A TASK FRAMEWORK FOR PREDICTING THE EFFECTS OF AUTOMATION

Research paper

Koorn, Jelmer J., Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, j.j.koorn@vu.nl

Leopold, Henrik, Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, h.leopold@vu.nl

Reijers, Hajo A., Vrije Universiteit Amsterdam, Amsterdam, the Netherlands, h.a.reijers@vu.nl

Abstract

The ongoing digitalization changes the nature of work. Nowadays, even complex tasks can be automated and reliably performed by machines. This new wave of automation has led to an increased interest in predicting the effects of automation on job design. A recent study suggests that around half of today's jobs could disappear in the coming twenty years. However, these results are heavily debated. Other studies claim that the effect of automation will be much less dramatic. A fundamental issue underlying all these studies is the question of how to categorize tasks. Some authors simply divide tasks into routine and non-routine tasks, others also consider which kind of cognitive abilities are required. Since the predicted effect of automation directly relates to the categories considered, a sound task framework is essential for useful predictions. Recognizing that existing task models are limited in terms of granularity and time, we use a literature study, interviews, and an analysis of historical data to systemically develop a new task framework for predicting the effects of automation. We conduct an evaluation of our framework to demonstrate the generalizability of the framework and compare the framework with existing models.

Keywords: task framework, automation prediction, digitalization.

1   Introduction

Fears about the effects of technological progress have been widespread since the beginning of the industrial revolution in the 18th century (Autor, 2015). At that time, this fear mainly related to the rise of machines in production processes. Today, the increasing digitalization and its potential effect on the automation of work dominates the public debate in particular (Eichhorst et al., 2017). However, reflections on the opportunities and effects of information technology on human work go back farther. Already in 1956, Howard Levin highlighted the value of information technology for eliminating manual interventions in data processing activities (Levin, 1956, p. 61). He concluded that technology will have a "*considerable effect on the office*". In 1961, Georges Friedmann analyzed the anatomy of work and came to a similar conclusion. He stated that automation will "*suppress most fragmentary and repetitive jobs*" (Friedmann, 1992, p. 114).

The disruptive changes that organizations might experience due to automation triggered an interest in predicting the effect of automation on job design. In 1962, Louis E. Davis conducted one of the first studies that systemically investigated this matter (Davis, 1962). However, he also admitted that many pieces of relevant information are simply not available and that his predictions should be rather understood as trends. While predicting the impact of technology remains a considerable challenge, contemporary researchers can exploit rich sources of data. One of the most recent studies was conducted by Frey and Osborne (2013)[1]. They concluded that around half of today's jobs in the USA are likely to disappear in the coming decade or two. Replications of their study for other OECD countries came to

---

[1] Please note that the working paper cited above got recently published in the journal *Technological Forecasting and Social Change* (Frey & Osborne, 2017). Since we discuss many works referring the version of 2013, we also refer to Frey and Osborne (2013) to avoid confusion.

similar conclusions (cf. Pajarinen & Rouvinen, 2014). However, several authors questioned the methodological choices made by Frey and Osborne. For instance, Arntz, Gregory, and Zierahn (2016) argue that considering jobs as a whole is too course-grained and that the analysis should rather be conducted on a task level. They replicated the study from Frey and Osborne with a focus on tasks and concluded that only about 9% of all jobs are in danger of being fully automated.

A fundamental issue underlying all these studies is the way how tasks are categorized. Some authors simply divide tasks into easy or hard to automate (Frey & Osborne, 2013). Others developed more sophisticated task models, categorizing tasks into routine and non-routine or categorizing them based on whether they require cognitive or manual abilities (Autor, Levy, and Murnane, 2003; Spitz-Oener, 2006). What all these task models have in common is that they are static from the perspective of *granularity* as well as from the perspective of *time*. This means that existing task models only define their categories on one level of granularity and only reflect the state of technology at the time of publication. Recognizing this, we use this paper to answer the following research question: "*How do tasks need to be categorized in order to predict the effects of automation on different types of jobs?*". To answer this question, we conducted a literature review, interviews, and an analysis of occupational data. On that basis, we systematically developed a new task framework. Our contribution, therefore, is a task framework that explicitly addresses the dimension of *granularity* and *time*.

The rest of the paper is organized as follows. Section 2 discusses the background of our work. Section 3 introduces the methodology of our research. Section 4 presents the results, that is to say, our new task framework. Section 5 elaborates on the evaluation of task framework. Section 6 discusses implications and limitations before Section 7 concludes the paper.

2 Background

In this section, we discuss the background of our work. We first elaborate on research that is concerned with predicting the extent of work automation. Second, we discuss the conceptual foundation for automation prediction, the so-called task models. Third, we reflect on the limitations of existing task models and identify the research gap we address in our research.

2.1 Predicting the Extent of Work Automation

The increasing digitalization on the job market in the last decade created the strong desire to predict the extent of automation in the future. One of the most notable attempts on the academic side was undertaken by Frey and Osborne (2013). Based on an extensive study of job market data, they concluded that around half of today's jobs in the USA are likely to disappear in the coming decade or two. The fundamental premise of their study was that all tasks *can* eventually be automated. Some tasks, however, are technically more difficult to automate and thus will be automated at a slower pace. According to Frey and Osborne (2013), tasks that are harder to automate require skills such as perception and manipulation, creative intelligence, or social intelligence. Tasks that do not require these skills, can be automated in a straightforward fashion. The study of Frey and Osborne was replicated for other OECD countries (cf. Pajarinen and Rouvinen, 2014), which led to similar results.

