluster analysis has been performed with more than 3600 items.

2. Data-Analysis

2.1. Data-Collection

User profile data of IT executives were collected through the XING API. To remove potential language or cultural biases the feasibility survey was restricted to Germany. The idea of using XING as base for the survey was born because of the disadvantages of classical survey methods: First, an online survey has a low response return rate [20] and we do not know which kind of candidates might answer [21-23]. The return rate is especially low if the there was no previous contact or if there is no relationship with the candidates. A semi-structured survey can in practise only be used for small number of interviewees [24] and has a different focus [25, 26]. Performing interviews without an initial explorative survey bears the danger of biasing the results as the survey proceeds [27]. In all cases, retrieving a sufficient amount of subjects is essential for the understanding of the domain and a data source like XING can help to overcome the limitations of classical survey methods. However, subsequent to our analysis, the results can be verified using one of the above methods.

XING [28] is a social network service for professionals. Registered members can befriend business associates and list contact data, professional experiences, work stations, qualifications and skills (aka ‘haves’) and wishes (aka ‘wants’). It is widespread in Germany with more than 10 million users and competes with the similar service LinkedIn. It is one of the major websites used by professional recruiters and companies [29, 30]. We assume that the candidates - by trying to present themselves in the best possible manner - reflect the perceived expectations of relevant stakeholders (i.e., business associates, recruiters and potential employers) with regards to which qualifications are required and useful in their current or potential future position. By doing so, we link the micro-level to the meso-level in terms of sociological science or the individual level to the team or company level in terms of organisational science. This differentiation is a standard approach in institutional or social research.

Whether the IT executives really possess these attributes, cannot be verified by us. However, a certain quality control can be assumed since the IT managers are also visible to colleges, business contacts and previous employers and may therefore shy away from overstating their skills. The XING platform provides an API for developers [31]. However, there are technical limitations. First, only 90 queries per day are allowed and, more severe, only a maximum of 100 hits per query are returned. Therefore, we had to split our queries into useful subqueries to retrieve the competencies of German IT executives. For this reason, we combined the actual search string with the name of each major cities (≥ 100.000 inhabitants) and medium-sized cities (≥ 20.000 inhabitants) in Germany and queried (in total 678 cities) [32]. This resulted in the following search strings:

Search string 1: “Leiter IT AND [city name]” (translated in English “Manager IT”) and
Search string 2: “CIO AND [city name]”

In total we pulled 19.587 users over a period of two months (November / December 2017).

2.2. Data Preparation

We anonymised, normalised and stored the data in an SQL database upfront of the data analysis. The attribute objects employment_status, professional_experience, business_address, educational_background, birth_date (and contained fields) offer predefined list values and are well-structured. The haves category is the most important feature of our analysis. Here, IT executives state their competencies, skills and qualifications in a keyword-based text form.
Therefore, the *haves* feature needed to be transformed with the help of natural language processing (NLP) operations. We realised that not only German words were used, but British and American English keywords were quite frequently stated as well alongside, with occasional spelling errors. We decided to deal with this irregularities manually at a later phase (2.3 Results) and apply no stemming or lemmatisation at all. To increase the quality and to analyse only the core information we applied a public stop word list [33], which was iteratively amended by own words after visual reviews. The resulting *haves* were normalised and stored in the SQL database for further analysis.

For all users or samples we found 1730 different characteristics for the *haves* feature. However single characteristics had a high variance in their distribution. For the analysis characteristics with a count <10 have been neglected. Each of these characteristics have been classified and assigned to one or more COBIT enablers. We found, that most of the stated characteristics fall in one of the seven COBIT enablers, while some needed to be grouped into multiple enabler categories [18]. In cases where the keyword-based qualification statement spawned several COBIT enablers, it was assigned to all of them with the same weight (for each enabler a column was added that contained either “1” if applicable or “0” if not applicable).

### Table 1. Example of attributes and their categorisation

<table>
<thead>
<tr>
<th>Attributes</th>
<th>like_or_equal</th>
<th>Principles, Policies &amp; Frameworks</th>
<th>Processes</th>
<th>Organisational Structures</th>
<th>Culture, Ethics, Behaviour</th>
<th>Information</th>
<th>Services, Infrastructure &amp; Applications</th>
<th>People, Skills &amp; Competencies</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>.net</td>
<td>like</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>Programming Framework</td>
</tr>
<tr>
<td>agile</td>
<td>equal</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>agile transformation</td>
</tr>
<tr>
<td>transformation</td>
<td>equal</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>Business Continuity Management</td>
</tr>
<tr>
<td>bem</td>
<td>equal</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>People skills</td>
</tr>
<tr>
<td>begeistergsfähigkeit</td>
<td>like</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>enthusiasm</td>
</tr>
<tr>
<td>itil</td>
<td>like</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>framework</td>
</tr>
</tbody>
</table>

Since the classification has admittedly a certain subjectivity, three different individuals performed this task independently. Afterwards, we merged the results and discussed differences in detail.

