PILLARS – Pathways to Inclusive Labour Markets: D1.1

Taking Stock of the Effects of Technological Change on Labour Markets: A Systematic Literature Review

First draft

Tommaso Ciarli[1], Tommaso Portaluri[2], Maria Savona[3], Ed Steinmueller[4], Simone Vannuccini[5]

This project receives funding from the European Union's Horizon 2020 Research and Innovation Programme under Grant Agreement No. 101004703.

[1] UNU MERIT (NL) and SPRU, University of Sussex, (UK)

[2] SPRU, University of Sussex, (UK)

[3] Corresponding Author: SPRU, University of Sussex, (UK) and Dpt. Economics and Finance, Luiss University, (IT)

[4] SPRU, University of Sussex, (UK)

[5] SPRU, University of Sussex, (UK)

Table of Content

List of Tables

List of Figures

1 Introduction

This paper addresses one of the core **objectives of WP1 within PILLARS**, that is taking stock of the extant literature and evidence on the effects of technological change on labour outcomes and represents Deliverable 1.1 (D1.1).

D1.1 systematically reviews the literatures that study at least one of the intersections between the three main factors (**technology, trade and regional industrial transformations**) that affect labour markets, in terms of **employment and wage distribution**.

The systematic literature review carried out in this paper focuses on the link between technological change and jobs and tasks.

Much of the (labour) economics literature provides overviews of a single main ingredient, included in the table below, reproduced from a seminal work by Léon Ledesma et al. (2010). This is a useful exercise of identification of the **typical labour elasticity coefficient of technology**. In particular, the coefficient σ (elasticity of substitution of labour to technology) is identified for different time samples by relevant contribution in the field. Different assumptions are formulated on the nature of technical change (Hicks Neutral or Factor augmenting) and several range of variation of σ are provided.

Unlike this, and the burgeoning number of contributions within labour economics, our objective is to "reverse engineer" this coefficient σ by identifying the relevant factors that are usually combined (without an indication of their respective importance) in a labour substitution coefficient. To do so, we start by providing a reasoned and **fine-grained classification** of the technologies included in the current black box of emerging "Automation". This allows us to dig into the idiosyncratic link between **sectoral exposure to** and **adoption of** each of these automation technologies (or a combination thereof) and labour outcomes in the exposed and adopting sectors.

The strategy of **increasing the granularity** of the investigation of both the types of automatized devices and processes and the types of tasks, occupations (and capabilities) that automation **replaces** on the one hand, and the **reconfiguration** of the tasks affected within each occupation on the other hand, is the one that we consider the most appropriate and novel in the context of PILLARS. In addition, the adoption of a finer level of granularity in this analysis also will allow us a higher degree of precision in the analysis of how the labour market outcomes of automation **are mediated by international trade and structural changes of economies**.

The present effort of systematic review allows us to unpack the relevant factors that contribute to a particular coefficient σ and increase awareness of what specific technology is affecting labour markets in the context of the latest wave of automation.

Table 1: Empirical Studies of Aggregate Elasticity of Substitution and Technological Change in the US. Source: Leon-Ledesma et al., 2010

Study	Sample	Assumption on technological change	Estimated elasticity of substitution $\hat{\sigma}$	Estimated annual rate of efficiency change: Hicks neutral: $\hat{\gamma}_N = \hat{\gamma}_K$	Labor augmenting: $\hat{\gamma}_N$	Capital augmenting: $\hat{\gamma}_K$
Arrow et al. (1961)[a]	1909–1949	Hicks neutral	0.57	1.8	—	—
Kendrick and Sato (1963)[b]	1919–1960	Hicks neutral	0.58	2.1	—	—
Brown and De Cani (1963)[a]	1890–1918	Factor augmenting	0.35	Labor saving ($\gamma_N - \gamma_K = 0.48$)		
	1919–1937		0.08	Labor saving ($\gamma_N - \gamma_K = 0.62$)		
	1938–1958		0.11	Labor saving ($\gamma_N - \gamma_K = 0.36$)		
	1890–1958		0.44	?		
David and van de Klundert (1965)[a]	1899–1960	Factor augmenting	0.32	—	2.2	1.5
Bodkin and Klein (1967)[c]	1909–1949	Hicks neutral	0.5–0.7	1.4–1.5		
Wilkinson (1968)[d]	1899–1953	Factor augmenting	0.5	Labor saving ($\gamma_N - \gamma_K = 0.51$)		
Sato (1970)[a]	1909–1960	Factor augmenting	0.5–0.7	—	2.0	1.0
Panik (1976)[a]	1929–1966	Factor augmenting	0.76	Labor saving ($\gamma_N - \gamma_K = 0.27$)		
Berndt (1976)[e]	1929–1968	Hicks neutral	0.96–1.25	?	—	—
Kalt (1978)[e]	1929–1967	Factor augmenting	0.76	—	2.2	0.01
Antràs (2004)[e]	1948–1998	Hicks neutral	0.94–1.02	1.14	—	—
		Factor augmenting	0.80	Labor saving ($\gamma_N - \gamma_K = 3.15$)		
Klump et al. (2007a)[f]	1953–1998	Factor augmenting	0.51	—	1.7	0.4

Notes: Key to estimation method: a = linear single equation; b = implicit; c = linear and nonlinear single equation; d = linear system; e = linear single equation and linear system, f = nonlinear system (Klump et al. 2007a report both constant and time-varying factor-augmenting growth cases; the results reported in this table are from the former case.) This table is updated from Kalt (1978) and Klump et al. (2007b). A full source of references to the above studies may be found therein.

More specifically, our strategy consists first of all in **offering a fine grained classification** of automation technologies; then in exploring **the technical and engineering literature** that focuses on each of the classified technologies, describing – among other things – prototypes and mature automation devices; technical characteristics and conditions of adoption; the extent of automation of the production process and success and failure of prototypes, all of which affects occupational tasks and their associated levels of skills, knowledge codification, levels of routinization, degree of substitution or complementarity with respect to humans, to various extent.

Our **classification** will underpin the rationale of the analysis of the whole of the project. The **methodology** of the present review allows organising the collected information into an **interpretative grid**, that serves as a general scheme to assess the employment impacts of fine-grained-defined technologies. The **results** of the systematic technical literature review will guide the interpretation of the economic results throughout PILLARS.

In sum, the main aim of this exercise is to perform a literature review that broadly addresses the following research questions:

- Can digital automation technologies (as identified here) potentially substitute, complement, and/or reconfigure specific technical tasks within occupations within sectors, that are executed by humans?

- What are these specific tasks across sectors? Can we infer the extent to which specific industries are exposed to specific automation technologies at the level of granularity identified above?
- Can we infer anything on the future of work in relation to the contemporary wave of potentially emerging automation technologies, characterised by the pervasive presence of AI?

The paper is structured as follows:

- We start by offering a brief historical review of previous waves of technological change based on mechanisation and electrification **(Section 2)**. This allows to contextualise our contribution, which illustrates the processes of automation linked to digitisation in comparison to previous waves of mechanisation. It is important to note that this historical contextualisation underpins the rationale of our proposed classification, which will be illustrated in detail following the historical review.
- Next, we lead the reader through the methodological and empirical journey to address the objective above **(Section 3 and in much greater detail in the Appendix – 8)**. This might represent a methodological benchmark of a multidisciplinary approach to look at the sectoral exposure and adoption of automation technologies and their impact on the future of work.
- Finally, we proceed with a preliminary analysis of the central focus of these papers, along a standardised set of dimensions that are relevant to PILLARS **(Section 4)**. We then draw summarise the key messages **(Section 5)** and next steps (**Section 6**).

2 Technical change, automation and employment: A brief historical overview

Contemporary discussions about automation and employment echo a long history of development of labour-saving innovation. This history has unfolded in waves of disruption to existing labour practices and routines and has been accompanied by anxieties and protests (Mokyr et al., 2015).

In previous eras, anxieties and protests declined as new occupations were created and economic growth continued to raise the demand for labour albeit in very different jobs than those that were lost. In every one of these historical occasions of automation anxiety, the claim was made that epochal changes were occurring and that if past incidences of concern were raised, the additional claim was '**this time it will be different**.' This history resembles the classical fable of the boy who cried wolf. After the boy made several unsubstantiated claims that a wolf threatened the village, the villagers ignored him and the wolf ravaged the village's herd. At its heart, therefore, discussion of modern forms of automation in the form of robots, AI, and other manifestations of information and communication (ICT) pose the question – is this the time different, is the wolf going to arrive?

Answering this question requires several empirical assessments. The most important of these is the assessment of **emerging capabilities for labour saving devices** associated with employment. The phrase 'associated with employment' may seem odd, but is an acknowledgement that some labour saving devices are related to activities that are generally not paid – e.g. household appliances. Emerging capabilities are those that extend and deepen the capacity of labour saving devices to substitute for *employed* labour.[6]

A considerable amount of attention has been devoted to **robotics and AI** and this seems to be the result of the proclivity of humans to anthropomorphise such devices – so a robot arm or a decision-making AI algorithm receives greater attention than an automated measuring system for filling containers or a logistical system that monitors the location of parcels in transit. In all these cases, of course, there are implications for human labour, and it is important to consider capabilities broadly to make an accurate assessment of emergent labour-saving solutions.

A second consideration is **the nature of jobs**. Job classifications and designations often reduce the complexity of the tasks that actual workers do. Consequently, the mapping of capabilities to employment may over-estimate the extent that labour-saving technologies may mean job losses. In many cases, deploying labour-saving devices results in the **reconfiguration rather than elimination of jobs** – with some net loss of employment and often an increase in output or productivity (which can then be reflected in lower prices that increase the size of a market or market

[6] Nonetheless, since almost all household services have paid equivalents (e.g. commercial laundry service is a substitute for home laundry), there is a two-way interaction between innovations in the paid and unpaid version of the services (e.g. home sewing machines are adaptations of commercial sewing machines).

share). In other words, the net impact on jobs of the introduction of labour-saving devices is complex and difficult to ascertain *a priori* – it often requires ***ex post* assessment** after allowing some time for adjustments to occur.[7] The complexity of this second assessment is why **greater precision in assessing emerging capabilities** is the most promising empirical strategy.

A third type of empirical assessment that takes a national or regional unit of analysis involves taking account of labour-saving technologies on the **international division of labour** and hence on a country or region's pattern of international trade. In the latter two decades of the 20th century and well into the present century, companies often resolved make or buy decisions by outsourcing production to lower wage countries, off-shoring, that could meet product or service quality standards and an immense infrastructure of international logistics was constructed that facilitated the movement of parts and finished goods, and in some cases, provided the means to deliver services across national or regional boundaries. For example, containerised shipping is not generally considered part of automation, but if one thinks about past arrangements for merchandise trade, there were many labour saving implications of the growth of this transport method. For a variety of reasons including rising wages, the Covid-19 crisis, and perceived inequities to domestic workers in consuming countries, **automation is taken as an opportunity to 're-shore' production**. It is, however, also a means for off-shore producers to improve the productivity of workers and hence to compete with re-shoring trends. Like the second consideration, this is also a very complex process whose outcome is very difficult to predict. Similarly, however, the examination of labour-saving technologies in terms of **capabilities** is an important first step for an empirical analysis of impact.

To assess the potential for studying improving capabilities, it is helpful to look a bit more closely at the **history of labour-saving innovation**. Along with the periodic eruption of concerns about the automation of jobs, there is a continuing development of the waves of the past (von Tunzelmann, 1995). Thus, **mechanisation** continues to develop and indeed is influenced by **electrification** and what we have come to call **digitisation** influences both the previous waves of development. The interactive and cumulative effects are important. For example, early mechanisation was driven (literally) by **steam power** which dictated that plants would be vertically organised due to the constraints in distributing motive power horizontally. With electrification, the organisation of factories, and the nature of jobs could be transformed, first by extension and development of horizontal '**assembly lines**' and more recently a plethora of different '**work station**' and '**machine cluster**' groupings in which partially finished goods continue to be brought to workers and then transferred for additional work or final packaging.

The basic principles of mass production involve the creation of interchangeable parts that can be assembled rapidly by the movement of these parts to the assembly worker (through an assembly line or other means). Interchangeability requires manufacturing to precise tolerances and rapid

[7] Part of the problem here is that task efficiency is often easier to estimate than production system efficiency because the latter involves multiple changes that have synergistic effects.

movement and assembly operations will be facilitated by flexible arrangements for delivering mechanical power to the worker (e.g. pneumatic tools and motorised conveyors). The technologies for achieving all these functions were developed early in the 20th century and accompanied the surge in development following electrification.

The 20th century is an extended history of **industrial mass production displacing craft production** in which the individual worker does a variety of operations in fabricating and finishing a product. The basic logic of mass production has been extended to operations in **service industries** such as processing payments in banks, the use of a battery of power tools and other equipment in dentistry, or the surgical operating theatre. In many services, there are residual craft elements that continue to rely on the skill of the 'operator.' In some of these cases, the logic of mass production has been reversed so that the customer becomes the operator as is the case with automated teller machines (ATMs) which largely displace the role of bank clerks who previously received and dispensed cash (Savona and Steinmueller, 2013). The ATM operates under the control of the customer rather than the bank employee, substituting customer 'labour' for the labour of the bank clerk.

Mechanisation using electrical motors (hence **electrification**) was also progressively refined throughout the 20th century. In many cases this involved the redesign of products to make them more compatible with mechanised technique – milk bottles became tetra-packs, butchers parcelled out meat in shrink wrap packaging, and stockings were knit and finished with very little human intervention. The idea of **co-opting customer labour** was further refined by shipping products with 'some assembly required' and Ikea's refinement of 'flat pack' furniture that could be transported by the customer from large retail showroom/warehouses and assembled by the customer at home.

It is important to note that parts assembly-based mass production is not the only means by which industrial progress has been achieved. The evolution of technique stemming from the exploitation of petroleum as a source of fuel and as a feedstock for chemical manufacture involved several innovations that, together, created the petroleum refinery. Petroleum refineries are examples of continuous flow manufacturing in which human intervention is only required to monitor and control the flows of raw petroleum to finished products (e.g. petrol, diesel fuel, asphalt base, fuel oils, heating oil, paraffin and liquefied petroleum gas). In chemical production, a mixture of continuous flow and 'unit process' technologies are employed with the transfer of materials undertaken by pumps or, in some cases, conveyors. These basic operations have been extended to a variety of other industries such as food processing, ore refining, and pharmaceutical production.

What is important to observe about this brief account of technological history is that processes of automation involving **mechanisation** and **electrification** have been underway for an extended time. In many cases, labour saving innovations have already greatly improved individual worker productivity and most of the job losses in global North manufacturing have already occurred, not so much as the result of outsourcing of production, but by the combined influences of mechanisation and electrification.

Digitisation combined with international logistics and transport have continued this process. The ability to codify designs, arrange supply contracts, communicate about production issues, trace and monitor transport of parts and partially finished goods, and efficiently manage inventories in relation to the flow of production and consumption have all been greatly improved by digitisation.

The uncertainties that now confront economies throughout the world arise from the technological potentials of newer generations of **cyber-physical systems** that have the potential to transform the mass production paradigm. These new potentials involve new physical technologies such as additive manufacturing which is evolving from a substitute for hand machining and other forming processes for the tools and dies involved in mass production toward efficient direct production of final parts and products. If this trajectory were to continue, it is possible to imagine city or even neighbourhood-based '**manufactories'** capable of producing a very wide range of products for business and household use. Competing with this possibility is an adaptation of the '**fabless integrated circuit company**' which produces designs that can be used to create physical products in larger-scale flexible manufacturing facilities. The location of these larger-scale facilities are one of the greatest uncertainties concerning the future of international division of labour. For example, it is possible to imagine an initial period of development in which currently labour-intensive factories in middle income countries are automated with this new generation of technologies. Alternatively, since the new technologies will be **far less labour intensive**, it is possible to imagine **a major trend to reshoring of production**.[8]

In the countries that are currently most wealthy, the absolute dominance of service sector employment raises questions about the role of **new technologies in replacing these service sector jobs**. For example, the historical occupation of data entry operator has experienced dramatic reductions as new technologies for data acquisition including user data entry displace centralised facilities for data entry and filing, creating a 'data cloud' which itself offers a myriad of opportunities for the application of machine-learning AI to create predictive models and, increasingly, to manage data-intensive service provision. In many of these cases, the challenge is to improve the **human-computer interface** so that opportunities, choices and services can be presented in ways that customised the users' needs. The flexibility and scalability of robotic equipment has major implications for employment in 'customer facing' jobs if what is facing the customer is a cyber-physical system rather than a human being. As in past waves of automation, it seems likely that **the initial applications will be specialised to tasks that can be clearly specified** (e.g. picking items from warehouse shelves to fulfil an online order). However, impacts can cascade – in the same example, if packing and delivery also are performed by autonomous systems (e.g. robotic stocking and packing and autonomous vehicle delivery), employment implications become much more significant. Assessing the potential for the emerging new wave of automation, and whether this time it will be different, begins with a **careful assessment of the emergence of new capabilities in the**

[8] The recent disturbances in international supply chains associated with the Covid-19 pandemic have provided a motivation for reconsidering the international division of labour and rebalancing off-shoring with new domestic capabilities.

cyber-physical systems that are the current subjects of research, development and initial deployment.

The novelty of this exercise with respect to extant reviews of the literature on the labour impacts of automation, as hinted at above, is in the purposefully sought classification of automation technologies that is fine grained enough to allow digging in depth into the specific capabilities of each of them and the assessment of the extent these capabilities to replace or reconfigure labour tasks.