After Frey and Osborne (2013) published their study, critique regarding their methodological choices was voiced. In their study, Frey and Osborne (2013) studied the automation potential mainly on a job level. Arntz et al. (2016) argue that the *job* level is too course-grained and that the analysis should be conducted on a *task* level. In order to show that the level of aggregation indeed affects the prediction outcome, Arntz et al. (2016) replicated the methodology from Frey and Osborne (2013). They, however, applied their classification scheme to tasks rather than jobs and only aggregated their results at the end of their analysis. The results from Arntz et al. (2016) indicated that only about 9% of all jobs in the OECD countries are in danger of being automated, rather than the 47% predicted by Frey and Osborne (2013). While this demonstrated the importance of the unit of analysis, it also sparked a more fundamental discussion on the methodology of categorizing tasks. Arntz et al. (2016) argue that also the way how task categories are defined has strong implications for the prediction outcome. At the same time, they question the suitability of the task categories used by Frey and Osborne (2013). Therefore, in the next section, we will elaborate on the problem of categorizing tasks in more detail.

2.2 Task Models for Automation Prediction

There are two main perspectives on how to categorize tasks for the purpose of analyzing and predicting automation potential. One stems from the field of *labor economics*, the other from the field of *ergonomics*. The former derives categories from a skill demand perspective, i.e., which skills are required to execute a task. The latter, by contrast, derives categories from a human-machine interaction perspective, i.e., the way a human uses a machine to execute a task.

Autor et al. (2003) presented one of the first task models in the field of *labor economics*. Their categories are based on the observation that, at the time of their analysis, technology was capable of automating routine tasks, but not non-routine tasks. In addition, they argue that manual tasks are easier to automate than cognitive tasks. Thus, they conclude that two distinctions are relevant when defining task categories: *routine* versus *non-routine* and *manual* versus *cognitive*. However, as automation progressed, the distinctions made by Autor et al. (2003) became outdated. Recognizing this, Spitz-Oener (2006) updated the task categories presented by Autor et al. (2003) in order to reflect the progress of automation. She argued that a more precise distinction within the non-routine cognitive task categories is necessary and defined five subcategories: (1) *routine manual*, (2) *non-routine manual*, (3) *routine cognitive*, (4) *non-routine analytical*, and (5) *non-routine interactive*. Frey and Osborne (2013) followed an alternative approach and redefined task categories based on specific skills that are hard to automate. They defined tasks as either automatable or hard to automate. The latter category consists of tasks requiring the skills: (1) *perception and manipulation*, (2) *creative intelligence*, and (3) *social intelligence*.

Researchers from the field of *ergonomics* argue that the aspect of human-machine interaction in particular is important for analyzing and predicting automation potential. In general, the field of ergonomics focuses on the productivity constraints rather than the technological constraints related to the automation of tasks. Onnasch et al. (2014) showed that plotting the extent of automation against the productivity results in a graph with a U-curved shape. Although the optimal point of automation is a much-debated topic, the shape of the graph shows how the interaction between human and machine restrains the pace of automation. Therefore, the task categories used in this field of research focuses on the interaction between human and machine. The Degrees of Automation (DOA) model proposed by Parasuraman, Sheridan, and Wickens (2000) is amongst the most recent models in this field. The model contains the four types of tasks in which human and machines are most likely to cooperate: (1) *information acquisition*, (2) *information analysis,* (3) *action selection*, and (4) *action implementation*.

2.3 Limitations of Existing Task Models for Automation Prediction

When reviewing the task models discussed above, it becomes apparent that they have two main limitations. They are static from the perspective of *granularity* as well as from the perspective of *time*.

The first limitation results from a lack of addressing different levels of *granularity*. Some researchers define highly specific task categories for certain types of tasks (Parasuraman et al., 2000). Others try to define task categories that are highly generic (Spitz-Oener, 2006). Although various levels of granularity are presented in the literature, there is no task model available that is adjustable to different levels of granularity. If, for example, a researcher is interested in looking into a particular job field, none of the presented task models takes this into account.

The second limitation results from the fact that existing task models only reflect the state of technology at the *time* of publication. This is illustrated by the various updates that have been proposed for older task models in the past (Frey & Osborne, 2013; Spitz-Oener, 2006). In general, this cannot be considered surprising since new technology can be expected to affect tasks and their execution. This, however, also means that an outdated task model is unlikely to lead to accurate and reliable conclusions about the automation potential.

Given this gap, it is the goal of this paper to develop a task framework that appropriately takes the dimensions of *granularity* and *time* into account. The next section elaborates on how we aim to achieve this.