After preparing the *category* table, these pattern were mapped to the individual values of the feature “*haves*” of each sample and thereby categorising the user qualifications by COBIT enablers. For the further analysis we had to deal with the different numbers of values for each sample. The count of values in a sample ranged from 4 to 20, some of the users provided only basic qualification, others went into more details. To make the characteristics comparable the values were normalised for each user (row oriented normalisation). That means that the sum of all of the seven enabler for a sample is always 1.

\[
X_{\text{enabler[user]}} = \frac{\sum_{\text{attributes}} \text{enabler}}{\sum_{\text{attributes}} \text{user}}
\]  

The result looked like:

### Table 2. Example of assigned and normalised characteristics

<table>
<thead>
<tr>
<th>user_id</th>
<th>Principles policies frameworks</th>
<th>Processes</th>
<th>Organisational structures</th>
<th>Culture ethics behaviour</th>
<th>Information</th>
<th>Services infrastructure applications</th>
<th>People skills competencies</th>
<th>sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>10002222_818bb0</td>
<td>0.1364</td>
<td>0</td>
<td>0.0909</td>
<td>0</td>
<td>0.0909</td>
<td>0.6364</td>
<td>0.0454</td>
<td>1</td>
</tr>
<tr>
<td>10003003_bd325a</td>
<td>0.1765</td>
<td>0.1765</td>
<td>0.2352</td>
<td>0.1765</td>
<td>0</td>
<td>0.0588</td>
<td>0.1765</td>
<td>1</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
</tbody>
</table>

To remove unnecessary influences furthermore we restricted the samples of the observation (aka users) by several conditions. First, we included only executives, i.e. people in charge of a business unit, since we were solely interested in candidates in a leading position. Furthermore, we considered only people who are working in companies with more than 50 people – due to the fact that below that the company’s organisational structures normally do not call for expressive hierarchies or division of labour in the IT department. That reduced the amount of people to 3680.

### Table 3. Filter conditions for survey

<table>
<thead>
<tr>
<th>Filter</th>
<th>Possible values</th>
<th>Included values</th>
</tr>
</thead>
<tbody>
<tr>
<td>employment_status</td>
<td>'employee', 'entrepreneur', 'executive', 'freelancer', 'public_servant', 'recruiter', 'retired', 'student'</td>
<td>'executive'</td>
</tr>
<tr>
<td>company_size</td>
<td>'1', '1-10', '10001+', '1001-5000', '11-50', '201-500', '5001-10000', '501-1000', '51-200', 'None'</td>
<td>'10001+', '1001-5000', '201-500', '5001-10000', '501-1000', '51-200',</td>
</tr>
</tbody>
</table>
3. Results

After having performed the pre-processing operations, we undertook an explorative analysis and applied a hierarchical cluster analysis algorithm on the data. As starting point we created a histogram using all samples aka candidates and investigated the distribution of qualifications over all COBIT enabler aka dimensions (Figure 2). Not surprisingly there is a relatively high proportion of counts in the technical dimension (enabler 6). All other bins are roughly equally distributed with each dimension being covered. Compared with the lower whisker we see a very high upper whisker and a considerable amount of outliers. For all bins we can see extreme outliers. That can further be confirmed by looking at the density chart - within the bins we can see left skewness and very long tails to the right (Figure 3).

![Figure 2. Histogram for all candidate characteristics with all COBIT dimensions](image1)

![Figure 3. Density chart for all candidates with all COBIT dimensions](image2)
Afterwards, a hierarchical cluster analysis (distance measure “ward”) was performed to see if additionally certain patterns are hidden in the data. In order to identify the best cluster solution we created a dendrogram (Figure 4) and the related elbow chart (Figure 5).