3 Methodology

This section and the related Appendix at the end of the document are purposely framed as a chronological journey that the research team at the UoS has endeavoured in to identify and refine in-the-making the literature search protocol. This allows us to articulate in detail the rationale underpinning our methodological choices and the solid grounding of the selected technical literature that will supports our findings. We intend this to become a methodological template and interpretative grid for any in-depth investigation of multidisciplinary literature that is able to shed light on such a complex phenomenon.

In particular, as we detail in the Appendix, we reconstructs how we dealt with the criteria of classification of automation technologies, from the initial scoping to the devising of a final classification that will serve as a standard ground for the whole PILLARS, particularly WP3 which deals with the new and emerging technologies; we then detail the complex search strategy of the relevant papers, from how the classification devised above informed the keyword strategy to the lengthy and scrupulous manual screening of the of the relevant papers to the identification of the relevant keywords that served as controlled terms to expand the initial selection of core papers.

We organized the literature review in **six main steps**, which are further detailed in the Appendix (Section 8).

First, given differences among technologies and their applications, we decided to run separate reviews for different families of digital automation technologies. We identified nine families:

A. **Robots –** *technologies that sense and (autonomously) act based on data*
B. Physical data acquisition technologies – *technologies that harvest and record information*
C. **Software based data management –** *technologies for storing, protecting, managing/handling and acquiring data*
D. Computing – *technologies used to compute/calculate*
E. AI (not directly as a cloud service) & Intelligent Information System – *technologies using algorithms and advanced methods to make sense out of the data*
F. Additive manufacturing (using any material – so e.g. powder metallurgy as well as bioplastic filament) – *technologies that produce bottom-up based on digital models*
G. Networking – *technologies for communicating between machines (data transmission) or connecting machines*
H. User interface – *technologies for human interaction with machines or data*
I. Other

In this paper we focus on two families, to study the difference between mainly manufacturing and mainly service automation technologies: robots (R) and software-based data management (DM).

Second, for each of the two families, R and DM, we identified the relevant records in the Scopus database (Figure 1). We used the literature and our own expertise to build a query to search titles, abstracts and keywords of publications. Because we are not interested in the technological

development per se, but in what the technology can do, and in particular what tasks it can perform, we built a query that combines keywords identifying the technology (e.g. robot OR human worker), its functions and applications (e.g. process OR routine), and tasks that it can perform (e.g. interact OR recognize OR weld) (detailed queries available in the Appendix). We selected only documents published after 2000, in the form of original articles, reviews, or conferences papers, and in the top percentiles by citations (by year). The number of percentiles varies between the two technologies and queries, to maintain a manageable number of documents to screen manually.

Third, to maximise precision in our selection of relevant papers, we manually screened this sample of papers (Figure 1). Each document was screened manually by two independent reviewers (title an abstract), and conflicting cases were decided by a third reviewer. Detailed screening rules are reported in the Appendix, but mainly consisted in filtering out papers that were not about the technology of interest, that were not about the production of goods and services (e.g. house appliances), which were only conceptual, or which did not clearly indicate what task the technology described could perform, even if when they clearly described their skills. For example, robots that can move in a small space, avoiding obstacles, but with no mention of their application to specific tasks, e.g. picking objects from shelves and bringing them to a different point of a warehouse.

Fourth, to increase recall, we sought to expand our initial query to documents in Scopus that, although not using terms in our initial query in their title and abstract (for example because they refer to tasks that we did not consider) may be relevant to the development of the technologies in the family under consideration. We took advantage of the Engineering Index Thesaurus, a thesaurus of controlled engineering terms that are manually allocated to documents in Scopus. Using the relevant documents selected after manual filtering (see previous step), we used text mining methods (TF-IDF) to select the engineering terms from the thesaurus that better identify these relevant papers (and the technologies they illustrate) (Figure 1). We combined engineering terms from the thesaurus that are ubiquitous, i.e. appear in documents related to different technologies (e.g. automation OR intelligent robots), and that relate to distinctive technologies (e.g. crops OR architectural design). We then used these engineering terms to run a second search in Scopus to identify papers that we did not find with our initial query, but which were assigned relevant engineering terms.

Fifth, to maximise precision in our selection of relevant papers, we manually screened this sample of papers as described in the third step (Figure 1).

Finally, as a result of manual screening of both first and expanded search, we were left with a number of selected documents per technology family (Figure 1), each referring to one or more technology. In our final step we read and coded the full text of the selected publications, extracting the following information about the technology, where available:

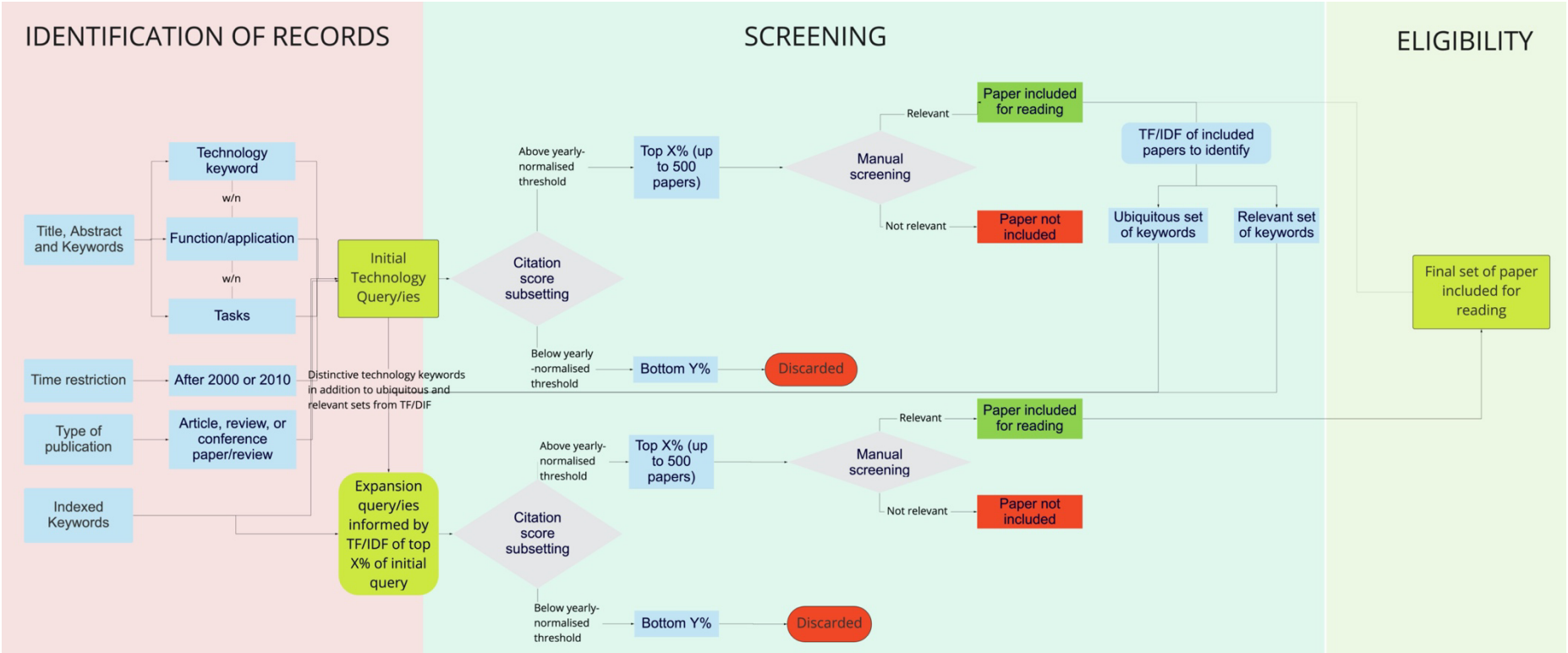
1. Level of adoption of the technology in the industry
2. Development stage of the technology
3. Routinisation: the ability to perform a task without any human intervention

4. Knowledge codification: the ability to make all instruction explicit (i.e. codified)
5. Whether the technology works with people, symbols or objects (Reich, 1991)
6. Level of skills required to use the technology
7. Whether the technology substitutes or complements human tasks
8. Whether the technology improves the product/service, or the process to produce products/services
9. Intended sectors of application
10. Intended tasks that the technology performs
11. Geographical area of development/use
12. Type of organization that is likely to use the technology
13. Size of the organization that is likely to use the technology

We describe all these steps in detail in the Appendix (Section 8).

The detailed reading and the coding allows to describe, for each technology family, whether the technologies identified in the literature potentially substitute, complement, and/or reconfigure specific technical tasks within occupations and sectors; and what are these specific tasks across sectors, and which are the specific industries are exposed to specific automation technologies and routinisation of tasks. The codification allows us to put us some numbers against our discussion of the literature, which provide some qualitative understanding of tasks, sectors, and type of impact that these different technologies may have on employment. When relevant, we split results into two sub-periods, according to the year of publication, before and after 2010. This is expected to give some indications on whether there has been any changes in the intended use and potential impacts of digital automation technologies.

Figure 1: Flow-chart of the literature review process



4 Results: A quantitative and qualitative analysis of technical papers

This section illustrates and discusses the results of the review of the extracted records of technical papers. The analysis organises the collected information into an interpretative grid, with the aim of producing a general scheme to assess impacts of fine-grained-defined technologies. Hence, the results come from a 'meta' analysis in the sense of data-driven construction of a typology, though not a conventional statistical meta-analysis. More precisely we codify the records according to a set of characteristics, which are singled-out in the literature and that capture essential features of the particular instantiations of the two technologies we study. We then look at frequencies of papers focused on the specific technology – as identified and classified above – across the relevant job-related characteristics identified below. The Appendix reports all the categories of classification for each of these characteristics (see tables Table 30-A and Table 31-A).

- ***Emerging technologies, exposure and adoption:***

1. **Level of Adoption of the technology.** Technologies experience different levels and patterns of adoption, which depend on an array of variables ranging from actors' features, propensities and thresholds, to the structures of interconnections in adopters' networks up to the very technical features of the technologies (Rogers, 2010). In particular emerging technologies (Rotolo et al., 2015), especially when articulated and complex, can show different adoption levels for different sub-technologies. In our work, we classify this variable in low, medium and high, depending on whether the technology is a prototype or displays a higher level of maturity to be adopted.
2. **Development (function design) stage.** This can be at the stage of invention; conceptual definition; or, more downstream, at the experimental level (prototype); definition of product; at the stage of ready to deploy; or mature. We use this variable to capture the level of maturity of the technologies (Albert, 2016) and to map in a stylised manner the stages of new product (or process) development (Takeuchi and Nonaka, 1986).

- ***Emerging technologies and tasks reconfiguration:***

3. **Task/routinisation.** Technologies execute functions at the level of single operations or activities, with different degrees of autonomy. This ability impacts the very structuring of how produces and services are made, i.e. the 'factory physics' of production (Hopp and Spearman, 2011); the reach of automation and its particular deployment depends on how much given tasks can be routinised (and the relevant knowledge for the activities to be executed codified – see below for another variable capturing this), which in turns also depends on how single

operations can be separated or consolidated (Combemale et al., 2021). For our purposes, we classify the technology on the basis of the ability to perform a task without any human intervention, including the possibility that a task is further decomposed into an automated and humanly supervised segments.

4. **Task/knowledge codification.** This variable classifies the technology on the basis of the ability to make all instruction explicit (i.e. codified) without the use of any tacit knowledge. The capability to develop and deal with tacit know-how is a cornerstone of the economics of knowledge (Cohendet and Steinmueller, 2000; Cowan et al., 2000; Foray and Steinmueller, 2003) and so far usually considered a monopoly of human action.
5. **Works with People/Symbols/Objects.** This characteristics is borrowed by the taxonomy suggested by Robert Reich (1991) and considers whether a technology deals mainly with people, integrating the use of objects or targeting the use of symbols. This simple categorisation has a very interesting (potential) explanatory power, as it allows to assess whether recent emerging technologies such as artificial intelligence algorithms, that seem to be lowering prediction costs in tasks usually involved in symbol manipulation, produce a similar impact compared to, for instance, robots.

- ***Emerging technologies and employment compensation:***

6. **Employment compensation: Skills.** The idea that technologies display different degrees of complementarity with humans along the skills distribution has been widely explored by the literature on skill-biased technological change, and refined by the approach of routine-biased technical change (Autor 2015; 2019; Ciarli et al., 2021; Goldin et al., 2020). As we identify the routinisation feature of technologies in variable (3), with this variable to single-out the direct relationship with skills. In our codification, technologies might be employed by low, medium or high skilled humans.
7. **Employment compensation: complement or substitute.** Technologies might (traditionally) complement or replace the use of human labour, producing what have been labelled respectively a productivity or substitution effect (Acemoglu and Restrepo, 2019). Recent work has been identifying the conditions and the mechanisms for the decrease of the labour share (hence, the substitution of human labour by capital) both at the aggregated and firm level (Ray and Mookherjee, 2021; Acemoglu et al., 2020). We record papers with a dummy indicator complement or substitute; in addition, we consider the possibility to identify segments of tasks (or sub-tasks) that are replaced and others within the same task that are complemented. In these cases records are identified as 'substitute; complement'.
8. **Employment compensation: Time saving or product/process innovation**
   The many sorts of automation technologies we cover in the analysis intervenes in different areas of firms' activities; some are process improvement aiming at restructuring production,

provide efficiency gains such as cost reduction or time saving, others are embedded into the design of new product offerings. We record this information to understand the direction of change within economic actors, loosely referring to Savona and Steinmueller (2013) for services.

- ***Emerging technologies and geo-sectoral domain of application:***

9. **Geographical area of provenience:** This codes the record on the basis of where the technology has been deployed.[9] Table 30-A in the Appendix includes all values.
10. **Sector of application.** This variable organises technologies according to the main sector of use based on 3-digit ISIC classification for manufacturing and 2-digit for other sectors (see Table 4 in Section 4.1.2 below); it is usefully cross-referenced with the tasks of application (point 11 below).
11. **Task of application.** This variable classifies technologies on the basis of what work task(s) it is likely to replace or be integrated into. We code these using O*NET aggregate work activities list. Table 31-A in the Appendix includes all values.

This type of analysis (qualitative reading and coding in a pre-structured grid) allows us to conduct a quantitative analysis (frequencies of occurrence) of the number of records that can be classified along each sub-category, both in terms of row-share of number of records and along the time dimension. With the filled table, we can explore an array of cross-tabulations to assess, for example, how frequencies in technology-, skills-, or task-related variables are distributed across sectors (and vice versa.

4.1 Robots-based Automation and Software-based Data Management: Analysis of Included Technical papers

In this sub-section, we focus on two major technologies (A, Robots and C, Software-based Data Management).

If, as we claim, the impact of automation technologies on labour is heterogeneous and depends on the specific, fine-grained technology considered, we should be able to detect fundamental

[9] This dimension is hard to codify as often the place/country of first development or application of the technology is not explicitly mentioned or deducible. When this is the case, the geographical area is attributed through the affiliation of the author. Because of this, we have not taken into consideration this category when looking at the results.

differences in our overview analysis of technical publications describing technical features of these technologies.

For instance, this heterogeneity should appear when computing cross-tabulations of (publications focused on) which tasks the technologies execute in which particular sector. Indeed, an interesting part of our analysis focuses on the concentration of tasks per each sector (Table 7) and how a task is concentrated (or spread) across sectors (Table 9). This is preliminary information to infer the degree of exposure of sectors and tasks to specific automation technologies. In addition, we compare other features of the technologies by task and sector, including the complementary/substitutive relationship with labour, the degree of knowledge codification, and the type of interaction of the technologies, whether with people, things, or symbols.

Let us first recall the granular list of emerging technologies included in the main category of Robots and Robot-related Automation, which has dictated the choice of keywords embodied in the queries (see previous section and Appendix for details). This includes machines that contribute to real time monitoring, self-driving, process automation robots, automated platforms; it includes semi-autonomous robots, service robots, co-bots, swarm robots. In sum, the technologies we explore in the literature are both what Sheridan (2016, p.525) labels *telerobots*, namely machines "capable of carrying out a limited series of actions automatically, based on a computer program, and capable of sensing its environment and its own joint positions and communicating such information back to a human operator who updates its computer instructions as required", and *teleoperators*, that is machines that "perform manipulation and mobility tasks in the remote physical environment in correspondence to continuous control movements by the remote human." To provide a few examples of the variety of technologies in development or use, in the analysis we reviewed studies on robotic arms, or welding robots, but also robot vehicles for the exploration of underwater mines as well as dismissed nuclear power plants. All of these robots might be adopted and used in applications that vary from bar tendering to very complex surgical operations, within which they replace or complement tasks that might be more or less routinised and involve the use of higher or lower skills or complex knowledge or very tacit knowledge.

The granular list of Software-based Data Management technologies includes Database systems, Relational Databases, Cryptography, security, blockchain and Big Data. Again, this has dictated the choice of the keywords included in the query (See Table 7 for details).

4.1.1 Maturity of technology development and exposure

As a starting point, we look at the distribution of R and DM technologies across different stages of functions' maturity, according to our paper coding.

Table 2 and Table 3 below show the absolute numbers and shares of technical papers that describe the development of robots and data management at different levels of technology maturity, from a conceptual stage to maturity of application, over the whole time span selected and across the two sub-periods, before and after 2010.

The large majority of papers describes robots at the experimental stage (almost two thirds), with only a minority at the mature stage. The prototypes and the conceptual stages have similar shares, interestingly regardless of the time span considered (e.g. before or after 2010). This confirms the relevance of the selection choice, which privileges the technologies which are currently emerging and have not yet been fully deployed. This reflects also the fact that engineering and technical publications over the last two decades are concentrated in novel applications of robotisation, mainly at the experimental stage, which most likely requires the academic community recognition to move forward towards downstream stages of development.