3   Methodology

To answer our research question of how tasks need be categorized in order to predict the effects of automation on different types of jobs, we followed the *three-step research design* illustrated in Figure 1. First, we collected data from three different sources: (1) literature, (2) historical data from the database O*NET, and (3) interviews. The overall rationale behind combining three different types of data collection was to triangulate the data such that the biases associated with each of the data source were minimized. The literature provides different views on how to categorize tasks. Historical data from the database O*NET provides quantitative insights about how the tasks associated with certain jobs have changed over time due to automation. Interviews provide qualitative insights about how experts perceive the impact of automation. Second, we synthesized the insights from the three data sources into a new task framework. To this end, we identified task categories by comparing the findings of the three data sources and positioned the task categories in a framework. Third, we evaluated the proposed task framework. The main goal of the evaluation is to demonstrate the generalizability and extent to which the proposed task framework leads to new insights. In the following sections, we describe the first two steps of our methodology in detail. The evaluation is discussed in Section 5.
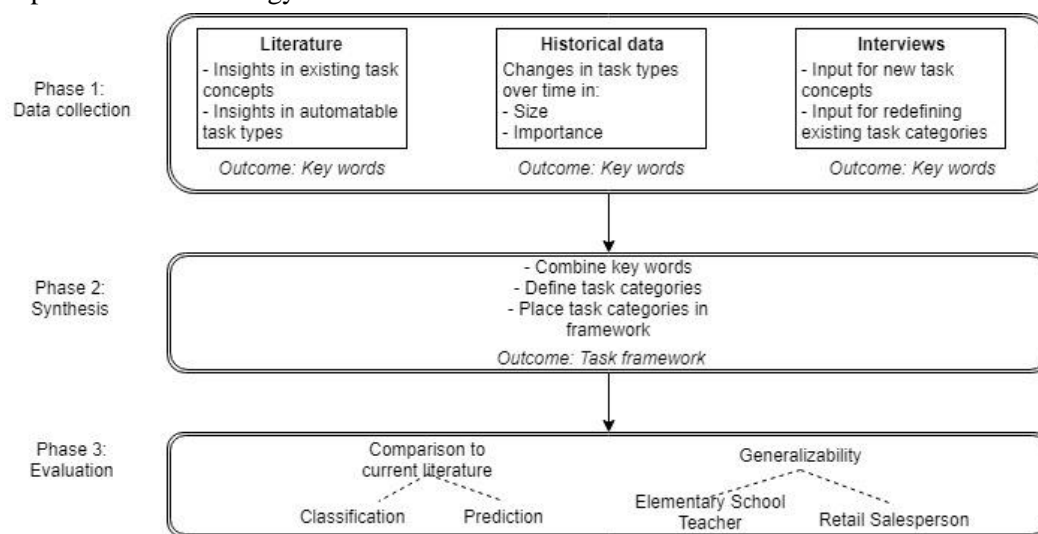

Figure 1.   Research design.

3.1   Data Collection

In this section, we elaborate on our data collection. We first introduce a reference task model, which we use as a basis for coding the considered data sources. Then, we explain the data collection for each of the data sources in more detail.

3.1.1   Reference Task Model

To be able to group the findings from the three considered data sources in a comparable way, we decided to pick reference task model. More specifically, we chose the task model presented by Spitz-Oener (2006). This decision was made based on two reasons. First, in the model from Spitz-Oener (2006) the task categories are defined in very general terms. Hence, it represents a viable choice to code our findings from literature and the interviews. Second, Spitz-Oener (2006) presents a list of verbs corresponding to each task category. For example, the verb "*negotiate*" refers to the *interactive* task category. We deemed this a particularly useful input for the interviews because it would allow us to make it easier for the interviewees to understand the notion of a task category.

3.1.2   Literature

To get a comprehensive overview of the possibilities to categorize tasks with respect to their automation potential, we carried out a structured literature survey. Since the topic of automation is discussed in a variety of fields, we did not opt for any domain-specific libraries or databases. Instead, we used Google Scholar, which can be considered as the most comprehensive search engine for academic literature covering all scientific domains. As search terms we used "*automation*", "*task automation*", "*job*

automation", "*routine automation*", ("*task*" OR "*job*") AND ("*categories*" OR "*categorization*"), "*job requirements*", "*degrees of automation*", "*level of automation*", and ("*human-machine*" OR "*robot*" OR "*technology*") AND ("*cooperation*" OR "*interaction*"). We selected papers based on their title, abstract, and keywords. Further filtering was done by full text reviewing. As a result, we identified 22 relevant papers. In order to extract specific insights with respect to task categories from these papers, we applied the following procedure:

1. We systematically analyzed all papers for relevant statements with respect to task categories. We identified 101 relevant statements from 12 papers.[2]

2. We characterized the statements relating to task categories using a number of keywords such as "*interactive*", "*analytical*" or "*creative*".

3. We took the most recent task model from literature (Spitz-Oener, 2006) as a starting point and checked whether the categories relating to the keywords from the previous step were covered.

4. We grouped all statements that did not fit into the task model from Spitz-Oener (2006) together. Out of the 101 statements, we assigned 39 to this group. By applying in vivo coding, we derived new task categories.

5. We revisited the analysis above and refined the derived categories by combining and splitting them.

3.1.3   Historical Data

To collect quantitative data about how the tasks associated with certain jobs have changed over time, we consulted the database O*NET[3]. The O*NET database contains information on hundreds of standardized and occupation-specific descriptors. It is continually updated by surveying a broad range of workers from each occupation. The O*NET database is described as "*the primary source of occupational information*" and is maintained by the U.S. Department of Labor. The O*NET data is collected with the goal to support the development of tool aimed at career exploration and job analysis (National Center for O*NET Development, 2017).

We selected two specific occupations: (1) the *accountant* and (2) the *auditor*. We chose these occupations because they are considered to be among the most likely candidates for automation in the near future (Arntz et al., 2016) and thus are likely to show clear trends. What is more, these two occupations have already been strongly influenced by automation in the past (Moffitt et al., 2016; Wilson & Sangster, 1992). To develop an in-depth understanding of how these two jobs and the tasks associated with them have changed over time, we used the full time range provided by the O*NET database. The earliest available data about the tasks associated with these jobs was from 1998, the most recent from 2017. We selected 2007 as an intermediate point of time to also be able to see gradual developments.

We selected the variables *tasks* and *work activities* from this data set. *Work activities* closely resemble *tasks* in terms of their definition in the O*NET database (National Center for O*NET Development, 2017), from here on we refer to both as *tasks*. These variables provided a basic description of the tasks along with a score of importance for each task. The tasks were deductively coded based on the task categories presented by Spitz-Oener (2006). The specific insights obtained from the O*NET included an overview of emerging and disappearing tasks and an importance score for each task of the two considered jobs for each selected year.

3.1.4   Interviews

To obtain insights into how experts perceive the impact of automation, we conducted a set of interviews. More specifically, we conducted one explorative interview and four in-depth interviews.