Figure 4. Dendrogram for all candidates for all COBIT dimensions

Figure 5. Elbow diagram for all candidates for all COBIT enablers
Both figures indicate that a good cluster solution might consist of 3 to 5 clusters. We decided that 4 clusters constitute the most significant and explainable cluster solution. With only 3 clusters we would miss the opposed slopes in the distribution over different company sizes as seen later in (Figure 6). In a 5 cluster solution however, as was revealed by a detailed investigation of the histograms, the first cluster (Cluster 1 in) would get divided on the 6th COBIT dimension (“Services, Infrastructure and Applications”) into a medium and an extreme shape - quite often an indicator of overfitting. For these reasons, we decided that the 4 cluster solution represents the underlying patterns in the data most accurately. After having performed the hierarchal cluster analysis we calculated the same histogram, as initially done for all candidates, for each cluster.

For ease of explanation, we gave all clusters “speaking names” and will subsequently refer to those. The largest cluster (named “generalist”, cluster 4) with 1620 out of 3680 samples shows a fairly homogeneous distribution over all dimensions (the median in all dimension not exceeding .21), demonstrating not only technical but a holistic range of skills. The second largest cluster (named “technically focused”, cluster 1) shows a strong concentration on technical skills with a median of .5 in the dimension “Service, Application and Infrastructure”. The number belonging to that cluster is 1284 out of 3680 samples. It might not be surprising that an IT Manager has certain technical knowledge, but for executive managers we did not expect such a large amount of samples with a strong focus in this area. With 570 of 3680 samples (named “socially focused”, cluster 2) quite surprisingly we found a type that emphasizes culture, ethics and behaviour (median of .39 in the dimension “Culture, Ethics and Behaviour”), but does not neglect organisational and technical skills at the same time. The smallest cluster contains only 204 of 3680 samples (named “organisationally focused”, cluster 3). Its members are very focused on organisational skills and the median there has the highest value of all medians (.57), even by far exceeding the technical skills median of the technically focused cluster.
The data do not allow insights into the depths of knowledge in the specified area. It can be only stated that the distribution in diverse areas is different. For the technical dimension it’s a fair assumption, that in smaller firms, employees need to cover a broader spectrum of technical skills than in larger firms. Furthermore, it can only be assumed that the concentration on less different products or technical skills leads to more expert knowledge in the specific technical area which in our investigation results in smaller numbers of characteristics. That might be the underlying reason of the broad inter quartile range (IQR) of the 6th enabler of cluster 1 (“technically focused”). In cluster 4 ("generalists"), besides a weak emphasis on the technical skills, we can see a second peak in the organisational skills, which goes well with the assumption of a “generalist” profile.

We wondered if the distribution of these clusters exhibits additional patterns. An additional interesting question arises, if the distribution of the clusters depend on the age or better the work experience of the samples. If different means could be identified, it might indicate an evolution from one cluster into another. Our first guess was that with growing work experience the members move from one cluster into another cluster. Therefore we grouped all clusters by work years of the members and compared the distribution. We might see a slight skewness to the right for the ‘organizationally’ focused cluster (cluster 3). Tus skewness could imply that membership in this cluster comes with certain seniority (Figure 7).

![Figure 7. Distribution of work year for each cluster](image)

We combined the analysis of the company size and work years feature, which indicates that with more work experience the member work in small companies and vice versa. However the population in each bin is not sufficient and for a detailed analysis additional data sources must be linked, which is out of scope of the paper at hand. Therefore the analysis neither can prove a strong concentration in any work experience bin for any cluster nor could we identify a clear trend. Therefore, based on the data at hand the clusters are independent from work years. A second idea was to investigate, if the cluster depends on certain regions, indicating local trends. Again - no concentration could be identified and the clusters can be regarded as region independent.

However, if we group the relative cluster distribution by company size (Figure 8), a clear trend becomes visible. The percentage of the ‘generalists’ cluster (from a share of 0.376 for small to 0.520 for major companies) as well as the ‘culture focused’ cluster (from a share 0.156 for small to 0.172 for major companies) increases for larger company sizes. At the same time the ‘technically focused’ clusters and the ‘organizationally focused’ clusters decrease (from 0.396 for small to 0.270 respective 0.071 for small to 0.038 for major companies) with more employees in an enterprise.
Figure 8. Distribution of clusters by company size

Looking at the slope (Figure 9) and the number of samples in a group, we can state that the bigger the firm the less technically focused people are found there. These seem to be replaced by generalists, but still, as later analysis will show, holistic minded people. Unfortunately, the data allow no insight on department size, but again it’s quite likely that teams are larger in companies with more employees. The ‘culture focused’ cluster is quite stable and shows only a small positive slope the larger the firms are. Overall, we can see a trend in senior management positions to have a more generalist and holistic profile for larger firms. Not so obvious on first sight is the cluster ‘organisationally focused’ – the slope is much smaller but - taking into account the percentage change – the ‘loss’ from the bin for the smallest firms to the bin for the largest firms is quite drastic – we see that the distribution shrinks by more than 46%.