Interestingly, the numbers and shares of papers describing data management technologies is slightly less concentrated in the experimental category, with a higher presence in the mature stage of application, despite the large majority of papers in DM have been published after 2010. This might be linked to the fact that the adoption and use of data management techniques, particularly those that are software-based, have a higher pervasiveness in terms of range and sectors of applications than the robots' applications. This might imply that the degree of novelty of publications is not necessarily correlated to the stage of technology development and the technical features of the robot-related automation, but in the domain of (service) use, rather than the technical feature of the software application used for data management (Saviotti and Metcalfe, 1984).

In either case, both sets of papers describe emerging applications of R and DM applications, the latter being more spread across sectors – as discussed in next section.

The large concentration of papers in the experimental stage of technological development might imply an idiosyncratic boundary of the tasks and labour functions actually affected by these technologies, due to their degree of diffusion and sectoral exposure being by definition still embryonal. We will focus on this in the next section.

Table 2: Technological Development classification of robot papers

	Conceptual	Experimental	Prototype	Ready-to-deploy	Mature	NA
Full (n=165)	0.10	0.59	0.19	0.08	0.03	0.01
<2010 (n=57)	0.10	0.57	0.17	0.06	0.00	0.10
>2010 (n=108)	0.09	0.57	0.19	0.09	0.05	0.01

Table 3: Technological Development classification of Data Management papers

	Conceptual	Experimental	Prototype	Ready-to-deploy	Mature	NA
Full (n=123)	0.17	0.43	0.15	0.13	0.12	0
<2010 (n=24)	0.17	0.29	0.17	0.25	0.12	0
>2010 (n=99)	0.17	0.46	0.14	0.10	0.12	0

In sum, based on the papers' coding, and looking at development stage, DM technologies seem more mature than R ones. This is reflected in the adoption rates, with R featuring studies that describe technologies with low adoption. An interpretation of the finding is that DM technologies belong to a technological trajectory that is established and ongoing – so that improvement and novelties described in the papers are incremental or they layer and integrate seamlessly on the existing technology 'platform'. Instead, robots tend to be in a different phase of the technological trajectory, with most publications focusing on novel equipment and devices. This is important because, as mentioned in our historical introductory section, current robotics is likely different from earlier instances of mechanisation. While they stand in continuity with processes of industrial automation, their flexibility or malleability (Wirkierman, 2022) means they have potential for broader and deeper impact. The fact that with our data we capture less mature robots is an indication that we are identifying the most recent wave of this technology.

4.1.2 Concentration of sectors in tasks and tasks in sectors

Table 4 reports a synthesis of the ISIC-rev4 classification of Aggregate Economic Activities that we use to classify the sector where the technologies are conceived, developed, adopted and that underpins the papers coding process in what follows.

Table 4: ISIC-rev4 classification of Aggregate Economic Activity (1-digit)

Aggregate Economic Activity			Sections ISIC- Rev. 4
Agriculture			A
Non Agriculture	Industry	Manufacturing	C
		Construction	F
		Mining and quarrying; Electricity, gas and water supply	B, D, E
	Services	Market Services (Trade; Transportation; Accommodation and food; and Business and administrative services)	G, H, I, J, K, L, M, N
		Non-market services (Public administration; Community, Social and other services and activities)	O, P, Q, R, S, T, U
Not elsewhere classified			X

Looking at where the two technologies are adopted and used, we find that the three ISIC sectors ranking highest in terms of applications of technologies are Professional, scientific and technical activities (M), Information and communication (J), and Manufacturing (C) for DM, with sector M appearing in more than 30 percent of the studies we reviewed. The ranking is persistent over time, when comparing the full sample and the literature published in the more recent period (after 2010). For what concerns R, the top-three sectors are Professional and Scientific Activities (M), Manufacturing (C), and Human health and social work activities (Q). Also in this case the ranking is stable, but M decreases in relative frequency while C increases after 2010. This seems to map the limited but growing uptake of robots from the science domain into the 'wild' of industrial and commercial activities (Benmelech and Zator, 2022).

Having an idea of the broad differences in development stage and adoption across sectors, we can move to a first coarse-grained comparison between the main tasks the two technologies are used for as they appear in our selected papers. In Table 5, we report the top-ten tasks for R and DM based on the full sample (hence including papers published from 2000 onwards). This comparison gives already a sense of the overlap (or lack of overlap) in the functions executed by the two technologies. Based on our analysis, only four tasks are present both in R and DM, and some such as object identification rank rather differently in R versus DM. This suggest that the two technologies are structurally different in the 'service' they perform, however with an overlap which uncovers the fundamental common knowledge base of being technology systems based on computing.

Table 5: Ranking of tasks frequency across Robots and Data Management technical papers

Rank	Robots (R)	Data Management (DM)
1	Handling and Moving Objects	Processing information
2	Identifying Objects, Actions, and Events	Analysing Data or Information
3	Performing General Physical Activity	Getting Information
4	Getting information	Scheduling Work and Activities
5	Inspecting equipment, structures, Material	Documenting/Recording Information
6	Assisting and caring for others	Monitor Processes, Materials, Surroundings
7	Monitor Processes, Materials, Surroundings	Making Decisions and Solving Problems
8	Controlling Machines and Processes	Monitoring and Controlling Resource
9	Operating Vehicles, Mechanized Devices	Judging the Qualities of Things, Services, People
10	Analyzing Data or Information	Identifying Objects, Actions, and Events

Next to the raking of tasks, we can provide some insight on their concentration across economic sectors, which can shed light on the variety of sectoral exposure of the technologies.

We look at sectoral and tasks exposure to the two technologies. The top-three tasks in R account for 50 percent of the literature mentions' frequency; the percentage rises to 64.6 percent for DM. This suggests that R has a higher penetration rate across functions. The top-three tasks for robots appear in, respectively, 13, 13, and 10 sectors, while the top-ranking ones for DM appear in 16, 17, and 12 sectors respectively. In a nutshell, contemporary robots, as hardware technology, are multi-purpose in nature, and malleable fixed capital. Different sectors tend to use them mostly for moving and performing physical activities, but also in tasks that require control of operations, processes and other device. The presence of object identification among the main tasks executed by robots hints at that robots are increasingly integrated in smart systems (cyber-physical systems) that combine motion capabilities with complementary technologies for the acquisition of data, such as sensors, which will be analysed among the remaining technologies of our PILLARS classification.

In Table 6, we cross-tabulate tasks and sectors, indicating for each sector (columns) the top-three tasks for each of the two technology (R, DM). The aim is to see whether there is an overlap in specialisation; rather, even though we present only the top tasks, we obtain confirmation of the fact that the two technologies are structurally different.

Table 6: Cross tabulation of tasks and sector in both robot and data-management technologies

Task/Sector	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	Q	R	S	T
Handling and moving objects	R	R	R		R	R		R		R			R			R			
Inspecting Equipment, Structures	R	R			R	R		R						R					
Identifying Objects, Actions, and E		R	R					R		R			R	R	R	R		R	R
Controlling machines and processes			R																
Performing General Physical Activity	R			R		R		R						R					R
Monitor process, materials, or surroundings	DM				R	DM		DM				DM							
Assisting and caring for others									R							R	R	R	
Getting information	DM		DM		DM								R; DM	DM	DM	DM		DM	
Analyzing Data or Information	DM	DM	DM	DM	DM	DM		DM		DM	DM		DM	DM	DM	DM		DM	
Processing information	DM	DM	DM	DM	DM	DM	DM	DM		DM	DM	DM	DM	DM	DM	DM		DM	DM
Interpreting the meaning of information		DM																	
Scheduling Work and Activities				DM						DM									
Documenting/recording information																		DM	

Note: we exclude a task if its frequency in the full sample is <1, unless the sector has only one mention (in that case we present the data to provide a broad idea of which task is at least executed in the sector).

Table 7 provides the full view of the task-sector cross-tabulation, the one displaying the number, shares and concentration of tasks executed in each sector. This representation captures the sectoral nature of our data. We report for each sector the number of unique tasks (UniqT), the total number of tasks (TotT) (as a task can appear multiple times, in different papers assigned to the same sector), the share and the concentration index of tasks for each sector (HHI). Each sector shows a varying degree of task concentration, from sector T (personal and social services) with max concentration of 3 unique tasks, to sector A (Agriculture) with the lowest concentration of unique tasks (15 out of a total number of tasks performed of 31). The uniqueness of tasks within a sector can be considered as a rough indicator of the degree of exposure of a sector to fully-fledged automation across the task-spectrum, which should occur in particular when the tasks are routinised. Conversely, a low concentration of tasks within each sector means that a higher number of tasks remain executed by humans, either because their specific mode of execution in a particular sector is not prone to automation or because routinisation and knowledge codification has not yet progressed enough in the context of that sector-task tandem. In sectors with lower robot-powered task concentration, existing operational practices, barriers or bottlenecks might provide protection from the risk of being fully exposed to automation-related substitution. Furthermore, the large gap between UniqT and TotT for sectors C and M indicates that the major difference between these and

other sectors is not necessarily the span of activities that robots can execute; rather, it is the scale of use of robots in these tasks, that appears repeatedly in a large number of papers – evidence of the fact that several research efforts are directed at such technology.

In most sectors, the task uniqueness refers to the period post 2010, exception made for sector M (business services), which shows quite a low concentration of tasks across both periods, though lowest over the latest decade. This might also mean that over time, the effects of robotization have not decreased the number of tasks (that is increased concentration) due to automation.

Table 7: Concentration of Tasks per Sector – robot papers

	Full (n=165)				<2010 (n=57)				>2010 (n=108)			
TaskShort	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI
A ~ Agriculture, forestry and fishing	15	31	0.48	0.093	5	5	1	0.2	14	26	0.54	0.097
B ~ Mining and quarrying	6	10	0.6	0.2	1	1	1	1	5	9	0.56	0.234
C ~ Manufacturing	20	153	0.13	0.134	9	33	0.27	0.188	17	120	0.14	0.131
D ~ Electricity, gas, steam [...]	4	4	1	0.25	NA	NA	NA	NA	4	4	1	0.25
E ~ Water supply; sewerage, waste management [...]	5	7	0.71	0.224	NA	NA	NA	NA	5	7	0.71	0.225
F ~ Construction	5	11	0.45	0.256	1	2	0.5	1	5	9	0.56	0.234
H ~ Transportation and storage	7	17	0.41	0.183	6	9	0.67	0.185	5	8	0.62	0.25
I ~ Accommodation and food service activities	5	8	0.62	0.312	2	2	1	0.5	4	6	0.67	0.334
J ~ Information and communication	4	4	1	0.25	2	2	1	0.5	2	2	1	0.5
M ~ Professional, scientific and technical activities	17	105	0.16	0.196	11	54	0.2	0.208	13	51	0.25	0.188
N ~ Administrative and support service activities	11	22	0.5	0.132	5	9	0.56	0.234	9	13	0.69	0.136
Q ~ Human health and social work activities	13	38	0.34	0.136	7	17	0.41	0.19	12	21	0.57	0.139
R ~ Arts, entertainment and recreation	5	6	0.83	0.222	1	1	1	1	5	5	1	0.2
S ~ Other service activities	6	17	0.35	0.253	5	12	0.42	0.278	3	5	0.6	0.36
T ~ Activities of households as employers [...]	3	3	1	0.333	NA	NA	NA	NA	3	3	1	0.333
NA ~ NA	6	9	0.67	0.185	1	1	1	1	6	8	0.75	0.188
Tot.	132	445	9.25	3.359	56	148	9.03	6.483	112	297	10.66	3.799

Table 9 reports the task-sector cross-tabulation, in terms of the number, shares and concentration of sector for each task. In particular with the HHI index here we aim at characterising the extent to which a certain task is ubiquitous and therefore the extent to which its potential automation would affect the economy at large. Some tasks become expectedly more diffused across sectors over time (e.g. Data analysis decreases its concentration over time); others instead become more concentrated, such as Machine control, which most likely gets substituted by robots over time. Some tasks are structurally highly concentrated and remain so over time, such as Coaching and Developing others. The most ubiquitous task – Handling and Moving Objects – is used in over 100 sectors and is unique to several sectors; interestingly it does not seem to change concentration over time. Similarly to a further, rather routinised and ubiquitous task, Identifying Object and Actions, remains fairly spread over time.

Looking at DM-related publications (Table 10) the most relevant tasks in absolute terms (getting, processing, and analysing information) are also those with the lowest concentration index – though more ubiquitous across sectors. This reflect the general-purpose nature of information processing, as the technologies overviewed in this domain are the most recent wave of established ICT techniques (Bresnahan and Yin, 2017)

In general, there is a large gap in terms of task concentration (uniqueness in specific sectors) between physical and non-physical tasks, most likely concentrated in several manufacturing

sectors rather than in a few services. The non-physical ones – such as administration, information processing, scheduling activities - are comparatively less diffused.

Table 8: Concentration of Tasks per Sector – DM papers

	Full (n=123)				<2010 (n=24)				>2010 (n=99)			
Sect	UniqT	TotT	Share	HHI	UniqT	TotT	Share	HHI	UniqT	TotT	Share	HHI
A	9	14	0.64	0.133	4	4	1.00	0.250	8	10	0.80	0.140
B	3	3	1.00	0.333	NA	NA	NA	NA	3	3	1.00	0.333
C	11	64	0.17	0.199	4	12	0.33	0.320	11	52	0.21	0.182
D	5	10	0.50	0.260	NA	NA	NA	NA	5	10	0.50	0.260
E	3	6	0.50	0.333	3	6	0.50	0.333	NA	NA	NA	NA
F	9	22	0.41	0.169	NA	NA	NA	NA	9	22	0.41	0.169
G	1	1	1.00	1.000	1	1	1.00	1.000	NA	NA	NA	NA
H	6	10	0.60	0.200	NA	NA	NA	NA	6	10	0.60	0.200
J	15	96	0.16	0.171	8	19	0.42	0.202	13	77	0.17	0.181
K	7	9	0.78	0.160	2	2	1.00	0.500	7	7	1.00	0.143
L	2	2	1.00	0.500	NA	NA	NA	NA	2	2	1.00	0.500
M	20	155	0.13	0.168	10	47	0.21	0.164	16	108	0.15	0.182
N	8	13	0.62	0.172	3	3	1.00	0.333	8	10	0.80	0.160
O	4	7	0.57	0.265	4	4	1.00	0.250	3	3	1.00	0.333
Q	9	15	0.60	0.156	NA	NA	NA	NA	9	15	0.60	0.156
S	4	4	1.00	0.250	NA	NA	NA	NA	4	4	1.00	0.250
T	7	8	0.88	0.156	NA	NA	NA	NA	7	8	0.88	0.156
NA	2	2	1.00	0.500	NA	NA	NA	NA	2	2	1.00	0.500

Table 9: Concentration of sectors per Task – Robot papers

	Full (n=165)				<2010 (n=57)				>2010 (n=108)			
TaskShort	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI
Coaching and Developing Others	1	1	1	1	NA	NA	NA	NA	1	1	1	1
Establishing and Maintaining Interp	1	1	1	1	NA	NA	NA	NA	1	1	1	1
Estimating the Quantifiable Charact	1	1	1	1	NA	NA	NA	NA	1	1	1	1
Evaluating Information to Determine	1	1	1	1	NA	NA	NA	NA	1	1	1	1
Interacting With Computers	1	2	0.5	1	NA	NA	NA	NA	1	2	0.5	1
Monitoring and Controlling Resource	1	1	1	1	NA	NA	NA	NA	1	1	1	1
Performing Administrative Activitie	1	2	0.5	1	NA	NA	NA	NA	1	2	0.5	1
Performing for or Working Directly	1	1	1	1	1	1	1	1	NA	NA	NA	NA
Scheduling Work and Activities	1	3	0.33	1	NA	NA	NA	NA	1	3	0.33	1
Controlling Machines and Processes	3	21	0.14	0.678	3	4	0.75	0.375	2	17	0.12	0.889
NA	2	3	0.67	0.556	NA	NA	NA	NA	2	3	0.67	0.556
Communicating with Supervisors, Pee	2	4	0.5	0.5	2	4	0.5	0.5	NA	NA	NA	NA
Provide Consultation and Advice to	2	2	1	0.5	NA	NA	NA	NA	2	2	1	0.5
Training and Teaching Others	2	6	0.33	0.5	NA	NA	NA	NA	2	6	0.33	0.5
Operating Vehicles, Mechanized Devi	5	19	0.26	0.385	NA	NA	NA	NA	5	19	0.26	0.385
Developing Objectives and Strategie	3	4	0.75	0.375	2	3	0.67	0.556	1	1	1	1
Judging the Qualities of Things, Se	3	4	0.75	0.375	NA	NA	NA	NA	3	4	0.75	0.375
Organizing, Planning, and Prioritiz	3	4	0.75	0.375	NA	NA	NA	NA	3	4	0.75	0.375
Documenting/Recording Information	3	5	0.6	0.36	1	1	1	1	2	4	0.5	0.5
Handling and Moving Objects	13	101	0.13	0.285	9	38	0.24	0.317	11	63	0.17	0.294
Assisting and Caring for Others	6	26	0.23	0.284	5	13	0.38	0.314	5	13	0.38	0.302
Making Decisions and Solving Proble	4	5	0.8	0.28	2	2	1	0.5	3	3	1	0.333
Getting Information	9	39	0.23	0.273	4	14	0.29	0.337	9	25	0.36	0.277
Interpreting the Meaning of Informa	5	7	0.71	0.265	1	2	0.5	1	5	5	1	0.2
Identifying Objects, Actions, and E	15	81	0.19	0.218	10	34	0.29	0.227	12	47	0.26	0.258
Monitor Processes, Materials, or Su	7	16	0.44	0.211	2	5	0.4	0.68	6	11	0.55	0.273
Analyzing Data or Information	7	12	0.58	0.208	2	5	0.4	0.52	6	7	0.86	0.184
Processing Information	6	8	0.75	0.188	2	2	1	0.5	4	6	0.67	0.278
Inspecting Equipment, Structures, o	9	25	0.36	0.149	2	3	0.67	0.556	9	22	0.41	0.145
Performing General Physical Activit	14	40	0.35	0.134	8	17	0.47	0.19	12	23	0.52	0.138
Tot.	132	445	17.85	16.099	56	148	9.56	8.572	112	297	17.89	15.762