For the *explorative interview*, the main goal was to (1) get a better understanding of the field of accounting and auditing, (2) develop an interview guideline for in-depth interviews, and (3) obtain a list of suitable candidates for the in-depth interviews. To meet these goals, we required a person to be familiar with both technology and accounting. Therefore, we consulted a full professor in accounting information systems from the Dutch Business School TIAS, which is affiliated with both Tilburg University and Eindhoven University of Technology. Based on the insights from this interview, we select-

---

[2] For a full overview of these papers, please refer to Appendix A.
[3] O*NET is available online: https://www.onetcenter.org/db_releases.html.

ed four candidates for *in-depth interviews*. Table 1 provides additional details about the interviewees and their background.

Interviewee	Gender	Background information
Interviewee 1	Male	Director digital innovation & assurance department in a consultancy firm
Interviewee 2	Male	Former accountant, now founder of a startup focusing on automation accountancy analyses
Interviewee 3	Female	Partner audit innovation and Ph.D. candidate working on the integration of process mining in auditing
Interviewee 4	Male	Professor accountancy at Tilburg University specializing in accounting information systems and has experience as a partner in an accountancy firm

Table 1.   Background information about interviewees.

The conducted interviews were semi-structured and consisted of two parts: (1) discussing which tasks in the context of accounting and auditing will and will not be automated;, (2) reflecting on the task categories as defined by Spitz-Oener (2006). The interviews lasted between 45 and 60 minutes. We recorded and fully transcribed all interviews.

In total, in the transcripts 201 quotes were found, of which 188 were included in the final sample. In total, thirteen quotes were deleted for being too vague (e.g. "*routine accountancy tasks can be automated*") or not specifically talking about a task that can be automated (e.g. "*automation will allow us to place all our tooling within the cloud*").

The same coding scheme as presented in the literature was applied to the interview quotes. First, the relevant quotes were extracted and characterized in terms of key words. Then, these key words are compared to Spitz-Oener (2006). If they did not fit their task categories, *in vivo* coding was applied to define new task categories.

3.2   Synthesis

In the synthesis phase of the research we took two steps: (1) we defined task categories, (2) we placed these task categories in a framework. The first step ensures the flexibility of the newly proposed task categorization with respect to *granularity*. The second step focuses on the flexibility of our framework with respect to *time*.

In the first step, we coded each of the data sources using the task categories from Spitz-Oener (2006). If key words were identified not fitting the task categories of Spitz-Oener (2006), we used inductive coding to identify groups of related key words. For each of these groups of key words, we chose an overarching term based on the literature. The task categories defined in this step are useful for a more high level focus of research.

We then further refined the definition within each of the task categories. We cross-referenced each of the key words with all the tasks/quotes. If a key word could not capture a task/quote, a distinction was made. We defined these distinctions based on literature (Rohrbach-Schmidt & Tiemann, 2013). Defining these distinctions within task categories is useful a for more fine grained research focus.

In the second step, we placed the task categories in a framework. Placing tasks in a framework rather than presenting them as a list of categories allows for changes in technology over time to become visible. In order to achieve this, we placed each task category on two dimensions. One dimension for degree of routine actions, and one dimension for perceived difficulty to automate.

4   Results

Section 4.1 first discusses the definitions of the new task categories and their subdivisions, as presented in Table 2. Section 4.2 then elaborates on the positioning of each category in a new framework.

4.1   Task Categories

Task category	Subdivision
Creative	
Adaptive	
Interactive	Routine
Analytical	(1) Evaluation (2) Standardization
System supervision	
Routine cognitive	
Information processing	
Information exchange	Data stream

Table 2.   New task categories and their respective subdivisions

We defined two task categories based on the task categories presented by Frey and Osborne (2013): *creative* and *adaptive*. The *creative* task category emerged from both the literature and interview data. It captures tasks such as "*developing new meaningful ideas/artefacts*". Usually, *creative* tasks occur in an unprecedented context, making it extremely hard to automate them. The *adaptive* task category only emerged from literature. Examples of quotes characterizing this task category are: "*orienting in a complex situation*" and "*reacting to (potential) failures and unstructured challenges*".

We found the *interactive* task category in all data sources. Examples of key terms describing tasks in this category well are: "*emphatically interacting face-to-face*" and "*training and instructing people*". The interview data shows that there exists an inherit split between routine and non-routine tasks within the *interactive* task category. "*routine*", "*repetitive*", and "*standardization*" are terms used by the interviewees to indicate which type of *interactive* tasks can be automated. These terms all refer to the division that was initially made by Autor et al. (2003): *routine* versus *non-routine*.

Based on contradictions found for the *analytical* task category among the three data sources, we redefined the task category using two concepts: (1) *evaluation*, and (2) *standardization*. The first axis that separates the analytical task category concerns the concept of *evaluation*, which is an term capturing all key words originating in the interviews. Patton (1987) suggests a way to separate concepts within evaluation. He argues that there are four levels: (1) findings, (2) analysis/interpretation, (3) judgement, and (4) recommendations. We split the *analytical* task category with an axis that starts with findings and ends in recommendation. The axis is continuous since the transitions from one to the other concept are very subtle. Moving from beginning to end, it is perceived to be increasingly difficult to automate these tasks.

The second axis that splits the *analytical* task category is the degree to which standardized data is used. This term was also brought forward during the interviews. One example is where the interviewee describes how once the template has been designed, exchanging and processing data can be automated. XBRL is a computer standard that is often mentioned as an example: "*After a company submits their data in XBRL format to the bank, without human intervention, the XBRL data can be processed within the banking system.*" Standardized data consists of two parts: (1) standardized structure, and (2) standardized meaning. Once both these elements are present, tasks are more susceptible to automation. Human action is still required to design a template for the exchange and processing of data. Once this template is designed, data can flow automatically. We depict *standardization* as a continuous axis that moves from standardized to not-standardized data with all degrees of partially standardized data in between. *Adaptive* tasks are very similar to *standardization*, since both deal with the input of the environment. Where standardization is concerned with the unification of incoming data into one standard, *adaptive* tasks capture the response to a change in the environment. Changes in the environment can also be seen as incoming data.