To explore the reasons are beyond the scope of the paper and must be postponed.

Figure 9. Slope of regression lines of clusters by company size
After analyzing the primary features and identifying the clusters, evaluating secondary features we explored how the distribution of the distinct values for the primary feature (aka enablers) looked like within each cluster. In other words we counted the number of enablers of each candidate and aggregated the results for the whole cluster. By doing so, we investigated how holistic minded the samples are overall (Figure 10) and within each clusters (Figure 11).

<table>
<thead>
<tr>
<th>Total no.</th>
<th>1 enabler</th>
<th>2 enablers</th>
<th>3 enablers</th>
<th>4 enablers</th>
<th>5 enablers</th>
<th>6 enablers</th>
<th>7 enablers</th>
</tr>
</thead>
<tbody>
<tr>
<td>percentage</td>
<td>4.9</td>
<td>7.7</td>
<td>20.1</td>
<td>25.5</td>
<td>17.5</td>
<td>25.9</td>
<td>19.9</td>
</tr>
</tbody>
</table>

If we look at the overall representation, it could be reasoned that all samples are holistic minded with approx. 46% have occupied 6 or 7 enabler, and approx. 63% have occupied 5 to 7 enabler. If we divide the data by the identified clusters the picture changes. It is no surprise that the majority of the “generalists” (i.e., more than 50%) occupy between 6 and 7 enablers. More surprising is, that the “technical focused” cluster comes second with approx. 44% of the members that stated 6 or 7 enablers in their XING competence portfolio. This means that in spite of having a very
strong technical focus, they still tend to demonstrate qualifications in many different areas and are holistic minded. On the other hand, the majority of the members of the organisational focused cluster (more than 60%) cover only 1 to 3 enablers. Therefore Vaughan’s concerns regarding focusing in one area leads to negligence in others [34] cannot be confirmed for the majority of samples and clusters - based on the data at hand.

4. Conclusion

As the initial explorative analysis demonstrated, by applying the COBIT enabler to the characteristics of IT senior executives, certain patterns could be identified. These patterns confirm that two major groups – "generalists" and "specialist" – can be found. Therefore, it has been proven, that COBIT and its enablers, in conjunction with social networks, can be used for classification of types of working attitudes.

Our analysis has revealed that company size influences competence profiles, whereas regional aspects or work years do not seem to be important. We have identified two major clusters – the technical-oriented generalist IT managers. Smaller companies seem to expect more from their IT management rather than technical skills (more than 40%), which declines to less than 30% for large companies. At the same time, generalists rise from less than 40% to well above 50%, thus making a shift in difference from -5% to more than 20%. The cluster with people having a focus on organizational skills decreases as well, from small to larger companies, but with a shallower slope. The number of socially or company culture-focused people increases at the same time with an equal slope. Therefore, it can be assumed that the bigger the company, the more general and social aspects become relevant. The analysis cannot make any predictions about the depth of knowledge or the kind of company culture, but we can make conclusions about what the people themselves think is important.

A very interesting result is that both major clusters still have a broad coverage of all COBIT enablers. Hence, among these individuals, a holistic mind-set can be assumed. While for the "generalist" cluster, this is no surprise, it is quite unexpected that the members of the technical-focused cluster still exhibit strong managerial skills. Further research efforts need to be undertaken to determine how these commonalities and differences have an influence on the organization of companies and their culture.

Our study has shown that the analysis of social network data can give valuable insights into the characteristics of senior IT management in Germany, confirming previous studies with different approaches and is therefore a suitable method for researching employee qualifications.

5. Declarations

5.1. Author Contributions

Conceptualization, S.G. and S.N.; methodology, S.G.; software, S.G.; validation, S.N. and L.W.; formal analysis, S.G.; investigation, S.G.; resources, S.G.; data curation, S.G.; writing—original draft preparation, S.G.; writing—review and editing, S.N.; visualization, S.G.; supervision: J.R.; project administration, J.R.; funding acquisition, J.R. All authors have read and agreed to the published version of the manuscript.

5.2. Data Availability Statement

The data presented in this study are available on request from the corresponding author.

5.3. Funding

The authors received no financial support for the research, authorship, and/or publication of this article.

5.4. Declaration of Competing Interest

The authors declare that there is no conflict of interests regarding the publication of this manuscript. In addition, the ethical issues, including plagiarism, informed consent, misconduct, data fabrication and/or falsification, double publication and/or submission, and redundancies have been completely observed by the authors.

6. References