Table 10: Concentration of sectors per Task – DM papers

	Full (n=123)				<2010 (n=24)				>2010 (n=99)			
TaskShort	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI	UniqS	TotS	Share	HHI
Analyzing Data or Information	16	87	0.18	0.200	8	27	0.30	0.248	15	60	0.25	0.186
Communicating with Supervisors, Pee	2	3	0.67	0.556	2	3	0.67	0.556	NA	NA	NA	NA
Controlling Machines and Processes	1	1	1.00	1.000	NA	NA	NA	NA	1	1	1.00	1.000
Documenting/Recording Information	8	24	0.33	0.302	3	9	0.33	0.630	8	15	0.53	0.191
Estimating the Quantifiable Charact	1	2	0.50	1.000	1	2	0.50	1.000	NA	NA	NA	NA
Evaluating Information to Determine	2	3	0.67	0.556	NA	NA	NA	NA	2	3	0.67	0.556
Getting Information	12	72	0.17	0.274	7	16	0.44	0.250	11	56	0.20	0.288
Identifying Objects, Actions, and E	4	8	0.50	0.344	NA	NA	NA	NA	4	8	0.50	0.344
Inspecting Equipment, Structures, o	4	4	1.00	0.250	1	1	1.00	1.000	3	3	1.00	0.333
Interacting With Computers	4	5	0.80	0.280	NA	NA	NA	NA	4	5	0.80	0.280
Interpreting the Meaning of Informa	3	3	1.00	0.333	NA	NA	NA	NA	3	3	1.00	0.333
Judging the Qualities of Things, Se	4	9	0.44	0.383	1	1	1.00	1.000	3	8	0.38	0.469
Making Decisions and Solving Proble	9	13	0.69	0.160	1	1	1.00	1.000	8	12	0.67	0.181
Monitor Processes, Materials, or Su	10	20	0.50	0.135	2	4	0.50	0.500	10	16	0.62	0.117
Monitoring and Controlling Resource	5	11	0.45	0.306	2	3	0.67	0.556	5	8	0.62	0.312
Organizing, Planning, and Prioritiz	3	6	0.50	0.389	NA	NA	NA	NA	3	6	0.50	0.389
Performing Administrative Activitie	2	2	1.00	0.500	NA	NA	NA	NA	2	2	1.00	0.500
Performing General Physical Activit	4	5	0.80	0.280	1	1	1.00	1.000	3	4	0.75	0.375
Processing Information	17	126	0.13	0.196	9	28	0.32	0.230	15	98	0.15	0.194
Provide Consultation and Advice to	1	2	0.50	1.000	1	2	0.50	1.000	NA	NA	NA	NA
Scheduling Work and Activities	10	32	0.31	0.252	NA	NA	NA	NA	10	32	0.31	0.252
Selling or Influencing Others	1	1	1.00	1.000	NA	NA	NA	NA	1	1	1.00	1.000
Staffing Organizational Units	1	1	1.00	1.000	NA	NA	NA	NA	1	1	1.00	1.000
NA	1	1	1.00	1.000	NA	NA	NA	NA	1	1	1.00	1.000

4.1.3 Knowledge codification, routinization, complementarity and substitutability of tasks

This sub-section focuses on some of the most important features of the impact of automation on labour. It summarises and discusses the results of the codification of technical papers on the basis of whether robots-related automation and data management are complementary or substitute human activities, and whether the tasks complemented or replaced are routinized and based on codified (or tacit) knowledge. These features are considered together, as much of the seminal economic literature has found that automation is most likely to replace routinised and codified tasks while it tends to complement non routinized tasks or those based on tacit knowledge, which remain pertinence of human execution.

Here, again valuing the extent to which we can fine grain this evidence, we try to unpack **which** technology, sectors and tasks underpin the interpretation of findings.

The Tables below report the number and row shares of papers describing the tasks being based on tacit and non-tacit (i.e. codified) knowledge, respectively for R and DM papers, published in the pre and post 2010 periods.

The large majority of papers focuses on tasks related to identifying and handling objects, which require comparatively less codified knowledge. The number of publications that describe information-related handling tasks or communication are relatively less frequent and require less codified knowledge.

The shares roughly change in the subsequent period, when the papers focus on a higher variety of tasks performed and a higher number of tasks related to information processing rather than object handling, including tasks of controlling or supervising machines.

Prototypes and initial concept development related to robots seem therefore over time be covering tasks that are in principle **less codified**, though presumably **complementary** to more traditional and codified tasks related to object handling.

After 2010, technical literature on R seems to have shifted focus, and covered mainly tasks of handling and moving objects, but also operating vehicles, getting information and inspecting equipment and structures, all of which require **more codified knowledge and are more routinized**.

Over time publications seem to privilege a larger variety of prototypes of robot related automated tasks which are seemingly technically incorporating more codified knowledge. In sum, more tasks are exposed to robotization, that requires more codified knowledge.

Table 11: Task/Knowledge Codification interaction (<2010) - Robot papers

	Absolute frequencies			Row shares		
TaskShort	Yes	No	NA	Yes	No	NA
Interpreting the Meaning of Informa	2	0	0	100.0	0.0	0.0
Performing for or Working Directly	1	0	0	100.0	0.0	0.0
Controlling Machines and Processes	3	1	0	75.0	25.0	0.0
Developing Objectives and Strategie	2	1	0	66.7	33.3	0.0
Assisting and Caring for Others	8	5	0	61.5	38.5	0.0
Monitor Processes, Materials, or Su	3	2	0	60.0	40.0	0.0
Making Decisions and Solving Proble	1	1	0	50.0	50.0	0.0
Performing General Physical Activit	7	9	1	41.2	52.9	5.9
Identifying Objects, Actions, and E	12	22	0	35.3	64.7	0.0
Inspecting Equipment, Structures, o	1	2	0	33.3	66.7	0.0
Handling and Moving Objects	11	27	0	28.9	71.1	0.0
Getting Information	4	10	0	28.6	71.4	0.0
Analyzing Data or Information	1	4	0	20.0	80.0	0.0
Communicating with Supervisors, Pee	0	4	0	0.0	100.0	0.0
Documenting/Recording Information	0	1	0	0.0	100.0	0.0
Processing Information	0	2	0	0.0	100.0	0.0

Table 12: Task/Knowledge Codification interaction (>2010) - Robot papers

	Absolute frequencies			Row shares		
TaskShort	Yes	No	NA	Yes	No	NA
Coaching.and.Developing.Others	1	0	0	100.0	0.0	0.0
Documenting.Recording.Information	4	0	0	100.0	0.0	0.0
Establishing.and.Maintaining.Interp	1	0	0	100.0	0.0	0.0
Evaluating.Information.to.Determine	1	0	0	100.0	0.0	0.0
Interacting.With.Computers	2	0	0	100.0	0.0	0.0
Monitor.Processes..Materials..or.Su	11	0	0	100.0	0.0	0.0
Monitoring.and.Controlling.Resource	1	0	0	100.0	0.0	0.0
Performing.Administrative.Activitie	2	0	0	100.0	0.0	0.0
Provide.Consultation.and.Advice.to.	2	0	0	100.0	0.0	0.0
Scheduling.Work.and.Activities	3	0	0	100.0	0.0	0.0
Controlling.Machines.and.Processes	15	2	0	88.2	11.8	0.0
Analyzing.Data.or.Information	6	1	0	85.7	14.3	0.0
Operating.Vehicles..Mechanized.Devi	16	3	0	84.2	15.8	0.0
Getting.Information	20	5	0	80.0	20.0	0.0
Making.Decisions.and.Solving.Proble	2	1	0	66.7	33.3	0.0
Interpreting.the.Meaning.of.Informa	3	2	0	60.0	40.0	0.0
Handling.and.Moving.Objects	34	28	1	54.0	44.4	1.6
Assisting.and.Caring.for.Others	7	6	0	53.8	46.2	0.0
Inspecting.Equipment..Structures..o	11	11	0	50.0	50.0	0.0
Organizing..Planning..and.Prioritiz	2	2	0	50.0	50.0	0.0
Processing.Information	3	3	0	50.0	50.0	0.0
Identifying.Objects..Actions..and.E	22	25	0	46.8	53.2	0.0
Performing.General.Physical.Activit	10	13	0	43.5	56.5	0.0
Training.and.Teaching.Others	2	4	0	33.3	66.7	0.0
NA.	1	0	2	33.3	0.0	66.7
Judging.the.Qualities.of.Things..Se	1	3	0	25.0	75.0	0.0
Developing.Objectives.and.Strategie	0	1	0	0.0	100.0	0.0
Estimating.the.Quantifiable.Charact	0	1	0	0.0	100.0	0.0

Table 13: Task/Knowledge Codification interaction (<2010) – Data Management papers

	Absolute frequencies			Row shares		
TaskShort	Yes	No	NA	Yes	No	NA
Analyzing.Data.or.Information	9	18	0	33.3	66.7	0
Communicating.with.Supervisors..Pee	3	0	0	100.0	0.0	0
Documenting.Recording.Information	4	5	0	44.4	55.6	0
Estimating.the.Quantifiable.Charact	0	2	0	0.0	100.0	0
Getting.Information	1	15	0	6.2	93.8	0
Inspecting.Equipment..Structures..o	0	1	0	0.0	100.0	0
Judging.the.Qualities.of.Things..Se	0	1	0	0.0	100.0	0
Making.Decisions.and.Solving.Proble	0	1	0	0.0	100.0	0
Monitor.Processes..Materials..or.Su	3	1	0	75.0	25.0	0
Monitoring.and.Controlling.Resource	3	0	0	100.0	0.0	0
Performing.General.Physical.Activit	0	1	0	0.0	100.0	0
Processing.Information	5	23	0	17.9	82.1	0
Provide.Consultation.and.Advice.to.	0	2	0	0.0	100.0	0
NA.	0	0	0	NaN	NaN	NaN

The range of tasks that are affected by the exposure to and adoption of software related data management technologies is lower and including information processing activities though also supervision , monitoring and control, all of which requiring codified knowledge.

The codified segment of the data value chain seem to be **upstream**, for tasks of monitoring, recording, and in general data generation-related tasks. The **downstream** segments of the data value chain, that include the processing of information and analysis of data till the production of information and knowledge, seem to be based on less codified knowledge.

These characteristics in terms of degree of knowledge codification do not change over time, though after 2010 we register a sensible increase in the number of papers focused on data management techniques.

Table 14: Task/Knowledge Codification interaction (>2010) – Data Management papers

	Absolute frequencies			Row shares		
TaskShort	Yes	No	NA	Yes	No	NA
Analyzing.Data.or.Information	16	43	1	26.7	71.7	1.7
Controlling.Machines.and.Processes	0	1	0	0.0	100.0	0.0
Documenting.Recording.Information	7	7	1	46.7	46.7	6.7
Evaluating.Information.to.Determine	2	1	0	66.7	33.3	0.0
Getting.Information	10	45	1	17.9	80.4	1.8
Identifying.Objects..Actions..and.E	4	4	0	50.0	50.0	0.0
Inspecting.Equipment..Structures..o	2	1	0	66.7	33.3	0.0
Interacting.With.Computers	1	4	0	20.0	80.0	0.0
Interpreting.the.Meaning.of.Informa	2	1	0	66.7	33.3	0.0
Judging.the.Qualities.of.Things..Se	0	8	0	0.0	100.0	0.0
Making.Decisions.and.Solving.Proble	2	10	0	16.7	83.3	0.0
Monitor.Processes..Materials..or.Su	6	10	0	37.5	62.5	0.0
Monitoring.and.Controlling.Resource	4	4	0	50.0	50.0	0.0
Organizing..Planning..and.Prioritiz	1	5	0	16.7	83.3	0.0
Performing.Administrative.Activitie	1	0	1	50.0	0.0	50.0
Performing.General.Physical.Activit	2	2	0	50.0	50.0	0.0
Processing.Information	19	78	1	19.4	79.6	1.0
Scheduling.Work.and.Activities	10	22	0	31.2	68.8	0.0
Selling.or.Influencing.Others	1	0	0	100.0	0.0	0.0
Staffing.Organizational.Units	1	0	0	100.0	0.0	0.0
NA.	0	0	1	0.0	0.0	100.0

One of the most interesting pieces of information that our analysis allows is the cross-tabulations between shares of papers that reports automation of routinised tasks and those that report substitution of or complementarity to tasks performed by humans. This is reported respectively over the whole sample of extracted records, and on the sub-samples of papers published before and after 2010.

Table 15: Routinisation/Substitution interaction (Full) – robot paper

	Absolute frequencies				Row shares			
Routinisation	Compl	Compl;Sub	Sub	NA	Compl	Compl;Sub	Sub	NA
Yes	53	72	142	0	19.9	27.0	53.2	0
Yes..No	6	0	0	0	100.0	0.0	0.0	0
No	119	20	29	0	70.8	11.9	17.3	0
NA.	0	1	0	3	0.0	25.0	0.0	75

Technical papers on robotization published before 2010 show a neater association between tasks that require tacit and creative (less codified) knowledge and complementarity between automated and human tasks. This association is less neat for codified tasks, which show similar shares of complementarity and substitutability. These shares are interestingly reversed in the most recent

papers, which illustrate prototypes and conception of robot related automation that are both complement and substitute despite based on codified knowledge and show a much less neat association between use of tacit and creative knowledge and task complementarity.

Table 16: Knowledge codification/Substitution interaction (<2010) – robot papers

	Absolute frequencies				Row shares			
Knowledge.codification	Compl	Compl;Sub	Sub	NA	Compl	Compl;Sub	Sub	NA
Yes	31	1	24	0	55.4	1.8	42.9	0
No	69	4	18	0	75.8	4.4	19.8	0
NA.	0	0	0	1	0.0	0.0	0.0	100

Table 17: Knowledge codification/Substitution interaction (>2010) – robot papers

	Absolute frequencies				Row shares			
Knowledge.codification	Compl	Compl;Sub	Sub	NA	Compl	Compl;Sub	Sub	NA
Yes	28	70	85	0	15.3	38.3	46.4	0.0
No	50	17	44	0	45.0	15.3	39.6	0.0
NA.	0	1	0	2	0.0	33.3	0.0	66.7

The same piece of information with respect to the data management technologies reveals a very different pattern. Both over time and across different degrees of knowledge codification (though more dominant in the non-codified knowledge category), data management technologies are largely **complementary** rather than substitutive of tasks performed by humans.

Table 18: Knowledge codification/Substitution interaction (<2010) – Data management papers

	Absolute frequencies				Row shares			
Knowledge.codification	Compl	Compl;Sub	Sub	NA	Compl	Compl;Sub	Sub	NA
Yes	24	0	4	0	85.7	0.0	14.3	0
No	66	4	0	0	94.3	5.7	0.0	0
NA.	0	0	0	0	NaN	NaN	NaN	NaN

Table 19: Knowledge codification/Substitution interaction (>2010) – Data management papers

	Absolute frequencies				Row shares			
Knowledge.codification	Compl	Compl;Sub	Sub	NA	Compl	Compl;Sub	Sub	NA
Yes	61	13	17	0	67.0	14.3	18.7	0.0
No	217	0	29	0	88.2	0.0	11.8	0.0
NA.	5	0	0	1	83.3	0.0	0.0	16.7

To summarise, **robots** seem to both complement and substitute labour, though in more recent publications, the technical literature displays an increasing focus on robots that substitute humans. This suggest that robots capabilities are accrued with time along their development stage, as well as the fact that it is the combination of robots with other technologies to enable them to expand the set of activities in which substitution can occur. For instance, the tasks of identifying objects and handling and moving objects are particularly interesting, as there substitution and complementarity co-exists. This points to the possibility that different cohorts of robots with different degree of capabilities co-exist, with some just improving efficiency and facilitating workers' operation, while other fully automating processes, for example opening the way for flexible factory-floors with reconfigurable assembly systems (Kousi et al., 2018).

In contrast, **data management** technologies are substantially more complementary, with cases of substitution detected only in sectors M and N. However, these cases account for a negligible share of the total. Notwithstanding that, the majority of papers overviewed for this technology describe DM use in routinised tasks. A possible interpretation of this fact can be explored when combining this information with the results on knowledge codification, as DM technologies used in tasks of getting information, processing information, and analysing data seem to be overwhelmingly grounded on the presence of tacit knowledge, especially in sectors C, F, and M. In brief, tasks are standardised, which would make them exposed to automation. However, labour cannot be easily displaced, as it holds key know-how required to complete the very task at hand. Human knowledge in these tasks acts as an enabler, or better an essential factor – without it, the task cannot be executed at any level of efficiency granted by automation technologies. Even tasks usually considered as conquest ground for automated systems, such as scheduling work and activities, seem to be grounded by the importance of non-codified knowledge, which acts as a 'reverse salient' – a front of activity that generates bottlenecks. A task for which we record more DM-papers suggesting that codification of instructions is possible is Document and recording information. The relatively higher presence of codifiable information suggests that this task could be considered as a possible 'target' for upcoming automation efforts.