The *system supervision* emerged from the literature analysis. *System supervision* captures the situation in which the human operator oversees a system, examples of tasks are: "*database maintenance*" and "*system monitoring*". The *interactive* task category only captures human-human supervision.

We defined the *routine cognitive* task category to be closely related to the task category with the same name presented by Spitz-Oener (2006). This task category emerged from all data sources. Key words such as "*simple calculations*", "*correcting texts/data*", and "*alphabetizing a list of names*" capture this task category best. Where Spitz-Oener (2006) categorizes standardized reporting tasks in her *analytical* task category, we include it in the *routine cognitive* category. We decided this on basis of the data brought forward during the interviews. Based on the literature, *routine cognitive* could be defined to include information acquisition tasks. However, the information acquisition tasks are included in the *information exchange* task category as described below.

We added two task categories, *information exchange* and *information processing*, which closely relate to the DOA model (Parasuraman et al., 2000). Key words for the categories relating to the DOA model were found in both the literature and interviews. We defined the *information exchange* category based on the interviews. *Information exchange* captures both in and outgoing data streams. Tasks such as "*extraction of data elements*", "*information acquisition*", and "*data transmission*" reflect the essence of this task category best. In addition, *information processing* forms its own category capturing the tasks that for a large part determine the level of standardization of data. This category includes tasks such as: "*integration of files*" and "*structuring of data*".

4.2   Placement in Framework

In this section we focus on placing the task categories in a framework. The new task framework is presented in Figure 2. We determined the position of a task category in the framework based on two concepts: (1) its perceived difficulty to automate, and (2) its perceived proportion of routine tasks.
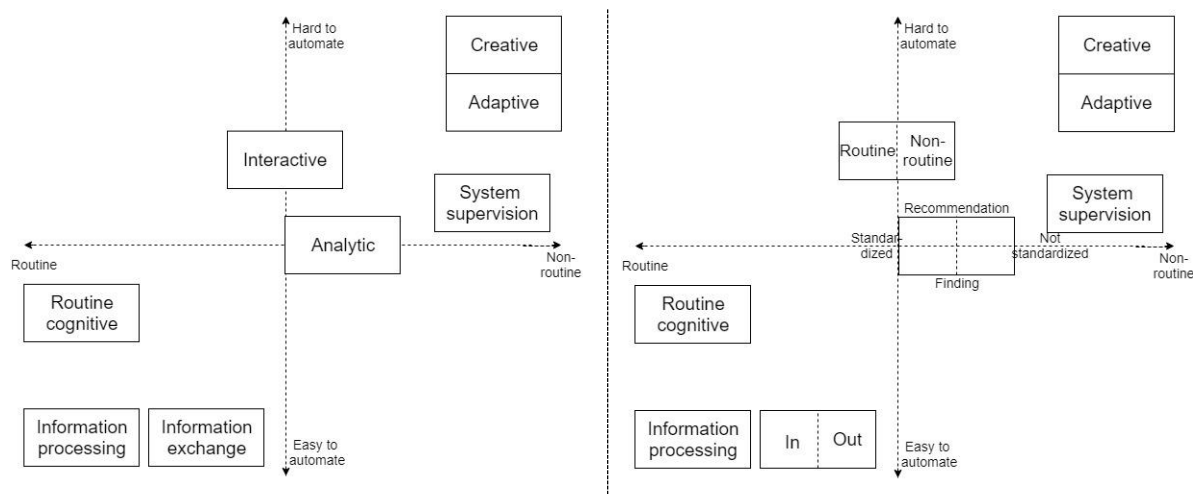

Figure 2.   New task framework: broad (left) and detailed (right).

We evaluated the perceived difficulty to automate by reviewing the categorization of the key words. More specifically, we categorized the key words into automatable and not automatable during the coding. Table 3 shows *the degree of perceived difficulty to automate* for each of the task categories. Note that the historical data is coded based on the task categories from Spitz-Oener (2006). Thus, this data cannot provide insights into the other task categories. Therefore, it has not been included in the table above.

The data clearly shows why the two task categories *creative* and *adaptive* are placed at the top of the framework. These categories are viewed as most difficult to automate and are exclusively mentioned as ´cannot be automated´. One of the underlying reasons for this might be that tasks from both categories often take place in unstructured and unprecedented environments.

The task categories placed in the middle of the spectrum, *interactive, analytical,* and *system supervision,* are much more subject to debate. *Interactive* tasks are mostly perceived as difficult to automate, for example, jobs tend to have a growing number of interactive tasks. The interview data provided insights in how especially routine *interactive* tasks can be automated. The *analytical* task category is discussed the most in the interviews. Historically for the accountant a declining number of *analytical* tasks can be found, but this number is growing for the auditor. The *system supervision* tasks were only

	Literature		*Interviews*	
	A	**NA**	**A**	**NA**
Creative	0	11	0	2
Adaptive	0	7	0	0
Interactive	0	14	2	6
System supervision	8	7	0	0
Analytical	9	2	12	8
Routine cognitive	9	0	6	1
Information exchange	0	0	8	0
Information processing	0	0	3	0

Table 3.   Number of key terms per data source named in the scenario of A = can be automated and NA = cannot be automated.

found in the literature and were mentioned almost as often *can be automated* as *cannot be automated*. Due to these debates it is most appropriate to place these categories in the middle of the spectrum.