In summary, R and DM are following very different path of automation, penetrating along different paths tasks and sectors. Robots display both complementary and substitutive relationships with labour, likely because different cohorts of robots with different degrees of multi-purpose capabilities co-exist. The nature of knowledge does not seem to be a major reverse salient in this context. The opposite holds for DM technologies, which, currently, mostly complement labour even in routinised tasks, due to the important of the know-how 'stored' tacitly in humans to complete tasks in which DM can be employed.

4.1.4 Integrating People, Symbols and Things

Finally we codify records of technical publications according to an interesting - and fairly under-estimated – contribution by Robert Reich in his book "The Work of Nations. Preparing Ourselves for

21st Century Capitalism" (Reich, 1991). Quite evocatively, Reich considers the North American labour markets at the dawn of the 1990s, and divides jobs into three categories. The "symbolic analytic" services, the "routine production" services, and the "in person" services.

The classification by Reich (1991) is revisited here to classify tasks within occupations on the basis of how the use of **technologies involves interactions mainly with people, things or symbols**. According to Reich (1991), the Symbolic Analysists are workers that mainly interact with technologies through symbols, and hence include, among other, what has been alternatively classified as "Knowledge Intensive Business Services": engineers, lawyers, scientists, academics, consultants and other intellectual activities. The routine and the in-person services respectively interact with technologies through Things –performing what has later been considered routinised tasks (as assembly workers, data processors, machine supervisors) – and through People, as in personal services, such as care workers, essential services, janitors and so on.

We predict that this classification will offer a good explanatory power in making sense of how the automation technologies considered here will reconfigure the tasks in terms of human-machine composition.

Currently technical papers show that the non-automated, non-routinised and complex tasks are less likely to interact mainly with things or people, unlike routinised ones, which tend instead to involve interactions with things.

Table 20 and Table 21 below show the absolute number and row shares of technical papers on robots that involve tasks and how these mainly interact with people (P), symbols (S) and things (T) or a combination of pairs of them. Interactions with the three of them seem to be infrequent. Most frequently, it has been possible to codify technical records according to a single, neat interaction target. Some tasks involving the use of robots only interact with People (Coaching others; establishing and maintaining personal relationships; Estimating the Quantifiable Characteristics of Products, Events, or Information; monitoring or controlling resources; performing for or working directly with the public) though these are shown in a negligible number of papers. The most frequent tasks affected by robotization are interestingly less neatly interacting with a main target: for instance, Identifying Objects, Actions, and Events and Handling and Moving Objects show similar shares of interaction with things, as expected, but also with people. To sum up, technical papers published between 2000 and 2020 on robots describe prevalently robotization processes that affect a very few number of tasks. These tend to incorporate technologies which mainly target interactions with things and people. Robots interacting with symbols belong to the future and, likely, to the interaction of symbol manipulation with more complex technology systems integrating both physical and data processing and analysis capabilities.

An opposite picture emerges for the technical papers focused on data management technologies. Several papers are focused on a larger spectrum of tasks (as expected, in large majority dealing with gathering, processing and analysing information) and all of them mainly interact with symbols, rather than people or things, or a combination of symbols and things.

The evidence emerging in this section is consistent with the features of knowledge codification, routinization and complementarity/substitutability considered above.

Table 20: Task/Works with interaction (full) - robot paper

	Absolute frequencies								Row shares							
TaskShort	P	S	T	P/S	P/T	S/T	P/S/T	NAs	P	S	T	P/S	P/T	S/T	P/S/T	NAs
Analyzing.Data.or.Information	0	6	1	5	0	0	0	0	0.0	50.0	8.3	41.7	0.0	0.0	0.0	0.0
Assisting.and.Caring.for.Others	19	0	1	1	5	0	0	0	73.1	0.0	3.8	3.8	19.2	0.0	0.0	0.0
Coaching.and.Developing.Others	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Communicating.with.Supervisors..Pee	0	0	0	4	0	0	0	0	0.0	0.0	0.0	100.0	0.0	0.0	0.0	0.0
Controlling.Machines.and.Processes	0	0	14	2	2	1	2	0	0.0	0.0	66.7	9.5	9.5	4.8	9.5	0.0
Developing.Objectives.and.Strategie	1	1	0	1	1	0	0	0	25.0	25.0	0.0	25.0	25.0	0.0	0.0	0.0
Documenting.Recording.Information	0	1	0	0	1	3	0	0	0.0	20.0	0.0	0.0	20.0	60.0	0.0	0.0
Establishing.and.Maintaining.Interp	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Estimating.the.Quantifiable.Charact	0	0	1	0	0	0	0	0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
Evaluating.Information.to.Determine	0	0	1	0	0	0	0	0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
Getting.Information	9	6	15	7	2	0	0	0	23.1	15.4	38.5	17.9	5.1	0.0	0.0	0.0
Handling.and.Moving.Objects	45	4	39	1	12	0	0	0	44.6	4.0	38.6	1.0	11.9	0.0	0.0	0.0
Identifying.Objects..Actions..and.E	39	1	27	2	11	1	0	0	48.1	1.2	33.3	2.5	13.6	1.2	0.0	0.0
Inspecting.Equipment..Structures..o	1	6	14	0	4	0	0	0	4.0	24.0	56.0	0.0	16.0	0.0	0.0	0.0
Interacting.With.Computers	0	0	1	0	0	1	0	0	0.0	0.0	50.0	0.0	0.0	50.0	0.0	0.0
Interpreting.the.Meaning.of.Informa	0	2	1	2	2	0	0	0	0.0	28.6	14.3	28.6	28.6	0.0	0.0	0.0
Judging.the.Qualities.of.Things..Se	0	0	1	2	0	1	0	0	0.0	0.0	25.0	50.0	0.0	25.0	0.0	0.0
Making.Decisions.and.Solving.Proble	1	0	1	0	2	1	0	0	20.0	0.0	20.0	0.0	40.0	20.0	0.0	0.0
Monitor.Processes..Materials..or.Su	4	4	4	0	2	2	0	0	25.0	25.0	25.0	0.0	12.5	12.5	0.0	0.0
Monitoring.and.Controlling.Resource	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Operating.Vehicles..Mechanized.Devi	1	0	16	0	2	0	0	0	5.3	0.0	84.2	0.0	10.5	0.0	0.0	0.0
Organizing..Planning..and.Prioritiz	0	0	1	0	0	3	0	0	0.0	0.0	25.0	0.0	0.0	75.0	0.0	0.0
Performing.Administrative.Activitie	0	0	2	0	0	0	0	0	0.0	0.0	100.0	0.0	0.0	0.0	0.0	0.0
Performing.for.or.Working.Directly.	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Performing.General.Physical.Activit	8	0	14	2	8	4	2	2	20.0	0.0	35.0	5.0	20.0	10.0	5.0	5.0
Processing.Information	1	0	7	0	0	0	0	0	12.5	0.0	87.5	0.0	0.0	0.0	0.0	0.0
Provide.Consultation.and.Advice.to.	0	2	0	0	0	0	0	0	0.0	100.0	0.0	0.0	0.0	0.0	0.0	0.0
Scheduling.Work.and.Activities	0	0	2	0	0	1	0	0	0.0	0.0	66.7	0.0	0.0	33.3	0.0	0.0
Training.and.Teaching.Others	2	0	0	0	4	0	0	0	33.3	0.0	0.0	0.0	66.7	0.0	0.0	0.0
NA.	1	0	0	0	0	0	0	2	33.3	0.0	0.0	0.0	0.0	0.0	0.0	66.7

Table 21: Task/Works with interaction (full) – Data Management paper

	Absolute frequencies								Row shares							
TaskShort	P	S	T	P/S	P/T	S/T	P/S/T	NAs	P	S	T	P/S	P/T	S/T	P/S/T	NAs
Analyzing.Data.or.Information	0	78	1	2	1	5	0	0	0	89.7	1.1	2.3	1.1	5.7	0	0
Communicating.with.Supervisors..Pee	0	3	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Controlling.Machines.and.Processes	0	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Documenting.Recording.Information	0	23	1	0	0	0	0	0	0	95.8	4.2	0.0	0.0	0.0	0	0
Estimating.the.Quantifiable.Charact	0	2	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Evaluating.Information.to.Determine	0	3	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Getting.Information	0	70	0	2	0	0	0	0	0	97.2	0.0	2.8	0.0	0.0	0	0
Identifying.Objects..Actions..and.E	0	7	0	0	0	1	0	0	0	87.5	0.0	0.0	0.0	12.5	0	0
Inspecting.Equipment..Structures..o	0	3	0	0	0	1	0	0	0	75.0	0.0	0.0	0.0	25.0	0	0
Interacting.With.Computers	0	5	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Interpreting.the.Meaning.of.Informa	0	3	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Judging.the.Qualities.of.Things..Se	0	9	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Making.Decisions.and.Solving.Proble	0	10	0	1	0	2	0	0	0	76.9	0.0	7.7	0.0	15.4	0	0
Monitor.Processes..Materials..or.Su	0	19	0	0	0	1	0	0	0	95.0	0.0	0.0	0.0	5.0	0	0
Monitoring.and.Controlling.Resource	0	7	0	0	0	4	0	0	0	63.6	0.0	0.0	0.0	36.4	0	0
Organizing..Planning..and.Prioritiz	0	3	0	0	0	3	0	0	0	50.0	0.0	0.0	0.0	50.0	0	0
Performing.Administrative.Activitie	0	2	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Performing.General.Physical.Activit	0	3	0	0	0	2	0	0	0	60.0	0.0	0.0	0.0	40.0	0	0
Processing.Information	0	115	2	2	1	6	0	0	0	91.3	1.6	1.6	0.8	4.8	0	0
Provide.Consultation.and.Advice.to.	0	2	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Scheduling.Work.and.Activities	0	25	0	0	0	7	0	0	0	78.1	0.0	0.0	0.0	21.9	0	0
Selling.or.Influencing.Others	0	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
Staffing.Organizational.Units	0	1	0	0	0	0	0	0	0	100.0	0.0	0.0	0.0	0.0	0	0
NA.	0	0	0	0	0	0	0	1	0	0.0	0.0	0.0	0.0	0.0	0	100

5 A summary of main findings

One of the main contributions of this paper is a methodological benchmark to approach from a multidisciplinary angle a literature review of technical papers that focus on the conception, experimentation and development of emerging technologies, such as robots-related automation and software-based data management.

We offer a very detailed reconstruction of the methodological journey that has led to a careful classification of emerging automation technologies, and a grounded multi-steps literature review of technical records that are most relevant for our purpose of fine graining technologies and tasks affected. This is included above in Section 3 and a very detailed Appendix. This allows a general interpretative grid to codify the records extracted.

We then provide a quantitative descriptive analysis of the qualitative codification of the papers extracted, focusing on the relevant employment-related variables illustrated at length in Section 4.

Overall, the analysis of the technical literature provides a rich zoomed-in picture of the economic evidence, which usually clusters automation in one single technology with a substitution effect on routinised tasks requiring codified knowledge. We start here by considering the technical literature focused on two of the most different ones, robotization-related automation and software-based data management.

First, automation technologies are fundamentally different and might have different effects. The future of work depends on their evolution, their idiosyncrasies, their stage of development and adoption, the specific sectors that are mostly exposed to each of them, the specific tasks they complement or replace. As highlighted in the introduction to this paper, the objective of unpacking different technologies is important and is so also in view of predicting the next trends of the future of work when the range of automation technologies changes shape, composition and effects over time. For instance, automation related to robotization is likely to become more and more substitutive of tasks performed by humans, notwithstanding some of these technologies are at the very experimental stage.

Second, Software-based data management technologies are more mature, they are more pervasive in services than in manufacturing sectors, though we expect new applications to emerge, particularly as a result of the novel regulatory framework emerging at the EU levels, such as the Digital Market Act, the Digital Service Act and the Data Act.

Third, over time, the literature on robots published post 2010 shows that they tend to become more substitutive than they were at the beginning of the years 2000s. Data management technologies are consistently more complementary to tasks performed by humans over time; they seem to be more pervasively used in services than in manufacturing, which might reflect the higher maturity of their development and adoption.

Fourth, the use of codified or tacit knowledge is fairly associated with routinization, whereby the most routinised tasks performed by these technologies seem to make use of codified more than tacit knowledge.

Finally, as one might expect, DM technologies interact largely more with symbols, rather than with things or people. In fact, the technologies included in this category emerge as covering tasks of processing and analysing information, unlike robotization-based automation is described as mainly interacting with things or people, in this latter case when employed to automatise processes that are supervised or managed by humans.

6 Next steps

The PILLARS proposal has been submitted before the full blown pandemics, in March 2020; it has been awarded in July 2020 with a contract finalised in October 2020. At the time of submission, we had scheduled the deliverable D1.1. at the end of Month 12, with no prior knowledge of what was to come. We have been asked to start the project in January 2021, which has ticked the clock for the submission period of this deliverable, due on the 1st of January 2022.

The work, development and submission of this deliverable have been delayed by a series of substantial bottlenecks, mainly due to the effects of the Covid-19 pandemics. In particular, this has delayed the recruitment of a Research Assistant, who has started working only in July 2021; it has affected all the team members in terms of sick leave at different stages of the ongoing work (the PI and Co-I for most of the month of January 2022). One of the team members still suffers from Long Covid, which has been proved in the medical literature to significantly affect normal functioning.

In terms of next steps, we plan to finalise the analysis of the other technologies and present the full results at the mid-term conference in July 2022 to get feedback and finalise an academic publication out of the present report.

7 References

Acemoglu, D., & Restrepo, P. (2019). Automation and new tasks: How technology displaces and reinstates labor. Journal of Economic Perspectives, 33(2), 3-30.

Acemoglu, D., Lelarge, C., & Restrepo, P. (2020). Competing with robots: Firm-level evidence from france. In *AEA Papers and Proceedings* (Vol. 110, pp. 383-88).

Albert, T. (2016). Measuring Technology Maturity: Theoretical Aspects. In Measuring Technology Maturity (pp. 9-113). Springer Gabler, Wiesbaden.

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

Autor, D. (2019). Work of the Past, Work of the Future (No. w25588). National Bureau of Economic Research.

Bakhshi, H., Downing, J. M., Osborne, M. A., & Schneider, P. (2017). *The future of skills: Employment in 2030*. Pearson.

Balland, P. A., & Boschma, R. (2021). Mapping the potentials of regions in Europe to contribute to new knowledge production in Industry 4.0 technologies. *Regional Studies*, 1-15.

Benmelech, E., & Zator, M. (2022). *Robots and Firm Investment* (No. w29676). National Bureau of Economic Research.

Bresnahan, T., & Yin, P. L. (2017). Adoption of new information and communications technologies in the workplace today. *Innovation policy and the economy*, *17*(1), 95-124.

Ciarli, T., Kenney, M., Massini, S., & Piscitello, L. (2021). Digital technologies, innovation, and skills: Emerging trajectories and challenges. Research Policy, 50(7), 104289.

Cohendet, P., & Edward Steinmueller, W. (2000). The codification of knowledge: a conceptual and empirical exploration. Industrial and corporate change, 9(2), 195-209.

Combemale, C., Whitefoot, K. S., Ales, L., & Fuchs, E. R. (2021). Not all technological change is equal: how the separability of tasks mediates the effect of technology change on skill demand. Industrial and Corporate Change, 30(6), 1361-1387.

Cowan, R., David, P. A., & Foray, D. (2000). The explicit economics of knowledge codification and tacitness. Industrial and corporate change, 9(2), 211-253.

Foray, D., & Steinmueller, W. E. (2003). The economics of knowledge reproduction by inscription. Industrial and corporate change, 12(2), 299-319.

Goldin, C., & Katz, L. F. (2020). Extending the race between education and technology. In AEA Papers and Proceedings (Vol. 110, pp. 347-51).

Hopp, W. J., & Spearman, M. L. (2011). Factory physics. Waveland Press.

IBM Classification of 4.0 technologies, accessible at https://www.ibm.com/topics/industry-4-0 (last access: March 14, 2022)

Inaba, T. and M. Squicciarini (2017), "ICT: A new taxonomy based on the international patent classification", OECD Science, Technology and Industry Working Papers, No. 2017/01, OECD Publishing, Paris, https://doi.org/10.1787/ab16c396-en.

Keisner, C.A. Raffo, J., & Wunsch-Vincent, S. (2015), *Breakthrough technologies – Robotics, innovation and intellectual property*, Economic Research Working Paper No. 30. WIPO

Kousi, N., Michalos, G., Aivaliotis, S., & Makris, S. (2018). *An outlook on future assembly systems*

introducing robotic mobile dual arm workers. Procedia CIRP, 72, 33-38.

Kousi, N., Michalos, G., Aivaliotis, S., & Makris, S. (2018). An outlook on future assembly systems introducing robotic mobile dual arm workers. *Procedia CIRP*, *72*, 33-38.

León-Ledesma, M. A., McAdam, P., & Willman, A. (2010). *Identifying the elasticity of substitution with biased technical change*. American Economic Review, 100(4), 1330-57.

Martinelli, A., Mina, A., & Moggi, M. (2021). The enabling technologies of industry 4.0: Examining the seeds of the fourth industrial revolution. *Industrial and Corporate Change*, 30(1), 161-188.

Ménière, Y., Rudyk, I., and Valdes, J. (2017). *Patents and the Fourth Industrial Revolution: The inventions behind digital transformation*. European Patent Office.

Mokyr, J., Vickers, C., and Ziebarth, N. L. (2015). The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different? Journal of Economic Perspectives, 29(3):31–50

OECD (2017), *OECD Science, Technology and Industry Scoreboard 2017: The digital transformation*, OECD Publishing, Paris http://dx.doi.org/10.1787/9789264268821-en

OECD (2017), *The Next Production Revolution: Implications for Governments and Business*, OECD Publishing, Paris. http://dx.doi.org/10.1787/9789264271036-enOECD.

OECD. (2018). Transformative Technologies and Jobs of the Future.

Penny, Simon (2013) 'Art and Robotics: Sixty Years of Situated Machines. [REVIEW],' *AI and Society* 28 (2):147-156.

Ray, D., & Mookherjee, D. (2021). Growth, automation, and the long-run share of labor. Review of Economic Dynamics.

Rogers, E. M. (2010). Diffusion of innovations. Simon and Schuster.

Rotolo, D., Hicks, D., & Martin, B. R. (2015). What is an emerging technology?. Research policy, 44(10), 1827-1843.

Saviotti, P. P., & Metcalfe, J. S. (1984). A theoretical approach to the construction of technological output indicators. *Research policy*, *13*(3), 141-151.

Savona, M. and W.E. Steinmueller (2013). Service output, innovation and productivity: A time-based conceptual framework. Structural Change and Economic Dynamics. Elsevier, ISSN 0954-349X. - Vol. 27.2013, C, p. 118-132

Servoz, M. (2019), *AI, the Future of work? Work of the Future! On How Artificial Intelligence, Robotics and Automation are Transforming Jobs and the Economy in Europe*. Brussels: European Political Strategy Centre (European Commission), Brussels

Sheridan, T. B. (2016). Human–robot interaction: status and challenges. *Human factors*, *58*(4), 525-532.

Takeuchi, H., & Nonaka, I. (1986). The new new product development game. Harvard business review, 64(1), 137-146.

Von Tunzelmann, N. (1995). Time-saving technical change: The cotton industry in the English industrial revolution. Exploration in Economic History, 32:1–27

Wirkierman, AL. (2022). Structural dynamics in the era of smart technologies, in Kurz, H. D., Schütz, M., Strohmaier, R., & Zilian, S., *The Routledge Handbook of Smart Technologies* (pp. 1-14). Routledge.

Zolas, N., Kroff, Z., Brynjolfsson, E., McElheran, K., Beede, D. N., Buffington, C., ... & Dinlersoz, E. (2021). *Advanced Technologies Adoption and Use by US Firms: Evidence from the Annual Business Survey* (No. w28290). National Bureau of Economic Research

8 Appendix

This section details the final literature search protocol, which is the outcome of several iterations, both to define technology families, as well as to select queries and their expansions. This allows us to articulate in detail the rationale underpinning our methodological choices and the solid grounding of the selected technical literature that supports our findings. We propose this as a methodological template for any in-depth investigation of multidisciplinary literature that seeks to shed light on such a complex phenomenon.

The remainder of the Appendix is organised as follows. The first part (Section 8.1) describes the rationales behind the classification into families of automation technologies that will be used across several PILLARS Work Packages, particularly WP3 which deals with the phase of emerging technologies. In this paper, we focus on two technologies, one mainly used in manufacturing (robots) and one mainly used in services (software-based data management). The second part (Section 8.2) details the complex and rigorous search strategy for relevant papers, which attempts to maximise both precision (selection of relevant papers) and recall (covering as many technology as are available in each family). We describe the basic query to search for documents in Scopus, the manual screening of the sampled papers, the identification of the relevant engineering controlled terms to expand the initial selection of core papers to a more comprehensive set.

8.1 Identification of Automation Technologies

To identify the families of digital technologies that contribute to the automation of production processes in all sectors of the economy that involve workers (e.g. excluding households' digital appliances). We followed four steps:

a) First, team members listed and classified technologies based on the literature in economics and management and on their expert judgement. This was followed by two validation exercises.
b) We used reports and academic publications from academia, industry, government and the public on Industry 4.0, future of work, and automation, to build a concordance table between the classification developed in step (a) and classifications used by experts across sectors.
c) We constructed a control classification, developed by experts in a different research team in the PILLARS project, to build another concordance table with the classification developed in step (a)
d) We studied the concordances between our first classification, the classifications used by experts across sectors (b) and the second PILLARS expert classification in point (c) and we produced a final classification. We discussed this final classification with the whole PILLARS consortium and made minor refinements.

For replicability, we describe each step in detail below

8.1.1 First classification

The initial classification into families of digital automation technologies was performed based on the literature and the expertise of the research team, especially by Ed Steinmueller. This classification was guided by industry observation and writing in the areas of industrial economics, innovation studies and technology history. It included eight classes, listed in Table 22-A, each with a list of prominent technologies. To some extent, the list of prominent technologies is arbitrary and reflects an early 21st century understanding of uses and purposes of these technologies. As we move through this century, there will certainly be additional subclassifications (e.g. 'robots' introduced into human bodies to diagnose or repair). The reported subclassifications reflect areas where substantial investments have occurred and that have a potential or demonstrable effect on human skills or employment.

Table 22-A: Classification of digital automation technology families (first classification)

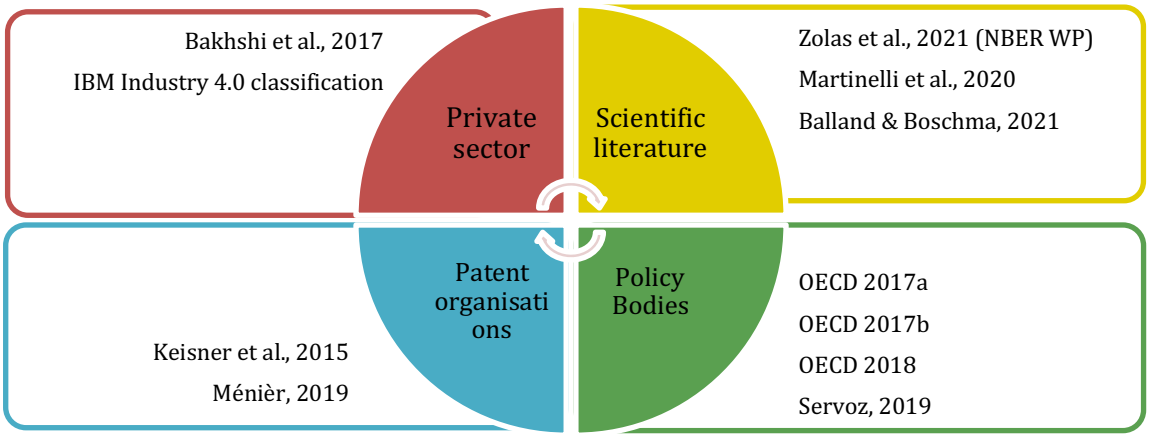
Codes	Automation technology families and prominent technologies
A	**Robots (Primarily stationary)**
1	Machine vision and real-time monitoring
2	Selective Compliance Assembly Robots (SCARA)
3	Articulated
4	Cartesian
5	Dual Arm
6	Co-bots (non-mobile)
7	Swarm robotics
8	Cylindrical
9	Service robotics
B	**Robots (Primarily mobile)**
1	Co-bots (mobile)
2	Semi-autonomous (e.g. bricklaying)
3	Automated platforms/vehicles
4	Tunnel boring and mining robots
5	Submersible robots
6	Drones
7	Space vehicles and rovers
8	Service robotics
C	**Data Acquisition Technologies**
1	IoT (including Radio-frequency identification (RFID) systems)
2	Scanners
3	Sensors
4	Remote Sensing

5	GPS
6	CCTV
7	Scientific and engineering instruments
8	Healthcare instruments (including Personal health instrumentation)
D	**Cloud, PC and Smartphone Computing and Services**
1	Big data analytics
2	Gaming
3	Streaming services
4	Automated storage and retrieval systems
5	Database systems
6	Relational databases
7	Computer architectures
8	Cryptography and security
9	5G
E	**AI (not directly as a cloud service)**
1	Simulation
2	Machine learning (predictive systems)/ Deep Learning
3	NLP
4	Machine vision (Image recognition)
5	Expert systems
6	Speech recognition and production
7	Text recognition and production
8	Machine Translation
F	**Additive manufacturing (using any material – e.g. powder metallurgy as well as bioplastic filament)**
1	Prototyping (including CAD)
2	Tools production
3	Production at scale
G	**User interface**
1	Conventional input devices (e.g. keyboard, mice, pens, webcams)
2	Display devices (conventional)
3	Augmented reality
4	Haptics and Tele-haptics
5	Virtual Reality (including 3D Visualisation)
6	Touchscreens/kiosks for customer interface
7	Sound technologies (e.g. noise cancellation)
8	Neuro-scanning
H	**Other**
1	Machine Tools
2	Wireless Identification Tags/Beacons

8.1.2 Concordance with classifications in the literature

A first validation of the classification presented in Section 8.1.1 (and Table 22-A) was done by reviewing existing classifications of automation technologies published by different organisations across different sectors of the economy: private sector, patent authorities, policy organisations, and academia (Figure 2-A). The literature refers to automation technologies under several frameworks, the most common one being Industry 4.0. We focused on documents especially focused in understanding the adoption of automation technologies and their impact on labour, specifically the automation of tasks. The selected documents include three OECD reports (OECD, 2017a; OECD, 2017b; OECD, 2018), one report each from the EC (Servoz, 2019), WIPO (Keisner et al., 2015), EPO (Ménière et al., 2017)), a report from Nesta with Pearson (Bakhshi, 2017), the IBM classification (https://www.ibm.com/topics/industry-4-0), and three academic papers (Balland & Boschma, 2021; Martinelli et al., 2021; Zolas et al., 2021)).

Figure 2-A: Reference documents from four sectors of the economy



From each of the documents selected we extracted the technology classification proposed. Whenever available, we also collected a list of prominent technologies within each family. Based on this information, we created a concordance table to compare our first classification with the ones received from the literature. The rationale for this exercise was to identify potential gaps, missing technologies or entire families of technologies, or mis-classifications of prominent technologies into the technological families proposed in our first classification. The concordance table is included below in Table 23-A.

Table 23-A: Concordance table of automation technologies classification (first classification and literature)

Literature			First classification	
Report	**Families**	**Sub-categories**	**Family**	**Techno logy**
WIPO 2015	Robotics - Remoted controlled	Telepresence robots	B	
		Remoted controlled humanoid robots		
		Robotic assisted surgical devices		
		Exoskeletons		
		Drones		B.6
	Robotics - Semi autonomous		B	B.2
	Robotics - Fully autonomous		B	B.3
	Artificial Intelligence		E	E
Pearson-Nesta, 2017 | The Future of Skills	Robots		A/B	
	Artificial Intelligence		E	E
	Big data		D	D.1
	Internet of Things		C	C.1
OECD STI 2017	Mobility		B	
	Cloud computing		D	
	Internet of Things		C	C.1
	Artificial Intelligence		E	
	Big data analytics		D	D.1
AI Future of Work | European Commission, 2019	Artificial Intelligence	General AI: broad cognitive abilities	E	
		Narrow AI: human-level intelligence		
	Robotics		A/B	
	Internet of Things		C	C.1
	Blockchain technology			
	New and advanced materials		F	N/A
	Autonomous devices		B	
Balland & Boschma, 2021	Additive manufacturing		F	
	Artificial Intelligence		E	E
	Augmented reality		G	G.3

	Autonomous robots		B	
	Autonomous vehicles		B	
	Cloud computing		D	
	Cybersecurity			
	Machine tools		D	N/A
	Quantum computers		D	N/A
	System integration		D	D.6
OECD Next Production Revolution 2017	Simulations		E	E.1
	Artificial Intelligence		E	E
	System integration		D	D.6
	Big data		D	D.1
	Cloud computing		D	
	Internet of Things		C	C.1
OECD 2018 | Transformative technologies & jobs of the future	Artificial Intelligence		E	E
	Internet of Things		C	C.1
	Blockchain technology			
IBM IR 4.0 classification	Internet of Things		C	C.1
	Cloud computing		D	
	AI and machine learning		E	E / E.2
	Edge computing		D	N/A
	Cybersecurity			
	Digital twin		D	N/A
Zolas et al., 2021 (NBER)	Augmented reality		G	G.3
	Automated guided vehicles (AGV)		B	B.3
	Automated storage and retrieval systems		D	D.4
	Machine learning		E	E.2
	Machine vision		E	E.4
	Natural language processing		E	E.6
	Radio-frequency identification (RFID) system		C	C.1
	Robotics		A/B	
	Touchscreens/kiosks for customer interface		G	G.7

	Voice recognition software		E	E.6
Martinelli et al., 2020	Internet of Things		C	C.1
	Big data / Industrial analytics		D	D.1
	Cloud manufacturing		D	
	Robotics		A/B	
	Artificial Intelligence		E	E
	Additive manufacturing		F	F
European Patent Office	Core technologies	Hardware	D/G	
		Software	C/D/E	
		Connectivity	D/G	
	Enabling technologies	Analytics	D	D.1
		Security	D	
		AI	E	
		Position determination	D	
		Power supply	N/A?	
		3D systems	G	
		User interfaces	G	
	Application domains	Home, Personal, Enterprise, Manufacturing, Infrastructure, Vehicles	N/A	

There were few discrepancies between our first classification and list of technologies and the various classifications produced in the literature by different groups of experts.

Terms in yellow cells identify prominent technologies which were not explicitly mentioned in our first classification, but which were covered in existing families either as synonyms or as subcategories of a wider technological group. These were mainly technologies part of Cloud, PC and Smartphone Computing and Services (D), such as Quantum computers, Edge computing, Digital twins, System integration and position determination (part of big data analytics). New and advanced materials was classified under additive manufacturing (F). And Position determination was classified as one of the enabling technologies part of the broader big data analytics, also in family D.

Terms in red cells identify prominent technologies which were not explicitly mentioned in our first classification, and which are not easily classified under our first classification. These are Blockchain technologies and Cybersecurity technologies. As we discuss below, following our second comparison, we define a new technology family called 'software based data management', which also includes Blockchain technologies and Cybersecurity technologies.

8.1.3 Control classification

The control classification into families of digital automation technologies was performed based on the literature and the expertise of the research team preparing a Delphi survey on the emerging automation technologies. This classification was also guided by industry observation, the literature, and the OECD classification of ICT (Inaba and Squicciarini, 2017). It included only seven classes, listed and described in Table 24-A, each with a list of prominent technologies.

Table 24-A: Classification of digital automation technology families (control classification)

Code	**Automation technology families and prominent technologies**
A	**Robotics**
1	Programmable Robots
2	(Semi) Autonomous Robots
3	Service Robotics
4	Exoskeletons
5	Autonomous Vehicles
6	Drones
7	Swarm Robotics
8	Remote sensing
9	Smart Equipment
B	**Networking**
1	Streaming Services
2	5G
3	Internet of Things
4	Wireless Identification Tags/Beacons
C	**Data Management**
1	(Relational) Databases
2	Big Data
3	Cryptography
4	Cyber Security
5	Blockchain
6	Cloud Storage
7	Data Mining/Scraping
D	**Computing**
1	Edge Computing
2	Cloud Computing
3	Quantum Computing
4	High Performance Computing
5	Grid Computing
6	Simulation/Digital Twins
E	**Artificial Intelligence**
1	Machine/Deep Learning

Code	**Automation technology families and prominent technologies**
2	Predictive Maintenance
3	Natural Language Processing
4	Machine Vision
5	Text Recognition and Production
6	Speech Recognition and Production
7	Machine Translation
F	**Additive Manufacturing**
1	3D printing
2	Computer Aided Design
G	**Human Machine Interaction**
1	Virtual Reality (VR)
2	Augmented Reality (AR)
3	(Tele) Haptics
4	Neuro-control (Brain Controlled)
5	Holograms
6	Telepresence
7	Voice control
8	Visual Interfaces (Displays, monitors, touchscreens)
9	Biometrics
10	Apps
11	Gamification (Serious games)

8.1.4 Final classification

After building a concordance table between the first (step a), section 8.1.1) and the control classifications (step c), section 8.1.3), experts who developed the two classifications met to discuss results from the two concordance table and converge on a final, integrated, classification. As a final robustness check, this final table was discussed with the entire consortium. The final classification of automation technologies contains nine families and is reported in Table 25-A.