The task categories perceived as least difficult to automate, *routine cognitive, information exchange,* and *information processing*, were consistently named as *automatable*. Interestingly, the *routine cognitive* tasks are evaluated as becoming more important over time. In contrast, there is no such trend in the number of tasks and only one quote in the literature indicating these tasks are hard to automate. For the other two task categories no quotes are found at all indicating these tasks cannot be automated. Therefore, these tasks are placed at the bottom of the axis.

In contrast to the perceived difficulty to automate, the ratio of routine versus non-routine tasks within a task category cannot be expressed in numbers. Therefore, decisions for placement of task categories on this axis were made based on the definitions we have proposed previously.

5   Evaluation

In the evaluation part of the research, we address two specific goals. First, we aim to demonstrate that the proposed framework is applicable beyond the exemplary case studies (auditor and accountant) used to create the framework. Second, we aim to show if the proposed framework leads to new insights compared to current models.

To get a better understanding into the generalizability of the proposed task framework, we applied the framework to two other occupations: the *retail salesperson* and the *elementary school teacher*. The choice for these specific occupations is based on two considerations. First, the nature of these jobs differs to a large extent from the one of an accountant and an auditor. While the occupations of an accountant and an auditor mainly require the analysis of documents, a retail salesperson and an elementary school teacher require much more interaction with people. Second, the selected occupations differ with respect to their vulnerability to automation. Based on the results from Frey and Osborne (2013) and Arntz et al. (2016), the elementary school teacher is very unlikely to be affected by automation, whereas the retail salesperson is very likely to be affected by automation. Given these differences, applying the proposed task framework to these occupations provides indications for the generalizability of the framework.

To also see whether the proposed framework (1) classifies task differently than current literature and (2) generates a prediction with respect to automation differing from current literature, we evaluated how the proposed framework positions itself with regards to existing models. The proposed framework is only of value if it leads to different insights than those that can be gained with the current models. To keep the evaluation consistent, the analysis was done using the newly selected occupations of a retail salesperson and an elementary school teacher.

To check *generalizability* of the framework to the occupations of a retail salesperson and an elementary school teacher, we extracted the tasks associated with these occupations from the most recent database from O*NET (i.e., the version of 2017). We then checked whether we could classify the tasks from O*NET using our task framework. We found that we were able to classify 110 out of 112 tasks. The two tasks that we were unable to classify were: "*interacting with computers*" and "*making decisions and solving problems*". We considered these tasks as too broad to be classified using our framework and therefore disregarded them. The fact that we were able to classify 98% of the tasks for the two occupations demonstrates that the proposed framework is very promising in terms of generaliza-

bility as the framework categorizes tasks not only relevant for the exemplary cases, but consists of task categories relevant for a wider spectrum of occupations.

We now turn to evaluating the new insights the proposed framework could provide. First, we focus on the *classification* part. We checked the number of tasks *classified differently* based on our framework compared to the model proposed by Spitz-Oener (2006). We choose Spitz-Oener (2006) because her task categories are phrased in general terms. The results of this analysis show that there is a clear indication that our framework classifies a substantial number of tasks and work activities differently than Spitz-Oener (2006). On an aggregate level, 37 out of 112 tasks have been either classified in another task category or have received an additional task category classification. Figure 3 visualizes the results for the elementary school teacher and the retail salesperson.
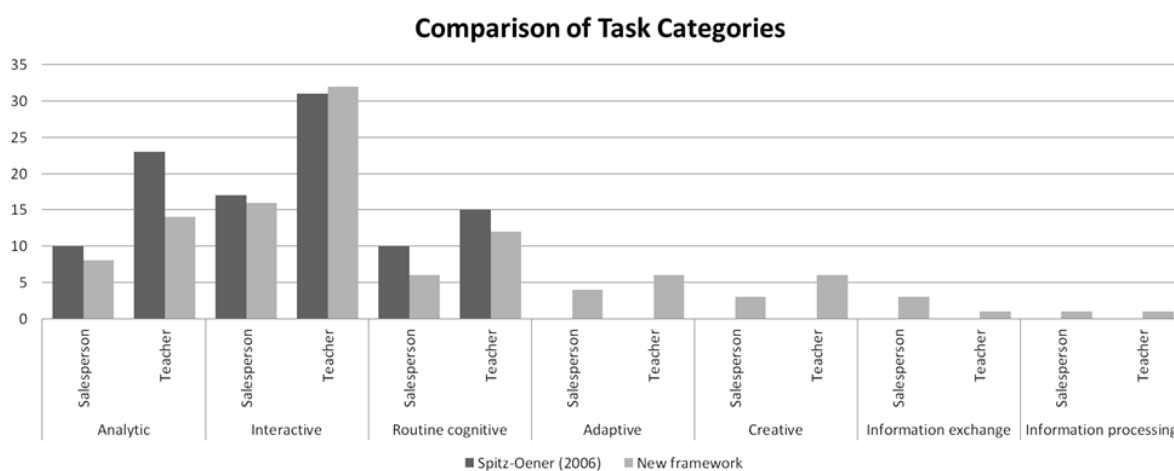
Comparison of Task Categories

Figure 3.   *Differences in task categorization between Spitz-Oener (2006) and the proposed framework for the occupations elementary school teacher (Teacher) and retail salesperson (Salesperson) using 2017 O*NET data.*

Second, we focus on the *prediction* part. We compared the *number of tasks that are predicted to be automated* according to our framework with the predictions from Frey and Osborne (2013). We chose the model from Frey and Osborne (2013) as reference here because they are most explicit about which types of tasks can and cannot be automated. The prediction of Frey and Osborne (2013) is that all tasks will be automated, except for tasks that fall into one of three engineering bottlenecks: (1) *creative intelligence*, (2) *social intelligence*, and (3) *perception and manipulation*. Our results show that the prediction of the vulnerability with respect to automation based on the new framework is substantially different from the prediction made by Frey and Osborne (2013). The percentage of tasks that are vulnerable to automation within a job is consistently estimated to be lower compared to Frey and Osborne (2013). Where Frey and Osborne predict that 96% of all tasks of the *retail salesperson* are susceptible to automation, our framework predicts 83%. For the *elementary school teacher* Frey and Osborne predict 69% of all tasks can be automated, this framework puts this number at 46%. However, in line with Frey and Osborne (2013), *retail salespersons* are substantially perceived to be more vulnerable to automation than *elementary school teachers*.