Table 25-A: PILLARS classification of digital automation technology families

A	**Robots (Articulated, Cylindrical, Cartesian, Dual Arm, SCARA)**
1	Machine vision and real-time monitoring
2	Co-bots
3	Swarm robotics
4	Service robotics
5	Semi-autonomous (e.g. bricklaying)
6	Automated platforms/vehicles
7	Tunnel boring and mining robots
8	Drones

9	Robotic vehicles (including Space vehicles and rovers, autonomous vehicles, Submersible robots)
10	Exoskeletons
11	Robotic Process Automation (RPA) (including software robots)
B	**Physical Data Acquisition Technologies**
1	Scanners
2	Sensors (including IoT and RFID)
3	Remote Sensing
4	GPS
5	CCTV
6	Scientific and engineering instruments
7	Healthcare instruments (including Personal health instrumentation)
8	Data scraping
C	**Software-based Data management**
1	Database systems
2	Relational databases (including API)
3	Cryptography, security, and blockchain (including data mining)
4	Big data analytics
D	**Computing**
1	Automated storage and retrieval systems (including cloud storage)
2	Computer architectures (including edge computing, cloud computing, HPC, grid computing)
E	**AI (not directly as a cloud service) & Intelligent Information System**
1	Simulation
2	Machine learning (predictive systems)/ Deep Learning
3	NLP
4	Machine vision (Image recognition)
5	Expert systems
6	Speech recognition and production
7	Text recognition and production (including Machine translation)
F	**Additive manufacturing (using any material e.g. powder metallurgy and bioplastic filament)**
1	Prototyping (including tools production, production at scale)
2	3D printing
3	CAD/CAM (prototype and/or production)
G	**Networking**
1	IoT (including Radio-frequency identification (RFID) systems)
2	Wireless communication (including 5G)
H	**User interface**
1	Conventional input devices (e.g. keyboard, mice, pens, webcams)
2	Display devices (conventional)
3	Augmented reality (including holograms)
4	Haptics and Tele-haptics (including all tele-operations of physical machinery by human operator requiring feedback, e.g. tele-dildonics)
5	Virtual Reality (including 3D Visualisation)
6	Touchscreens/kiosks for customer interface
7	Sound technologies (e.g. noise cancellation)
8	Neuroscanning
9	Gamification

I	**Other**
1	Machine Tools
2	Factory control system

As any other classification, the one presented in Table 25-A is based on a specific interpretation of these digital automation technologies. Our classification serves the purposes of the analysis of the literature in this paper, and the analysis of emerging technologies both using documents metadata (publications and patents) and using expert's view in the Delphi survey. To provide more clarity on our classification, below we provide our own definition and description of these technologies by family. It should also be noted that several of the prominent technologies listed in Table 25-A may be allocated in more than one family. As any classification, the boundaries between the technology families proposed here overlap, and technologies may fall under more than one family.

Technology family A – Robots (Articulated, Cylindrical, Cartesian, Dual Arm, SCARA)

Definition: Technologies that sense and (autonomously) act based on data

Although robots were initially imagined as autonomous and human-like in science fiction, their actual implementation reflects a series of incremental steps to integrate specific functionalities, e.g. initial efforts to develop software that could recognise and differentiate objects (e.g. wooden alphabet blocks led to the development of robot 'hands' and arms capable of arranging blocks or other objects). This basic capability was then further developed through the form of the robot during which the issue of robots and humans occupying the same space needed to be addressed (a robot that is sturdy enough to move large objects at speed is a potential hazard to human co-workers). Robot mobility is almost inevitably linked with human presence (e.g. exceptions did exist such as robots used to move into hazardous spaces such as nuclear accident sites) and also led to robots navigating different terrains and performing a few specialised functions (e.g. tunnel boring).

Technology family B – Physical Data Acquisition Technologies

Definition: technologies that harvest and record information

Computer systems originated as engines for data processing. The data that they employed needed to be fed carefully prepared (in order to be machine readable), a process that began as translation of paper records and proceeded to direct entry by data terminals. More recently, a collection of technologies has been developed for data acquisition that do not require this careful preparation. These new technologies stem from the use of digital measurement devices in laboratories and the use of locational beacons (e.g. GPS). The sensory capabilities of these devise were extended outside of the laboratory in the world, paralleling the rise of computer vision applications in robotics.

Technology family C – Software-based data management

Definition: Technologies for storing, protecting, managing/handling and acquiring data

The stores of data themselves became a site for innovation. What were once paper records generated by human activities became electronic records, audiovisual content became data streams and means of securing data became security and cryptographic innovations. The software used for these purposes is also included in this family.

Technology family D – Computing

Definition: Technologies used to compute/calculate

Computers were initially employed to automate the work of 'human computers' and quickly exceeded human capabilities in computational tasks. As a consequence, several activities that were previously performed by humans such as compiling tables of the values of mathematical functions for different parameters (e.g. values of cosine by angle) are so well established that they are excluded here. Innovation in the architectures for data storage and for performing computations are still occurring and constitute this family.

Technology family E – AI (not directly as a cloud service) & Intelligent Information System

Definition: Technologies using algorithms and advanced methods to make sense out of the data

The existence of large data stores (B) has opened a new range of computational capabilities commonly referred to as Artificial Intelligence (AI) which (so far) involves a collection of algorithms and some related hardware to implement machine learning. Machine learning has proven a powerful means of addressing the older issue of machine vision and the previously unreachable goal of Natural Language Processing (NLP) as well as more limited goals such as speech recognition or production.

Technology family F – Additive manufacturing (using any material)

Definition: Technologies that produce bottom-up based on digital models

A by-product of the development of mechanical and materials control for printing was the recognition that material objects could be 'printed' by precisely positioned deposits of material, what has come to be known as additive manufacturing.

Technology family G – Networking

Definition: Technologies for communicating between machines (data transmission) or connecting machines

The distribution of data entry and computation created a demand for rapid and reliable data communication between computer systems and the technologies for performing this function are networking technologies. As new technologies from family B were developed, they came to employ networking technologies as well. Networking technologies employ both 'wired' (physical connection) and 'wireless' (connection using the radiofrequency spectrum) methods. The extension of these networks through 'inter-networking' led to the Internet and with the Internet previous technologies such as telephony and broadcast media were increasingly accomplished with data communications (as these previous technologies were already automated, their applications are omitted. The areas of active innovation in networking are the integration of technologies from family B into the Internet (hence IoT) and the use of mobile phones as both receivers of broadcasts and as interactive devices (e.g. for voice and text communication).

It shall be noted that there are still areas where innovation is occurring that are not covered by IoT and 5G – e.g. edge networks and variants of the Internet protocol (e.g. IPv6). However, innovations in these areas are variants in the form of technology in already automated technologies and are therefore not directly implicated in employment (except to the extent that they extend or modify the skills needed to work in the data communications areas)

Technology family H – User interface

Definition: Technologies for human interaction with machines or data

The interaction between humans and computational capabilities is achieved through human computer interfaces, an area of continuing innovation. The technologies under this heading reflect those interactions that involve the presence of a human being exchanging information with a computer system through a variety of methods.

Technology family I – Other

Two important groups of technologies are machine tools and factory control systems. In principle these can be fit in the earlier categories. Machine tools can be seen as robots in their similarity with robot arms and hands being directed by software to perform precise operations, but this might be seen as an idiosyncratic interpretation of robots. Factory control systems are technologies that integrate components from most of the other families.

This classification is used to divide the systematic literature review into different blocks of coherent literature. To study the properties of these technologies in relation to working tasks, we conduct a separate literature review for each of the nine technology families, following the same procedure described below. The prominent technologies within each family are used to construct detailed search queries to identify relevant literature at a rather fine-grained level of technological detail. In

this deliverable we focus on two very different technologies: robots and software-based data management.

8.2 Systematic literature review of automation technologies: the protocol

As mentioned in the introduction, the main aim of this sub-project is to perform a literature review that broadly addresses the following research questions:

- Can digital automation technologies (as identified here) potentially substitute, complement, and/or reconfigure specific technical tasks within occupations within sectors, that are executed by humans?
- What are these specific tasks across sectors? Can we infer the extent to which specific industries are exposed to specific automation technologies at the level of granularity identified above?
- Can we infer anything on the future of work in relation to the contemporary wave of potentially emerging automation technologies, characterised by the pervasive presence of AI?

The novelty of the rationale proposed here with respect to existing reviews of the literature on the impact of selected technologies on employment is the focus on the technical literature rather than the economic literature, to go beyond the coarse view that is available in current measures of, e.g., robots and AI. The focus on technical features of automation technologies, as described by academics working on them, has the power of offering a better understanding of the nature of human-machines interdependence, depending on the nature of the technology, the tasks they can perform, and on the nature of human-machine interactions.

To this end, we followed a six-steps protocol, each including several sub-steps. It was first validated on robots, on which we have performed further checks (explained below). Figure 2-A shows the final sequence of steps, whereas Figure 3-A shows the sequence including the validation (on the literature on robots).

This protocol was applied to robots and software-based data management (see Table 25-A), and will be applied to all other technology families in future work.

8.2.1 Step 1: Identification of records: query

Because we are not interested in the technological development per se, but in what the technology can do, and in particular what tasks that it can perform, we first build a query that is composed in the following way. We define three different sets of keywords, each identifying

- the technologies in a given family (and synonymies),
- their functions or applications, and
- the tasks that they can perform.

To increase the precision of the documents retrieved by the query, the three sets of terms were connected using the proximity operator (W/n) which requires terms in the three sets to be no more than *n* terms away. For instance, if n=2, the search would retrieve only documents in which the technology term is followed (or preceded) by that term identifying the function at most two terms away, and the term identifying the task at most two terms away from the term identifying the function. In practice, this means that those three terms, on average, would appear in the same sentence of the abstract.

The identification of relevant keywords for the three sets of terms is based on our technology classification and on additional keywords identified via relevant papers from the Scopus and/or from core papers in the literature which study these technologies, especially in relation to labour. Although we present here our protocol as a linear process, the refinement of queries was an iterative process that necessarily required expert assessment, also based on the documents retrieved by several queries. We also experimented with different values of n, ranging between one and three.

8.2.2 Step 2: Identification of records: extraction of documents

We applied the query in Scopus and downloaded all documents which

- Were retrieved by our query (Step 1) in title, abstract, or keywords (TITLE-ABS-KEY)
- Were Published after 2000 or 2010 (depending on the technology) ((PUBYEAR)).
- Are among the following type of publication (DOCTYPE): article (ar) OR Conference Proceedings (cp), OR Conference Review (cr) OR review (re). We decided to include conference proceedings and reviews, despite the lack peer review quality check, because of the key goal of the analysis to capture emerging applications of technologies and the tasks they may be able to perform.
- Include one or more keyword(s) among from the indexed terms (INDEXTERM), to restrict to specific technologies, when available in the engineering thesaurus used by the Ei Compendex.[10]

We select the top X% of cited papers (normalized by year). The exact percentile is determined based on the size of the corpus of papers downloaded with the query; we aim at more or less 500 papers to be screened for each technology family. See Figure 6-A and Figure 7-A for details on the number of documents included and excluded in each step for robots and software based data management.

[10] https://www.elsevier.com/about/press-releases/archive/science-and-technology/5th-edition-of-the-ei-thesaurus-now-available

8.2.3 Step 3: Screening: documents selection

To create a sample of relevant documents to read and code for our literature review, we first screened the title and abstract of all the ~500 documents sampled in Step 2. Each paper was screened by two random reviewers independently.

The five reviewers agreed on the following rule to include or exclude documents from the sample of relevant documents. They excluded documents that were conceptually describing a technology, documents which were describing abstracts proofs of concept, and documents which did not explicitly refer to the performance of specific tasks. For instance, in the case of robot technologies, in relation to tasks, we included documents that explicitly referred to the automation of tasks, or robotic control in specific domains, both supervised and unsupervised, and the interactions between humans and robots. We instead excluded document referring to basic research on robotics, basic research on the capacity to move, or sense, but without a specific explanation about the task, basic research on the performance interaction between robots, again without a specific description of the tasks performed.

Figure 3-A: Flow-chart of the literature review process (including robot validation)

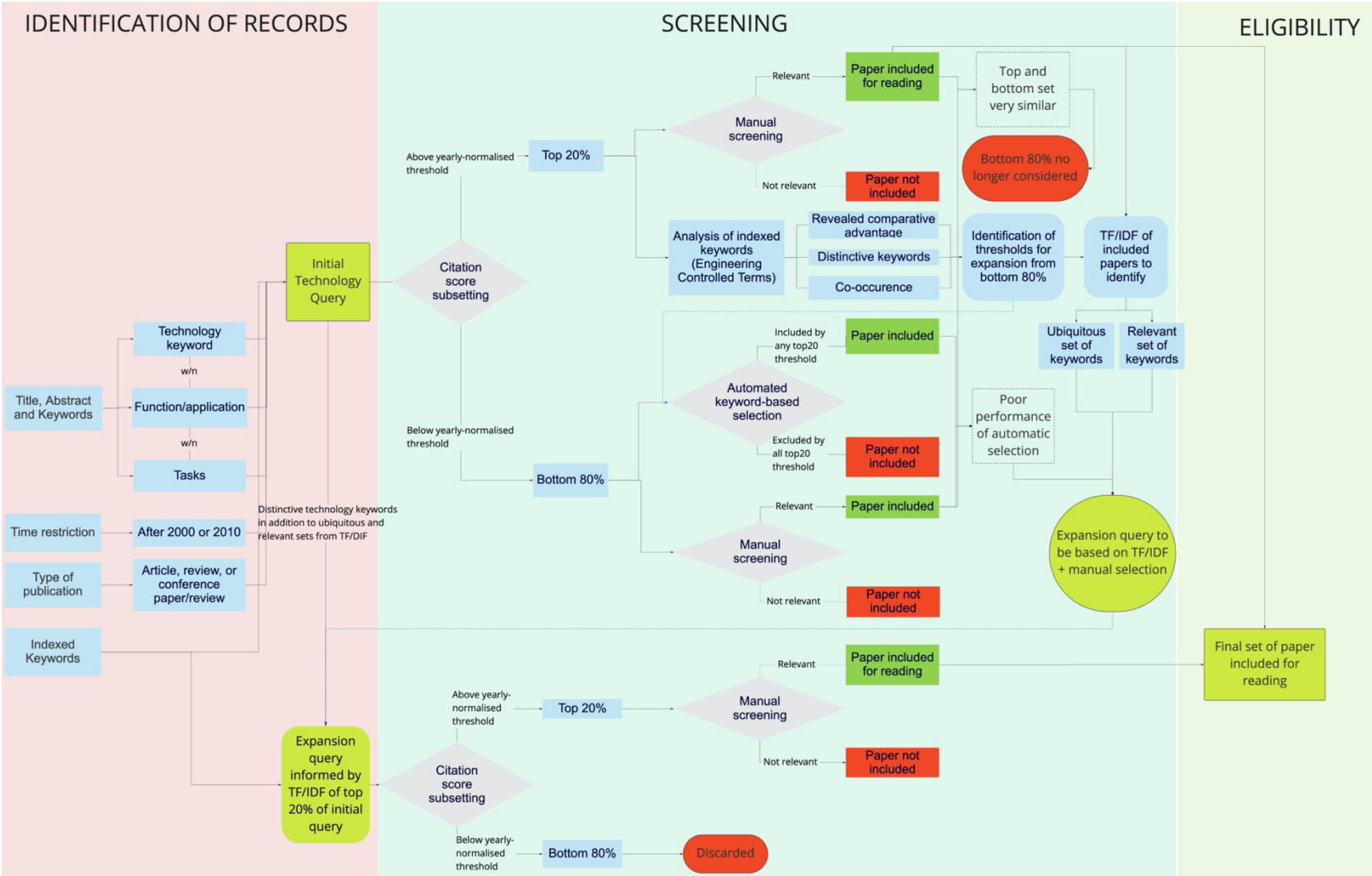
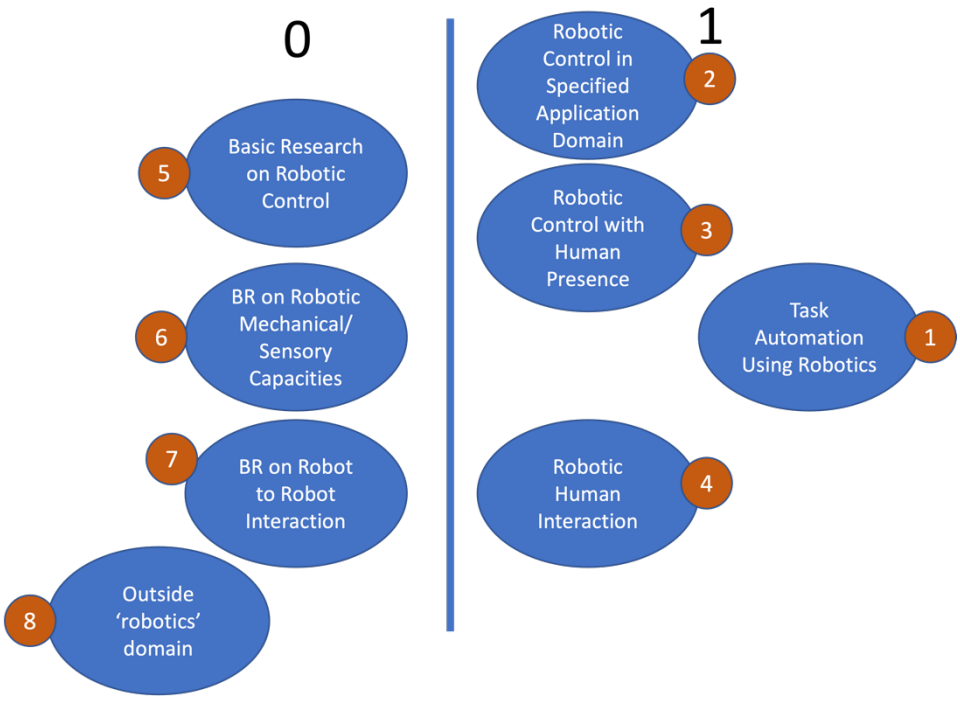


Figure 4-A illustrates the above distinction, were the section with 0 refers to document that were excluded from the selected sample and 1 refers to document that were included in the selected sample.