In summary, we can state that this evaluation provided some clear indications that the proposed framework is sufficiently generalizable to other occupations. We also found differences in comparison to existing literature with respect to the way our framework classifies tasks and predicts automation. This indicates how this study provides new insights that could not be obtained using existing models.

6   Discussion

In this section, we discuss the implications of our research and reflect on its limitations.

6.1   Implications

The presented task framework provides insights into how automation may affect jobs and their respective tasks in the future. Most importantly, it gives an idea of which tasks will be performed by machines rather than by humans. What is more, it shows the interaction between humans and machines.

In comparison to the model presented by Spitz-Oener (2006), our framework contains a substantially larger amount of human-machine interaction tasks. This increased amount and different nature of human-machine interaction shows that the interaction between human and machine can be expected to become more complex. These insights have particular relevance for information systems use, as well as for information systems design. Research on information systems use is, among others, concerned with patterns of how users interact with information systems (Deng & Chi, 2012; Ortiz de Guinea & Webster, 2013; Sun & Teng, 2012). The results presented in this paper indicate that these patterns of use might change and that a more fine granular view might be required on how automated and manual work will be blended to effectively use information systems in the future. In a similar way, our results inform research on information systems design (Te'eni, Carey, & Zhang, 2005). The increased level and different types of interactions might require new ways of designing information systems. As an example, consider the use of Robotic Process Automation (RPA) in the service domain (Davenport & Kirby, 2016). The vision of RPA is to fully automate routine service requests and, in this way, give humans more time to focus on complex, non-routine requests. Although RPA solutions are successfully employed in many organizations, they also come with completely new challenges. One of the most prominent questions is how information system design can ensure effective collaboration between humans and machines. While the research in this paper cannot answer this question, the task framework provides a valuable starting point for further investigating this matter.

6.2 Limitations

The results presented in this paper are subject to certain limitations. First, it has to be noted that the scope of this research is mainly limited to cognitive tasks. Hence, the presented framework only provides little insights with respect to the nature of manual tasks. It would be particularly interesting for future research to focus on occupations in which a human operator performs manual tasks in a highly automated environment. This, however, was a deliberate choice we made due to the complexity associated with studying automated tasks. Second, there are a number of limitations related to the data sources: (1) assumption of causality, (2) lack of four eye principle while coding, and (3) small number of interviews. One of the assumptions of the historic data analysis was that all the observed effects were caused by automation. Although the findings seem to be consistent with the literature, there are numerous other factors besides automation. For example, changes in legislation can influence the nature of tasks, too. We did not account for these factors in this work. Second, for both the literature and interviews, the coding was conducted by a single researcher. Ideally, the four eyes principle should be applied in order to guarantee the consistency and validity of the coding scheme. In order to mitigate the impact of this limitation, the coding has been subjected to multiple cycles, in which the researcher had to review and compare the codes. Third, the number of interviewees was rather low. In an ideal scenario, more people would have been interviewed to reach a higher degree of saturation.

7 Conclusion

In this paper, we addressed the question of how to categorize tasks in order to predict the effects of automation on different types of jobs. More specifically, we used a literature analysis, interviews, and an analysis of occupational data from the database O*NET to systematically develop a new task framework. In comparison to existing task models, our framework explicitly accounts for *granularity* and *time*, that is, it accounts for both general and specific tasks and it abstracts from today's technology as far as possible. An evaluation with the jobs of a retail salesperson and an elementary school teacher demonstrated that the proposed framework appears to be sufficiently generalizable beyond our exemplary cases. What is more, it shows that our framework can generate insights that could not be obtained using existing models.

The results presented in this paper particularly inform research on information systems use and design. Our results suggest that the patterns of information systems use might change. Also, a more fine granular view might be required on how automated and manual work will blend together. Our results also suggest that the increased level and different types of interactions with information systems might require new strategies for information system design. From a practical perspective, the presented framework can provide valuable information for organizations. It reveals which jobs are vulnerable to automation and it can serve as a blueprint for training and personnel planning.

References

Arntz, M., Gregory, T., & Zierahn, U. (2016). The Risk of Automation for Jobs in OECD Countries: A Comparative Analysis. *OECD Social, Employment and Migration Working Papers, 2*, 47-54. doi:10.1787/5jlz9h56dvq7-en

Autor, D. H. (2015). Why are there still so many jobs? The history and future of workplace automation. *The Journal of Economic Perspectives, 29*(3), 3-30.

Autor, D. H., Levy, F., & Murnane, R. J. (2003). The skill content of recent technological change: an empirical exploration. *The Quarterly journal of economics, 118*, 1279-1333.

Davenport, T. H., & Kirby, J. (2016). Just How Smart Are Smart Machines? *MIT Sloan Management Review, 57*(3), 21.

Davis, L. E. (1962). The effects of automation on job design. *Industrial Relations: A Journal of Economy and Society, 2*(1), 53-71.

de Witte, M., & Steijn, B. (2000). Automation, job content, and underemployment. *Work, employment and society, 14*(2), 245-264.

Deng, X., & Chi, L. (2012). Understanding postadoptive behaviors in information systems use: A longitudinal analysis of system use problems in the business intelligence context. *Journal of Management Information Systems, 29*(3), 291-326.

Eichhorst, W., Hinte, H., Rinne, U., & Tobsch, V. (2017). How Big is the Gig? Assessing the Preliminary Evidence on the Effects of Digitalization on the Labor Market. *Management revue, 28*(3), 298-318.

Frey, C. B., & Osborne, M. (2013). The Future of Employment: How susceptible are jobs to computerization? *Working paper*, 1-79.

Frey, C. B., & Osborne, M. A. (2017). The future of employment: How susceptible are jobs to computerisation? *Technological Forecasting and Social Change, 114*, 254-280.

Friedmann, G. (1992). *The anatomy of work: Labor, leisure, and the implications of automation*: Transaction Publishers.

Hancock, P. A., Jagacinski, R. J., Parasuraman, R., Wickens, C. D., Wilson, G. F., & Kaber, D. B. (2013). Human-Automation Interaction Research Past, Present, and Future. *Ergonomics in Design: The Quarterly of Human Factors Applications, 21*, 9-14. doi:10.1177/1064804613477099

Kennan, M. A., Willard, P., & Wilson, C. S. (2006). What do they want?: a study of changing employer expectations of information professionals. *Australian Academic & Research Libraries, 37*(1), 17-37.

Levin, H. S. (1956). *Office work and automation*: Wiley.

Moffitt, K. C., Richardson, V. J., Snow, N. M., Weisner, M. M., & Wood, D. A. (2016). Perspectives on Past and Future AIS Research as the Journal of Information Systems Turns Thirty. *Journal of Information Systems, 30*(3), 157-171.

Müller, R., Vette, M., & Mailahn, O. (2016). Process-oriented Task Assignment for Assembly Processes with Human-robot Interaction. *Procedia CIRP, 44*, 210-215. doi:10.1016/j.procir.2016.02.080

National Center for O*NET Development. (2017). *O*NET Analyst Ratings of Occupational Skills: Analysis Cycle 17 Results*. Retrieved from Alexandria,Virginia:

Onnasch, L., Wickens, C. D., Li, H., & Manzey, D. (2014). Human Performance Consequences of Stages and Levels of Automation. *The Journal of the Human Factors and Ergonomics Society, 56*, 476-488. doi:10.1177/0018720813501549

Ortiz de Guinea, A., & Webster, J. (2013). An Investigation of Information Systems Use Patterns: Technological Events as Triggers, the Effect of Time, and Consequences for Performance. *MIS Quarterly, 37*(4).

Pajarinen, M., & Rouvinen, P. (2014). Computerization threatens one third of Finnish employment. *ETLA Brief, 22*(13.1), 2014.

Parasuraman, R., Sheridan, T. B., & Wickens, C. D. (2000). A model for types and levels of human interaction with automation. *IEEE Transactions on Systems, Man, and Cybernetics - Part A: Systems and Humans, 30*, 286-297. doi:10.1109/3468.844354

Patton, M. Q. (1987). *How to use qualitative methods in evaluation*: Sage.

Rohrbach-Schmidt, D., & Tiemann, M. (2013). Changes in workplace tasks in Germany—evaluating skill and task measures. *Journal for Labour Market Research, 46*, 215-237. doi:10.1007/s12651-013-0140-3

Spitz-Oener, A. (2006). Technical Change, Job Tasks, and Rising Educational Demands: Looking outside the Wage Structure. *Journal of Labor Economics, 24*(2), 235-270.

Sun, J., & Teng, J. T. (2012). Information Systems Use: Construct conceptualization and scale development. *Computers in Human Behavior, 28*(5), 1564-1574.

Te'eni, D., Carey, J. M., & Zhang, P. (2005). *Human-computer interaction: Developing effective organizational information systems*: John Wiley & Sons.

Vagia, M., Transeth, A. A., & Fjerdingen, S. A. (2016). A literature review on the levels of automation during the years. What are the different taxonomies that have been proposed? *Applied Ergonomics, 53*, 190-202. doi:10.1016/j.apergo.2015.09.013

Wickens, C. D., Li, H., Santamaria, A., Sebok, A., & Sarter, N. B. (2010). *Stages and levels of automation: An integrated meta-analysis.* Paper presented at the Proceedings of the Human Factors and Ergonomics Society Annual Meeting.

Wilson, R., & Sangster, A. (1992). The automation of accounting practice. *Journal of Information Technology, 7*(2), 65-75.

Xu, H. (2011). The impact of automation on job requirements and qualifications for catalogers and reference librarians in academic libraries. *Library Resources & Technical Services, 40*(1), 9-31.

Appendix A

	Routine manual	Routine cognitive	Analytical	Interactive	System supervision	Adaptive	Creative
Arntz et al. (2016)				NA		NA	NA
Xu (2011)		A		NA	NA		
Kennan, Willard, and Wilson (2006)				NA	NA	NA	NA
de Witte and Steijn (2000)					NA		
Onnasch et al. (2014)		A	A				
Vagia, Transeth, and Fjerdingen (2016)	A	A	A/NA		NA		NA
Parasuraman et al. (2000)		A	A				
Hancock et al. (2013)					NA		
Müller, Vette, and Mailahn (2016)	A		A				
Wickens, Li, Santamaria, Sebok, and Sarter (2010)		A	A				
Autor (2015)	A	A	A/NA	NA		A/NA	NA

Table 4.   *Quotes from the literature assigned to (new) task categories per article. If the authors mentioned the category in scenario 'can be automated' an A is marked, if the authors mentioned a category in scenario 'cannot be automated' an NA is marked. If the authors mention a category in both scenarios it is marked A/NA. If the authors did not mention the category the cell is left blank.*