Figure 4-A: Task criteria for excluding/including documents

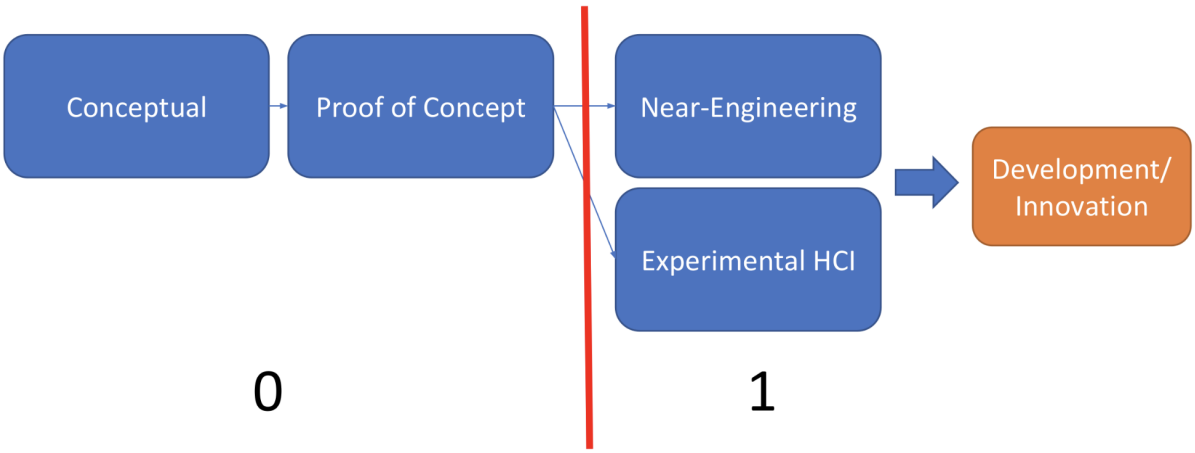


Instead, we included documents that describe the development of specific technologies, including proof of concepts that lead to such developments, documents that describe experiments of the technologies, and which refer to tasks. Figure 5-A provides a visual description of the criteria used to select screened documents based on their level of applicability. Below we define the four levels of applicability:

- Conceptual – A new framework or approach is described and argued to be valuable
- Proof-of-Concept – A particular approach is tested in a limited number of trials, sometimes with reference to existing practice but more commonly in relation to the claim made by the concept
- Near-engineering – Proof of concept is implicit but further work is done to characterise and test the robustness of the solution in ways that would be useful in the development process
- Experimental HCI – Explicit tests and co-design with users of new HCI approach. In one sense this is proof of concept but because it involves 'live subjects' it is likely to inform development efforts.

In case of disagreements between the two reviewers a third reviewer made the decision on whether to include or not the document in the selected sample. Extracted data included: title, abstract, keywords, authors, affiliation, journal, publication year, type of studies (article or research letter/comment/editorial). See Figure 6 and Figure 7 for details on the number of documents included and excluded in each step for robots and software based data management.

Figure 5-A: Application criteria for excluding/including documents



8.2.4 Step 4: Identification of records: query expansion

Any strategy to select the literature on the basis of a string of keywords may identify documents that are not relevant to the topic (low precision) and not identify documents that are relevant to the topic (low recall). Low precision occurs because the terms used are too broad. For instance, "robots" may retrieve documents from science fiction. We controlled for precisions using a relatively complex query with proximity operators, and by manually screening titles and abstracts as described in the steps above. To control for recall (i.e. retrieve documents that were not found by our query, but which are relevant for a given technology), we exploited the *engineering controlled terms*, a set of keywords in the Scopus documents metadata assigned by librarians manually. *engineering controlled terms* provide a list of subject terms for the content of a document in a specific and consistent way, because they for a curated list of terms reviewed and standardised. According to Elsevier, "controlled terms are assigned by professional indexers to capture the concepts a document deals with in a standard manner".[11]

[11] https://www.sciencedirect.com/topics/computer-science/controlled-vocabulary

Although we did not have access to the "ei Thesaurus", which is integrated in Scopus, for each document we retrieved the engineering controlled terms as part of the metadata. Only for the case of robots, we examined different ways in which we could use this controlled vocabulary to identify an automatic way to select distinctive keywords for papers in the selected sample (after screening) (Figure 3-A). And then use these distinctive controlled terms to search for more paper in Scopus, which were assigned the same key terms, but which were not retrieved by our query.

In the robot pilot test case, out of the 498 papers screened, 87 were included in the selected sample. These 87 papers have been labelled with a total number of 511 distinctive engineering controlled terms. The top ten most frequent terms are presented in Table 26-A as an example.

Table 26-A: Top 10 most frequent engineering controlled terms for robot technologies

Keywords	**n**
Robotics	29
Robots	17
Robot programming	9
Industrial robots	8
Intelligent robots	8
Man machine systems	7
Mobile robots	7
Artificial intelligence	6
Motion planning	6
Agricultural robots	5
Automation	5

To identify engineering controlled terms that are specific to the 87 documents in the selected sample, and are not relevant for documents in the residual non-selected sample (411 documents) we used four different methods:

- **Balassa index (RCA):** considering *s* a sample of documents s and k the engineering controlled terms, we define the RCA of a given engineering controlled term in a given sample of documents as the ratio between the normalized frequency of an

engineering controlled term[12] *k* in the sample *s* over the normalized total number of engineering controlled terms *K* in that set and the normalized frequency of the same engineering controlled term *k* in the whole sample of screened publications *S* over the normalized total number of engineering controlled term in the whole sample of screened publications. Formally:

$$RCA_{sk} = \frac{F_{sk}/\sum_{k\prime\in K} F_{sk\prime}}{\sum_{s\prime\in S} F_{s\prime k}/\sum_{s\prime\in S,k\prime\in K} F_{s\prime k\prime}}$$

where F_{sk} is the normalized frequency of engineering controlled term sets *k* and *k'* in the samples *s* and *s'*

- We used two samples of documents, based on the screening (Step 3): the sample selected for final analysis and the sample of all documents excluded by at both reviewers.[13]
- We consider the following set of engineering controlled terms that are specific to the selected sample: all terms in the sample of selected publications with RCA > 1.
- Note that these indicators include engineering controlled terms that may appear both in the selected sample and in the excluded sample, but which are more common to appear in the selected sample, and therefore may characterize the topic of the sample that we manually selected as relevant.

- **Co-occurrence**
  - We compute all combinations of two or more engineering controlled term that appear in any one document in the selected sample.
  - We consider the following two sets of engineering controlled terms that are specific to the selected sample: we restrict our analysis to the documents including at least 3 or 4 terms out of the overall set of engineering controlled terms in the selected sample (excluding robots and robotics because they are too common also in the non-selected sample of documents, by construction)
- **Distinctive keywords:**
  - We consider the set of engineering controlled terms that appear only in documents in the selected sample and not in documents in the non-selected sample.

[12] Because documents differ in relation to the number of key terms they are labelled with, and documents with more terms are more likely to be labelled against one term, we normalised the frequency (F) of engineering controlled terms by the number of engineering controlled term of each document. That is, if a document had 10 engineering controlled term, each term was assigned 1/10 as frequency; the same term was assigned 1/6 a document with 6 keywords.

[13] We did not include documents excluded by one reviewer, because their non-relevance is more ambiguous.

- **Single keywords selected**
    - We rank engineering controlled terms by their frequency in the selected sample of documents. We consider the set of engineering controlled terms with an absolute frequency higher or equal than 3 (top 6%).

We then assess if any, or a combination, of these four methods and sets of engineering controlled terms distinguishes relevant documents in a larger corpus. If this were true, we could then use the selected set of engineering controlled terms to search in Scopus to expand our query to other relevant documents and increase recall. To do this assessment, we first manually screened the 865 documents retrieved by our query (Step 1), not included among the 20% most cited (normalised by year) (Step 2),[14] and retrieved by at least one of the four methods explained above to select engineering controlled terms. For the screening we followed the same procedure described in Step 3.

As a result of the manual screening, 124 documents were included in the selected sample by two reviewers (14%), 592 were excluded (68%), and 149 were included by only one reviewer (17%). The controversial papers were analysed by a third reviewer, who included 76 of those in the selected sample. In total, 200 papers out of the 865 were included in this selected sample (for the 80% least cited documents) by means of manual screening.

We compared these documents allocated to the selected sample manually with the documents selected by each of the four methods (i.e. documents that contain engineering controlled terms in the sets selected by each of the methods). Results were not encouraging.

Most documents (815) were retrieved by at least one of the methods – the most inclusive being the single keyword selection (79%; 682 documents) and the most selective being co-occurrence (33% for three keywords; 287 documents). The overlap between the sample of documents selected manually, and the sample of documents selected by the four methods is rather low. Only between 32% (co-occurrence) and 78% (single keyword selected) of documents selected manually were also selected by any of the four methods. That is between 22%-68% of the documents selected manually, were not identified by any of our methods. Combining the four methods did not improve much the performance of the automatic methods to select relevant document in a larger corpus[15].

This pilot test on robots led to two important findings for our protocol. First, in relation to extending our initial query to other relevant papers in Scopus, the sparsity of engineering

[14] That is, the bottom 80% least cited documents (normalised by year).

[15] Both pair-wise combinations and a higher number of concurrent methods were analysed.

controlled terms (in the corpus of papers related to robots) is too high and their frequency too low to determine an automated method to select relevant documents.[16]

Second, in relation to the choice of focusing on the top X% of top cited documents (normalized by years) the documents in the bottom 80% cover very similar topics to those covered by documents in the top 20%. There does not seem to be a substantial added value/information in considering all documents for screening.

To extend our initial query, for all technologies, we thus followed a different strategy, which addresses the issue of the sparsity of the engineering controlled terms in the corpus of documents downloaded from Scopus. We identify relevant engineering controlled terms using Term Frequency Inverse Document Frequency (TF-IDF). TF-IDF is a text mining technique used to identify relevant words in a corpus of text. It combines the frequency of a term in the corpus (TF),[17] with their inverse document frequency (IDF), which measures the extent to which words are used frequently within a given document but not in the entire corpus of documents. By multiplying TF and IDF, we obtain the frequency of a term adjusted for how rarely it they are used. Where IDF is mathematically defined as

$$IDF(term) = ln(\frac{n_{docs}}{n_{docs\ contaning\ term}})$$

We first build the matrix of co-occurrences of all engineering controlled terms across the documents in our selected sample (in the top X%). Columns and rows are the engineering controlled terms, and the cells measure the number of documents in which each pair of engineering controlled terms appears together in the same document.

We then compute the TF-IDF using the co-occurrence matrix in the following way. We consider the weighted combination of each engineering controlled terms with all the engineering controlled terms with which it appears as the record in which a term can appear – that is, all columns of the co-occurrence matrix. In other words, we consider all the terms that appear together with each of the other engineering controlled terms in the same document. Table 27-A provides an example extracted from the case of software based data management. “State estimation” and “internet of

[16] We also made an attempt using machine learning methods (Word2vec), but the size of the corpus is too small for the algorithm to recognise any pattern in the selected sample.

[17] Measuring only how frequently a term occurs in documents would privilege terms that may not be as relevant to identify a specific corpus. For instance stop words, or in the case of the robot technology “robots”.

things” in the rows both appear together with four other terms, but "internet of things” appear more frequently with some terms than with others, which means they will have a higher IDF.

Table 27-A Extract of co-occurrence matrix of engineering controlled terms – Software-based data management

	5g mobile communication systems	internet of things	state estimation	wireless sensor networks	wireless telecommunication systems
5g mobile communication systems	0	2	1	2	1
internet of things	2	0	1	3	2
state estimation	1	1	0	1	1
wireless sensor networks	2	3	1	0	2
wireless telecommunication systems	1	2	1	2	0

Another way to look at this is to consider each engineering controlled term as a record, whose text is the sequence of all engineering controlled term with which it appears together in the same documents, repeated for the number of times they co-occur (as per the co-occurrence matrix). Table 28-A provides an example of these records for three engineering controlled terms included in Table 27-A.

Table 28-A From the co-occurence matrix to TF-IDF records and terms

Doc	**Text**
5g mobile communication systems	internet of things, internet of things, state estimation, wireless sensor networks, wireless sensor networks, wireless telecommunication systems, quality control, global positioning system, optimization, quality of service, accident prevention, automotive industry, military applications, surveys, wireless local area networks (wlan)
internet of things	5g mobile communication systems, 5g mobile communication systems, state estimation, wireless sensor networks, wireless sensor networks, wireless sensor networks, wireless telecommunication systems, wireless telecommunication systems, automation, network architecture, network architecture, network security, sensor nodes, sensor nodes, sensor nodes, zigbee, process control, semantics, diseases, data privacy, distributed database systems, fog, fog, fog, fog, fog computing, fog computing, health care, health care, middleware, middleware, global positioning system, cloud computing, cloud computing, cloud computing, digital storage, digital storage, distributed computer systems, distributed computer

	systems, manufacture, data acquisition, data acquisition, optimization, security of data, quality of service, accident prevention, accident prevention, automotive industry, automotive industry, military applications, surveys, wireless local area networks (wlan), budget control, decision making, decision making, edge computing, edge computing, electric batteries, energy efficiency, green computing, job analysis, monitoring, parallel algorithms, solar cells, agriculture, heart, data handling, intelligent buildings, internet, web services, matrix algebra, embedded systems, agricultural robots, chemical sensors, gas detectors, knowledge based systems, gas sensing electrodes, risk assessment, data transfer, safety engineering, conducting polymers, design, integration, security systems, medical applications, remote patient monitoring, planning, urban growth, complex networks, intelligent systems, blood pressure, cardiology, patient treatment, infrastructure as a service (iaas), knowledge representation, multimedia services, multimedia systems
state estimation	5g mobile communication systems, internet of things, wireless sensor networks, wireless telecommunication systems, trees (mathematics), global positioning system, kalman filters, digital storage, vehicles, online systems, uncertainty analysis, cellular telephones, location based services, telecommunication services, mobile phones, telephone sets, probes, highway traffic control, hybrid systems, lagrange multipliers, street traffic control, telecommunication equipment, traffic control

We then compute TF as the number of times a term appears in the column (record), and IDF as the log of the share between the total number of columns (records) and the number of columns (columns) in which it appears (records). In this way, we consider both the frequency of a given engineering controlled term (TF) and a measure of their relevance in combination with other terms, that is, with how many other engineering controlled terms they appear together. For example, in Table 29-A the TF of the term “human” in the record (engineering controlled term) “article” is the ratio between n (its frequency) and the total number of terms in that record (total). The IDF instead is computed as the log of the ratio between the total number of records (2060) over the total number of records in which “human” appears (256), that it the total number of engineering controlled term with which “human” co-occur.

Table 29-A: TF-IDF for extension query (example)

Record	Term	n	total	tf	idf	tf_idf
article	human	14	331	0,042296	2,043213	0,08642
human	article	14	379	0,036939	2,158191	0,079722
chemical sensors	gas detectors	13	206	0,063107	2,967022	0,187239
gas detectors	chemical sensors	13	158	0,082278	2,667617	0,219487
process control	sensors	12	256	0,046875	2,352347	0,110266

sensors	process control	12	254	0,047244	2,372966	0,112109
article	priority journal	11	331	0,033233	2,245966	0,074639
priority journal	article	11	278	0,039568	2,158191	0,085396
human	priority journal	10	379	0,026385	2,245966	0,05926
priority journal	human	10	278	0,035971	2,043213	0,073497

Note: **Record** is the name of the columns in the engineering controlled term co-occurrence matrix, that is all engineering controlled terms; **Term** is the list of engineering controlled term in the column of the engineering controlled term co-occurrence matrix (also all engineering controlled terms); **n** is the total number of times an engineering controlled term appears together with other engineering controlled terms; **total** is the number of times that all engineering controlled terms co-occur with an engineering controlled term; **tf** is term frequency as defined above; **idf** is Inverse Document Frequency as defined above; **tf-idf** is the product between tf and idf as defined above.

We use both TF and IDF to construct the expansion query, which is meant to identify additional areas of application of the technology that were not included in the string of keywords in our initial query. We build the query using two sets of engineering controlled terms. A set of ubiquitous terms, i.e., engineering controlled terms that are very frequent and appear in most of the documents in the selected sample, which we assume characterise the technology family of interest. And a set of relevant specific terms, i.e. engineering controlled terms that are specific to sub-technologies or sub-applications, which are frequent only in homogeneous subsets of the documents in the selected sample.

The first set is identified ranking engineering controlled terms by term frequency (TF); the second set is identified ranking engineering controlled terms by inverse frequency (IDF). The selection is always manual, although informed by this methodology, as the expansion is meant to cover additional areas which might be missed by the initial query, hence human judgment is always needed.

The full expansion query based on the engineering controlled terms is reported in Table 32-A. See Figure 6-A and Figure 7-A for details on the number of documents included and excluded in each step for robots and software based data management.

8.2.5 Step 5: Identification and screening of expansion documents

We repeat Steps 2 and 3 to download from Scopus the documents retrieved with the expansion query, select the top X% most cited (by year)[18], and screen them to build a second sample of selected documents to be included in the literature review.

8.2.6 Step 6: Reading and coding selected documents

We finally read the text of all documents in the selected sample. A small share of these documents were discarded from the review because considered not relevant in relation to the criteria discussed under Step 4 when reading the full text, or because we could not find in electronic form to be downloaded. See Figure 6-A and Figure 7-A for details on the number of documents included and excluded in each step for robots and software based data management. The remaining papers were coded to capture the following features, when available. These are described at length in the text and recalled here for convenience:

Emerging technologies, exposure and adoption:

1. **Level of Adoption of the technology:** Depending on whether the technology is already used or not in industry.
   Values: low, medium or high.
2. **Development stage:** Depending on the level of maturity of the technology described
   Values: conceptual; experimental; prototype; ready to deploy; mature

Task Routinisation:

3. **Routinisation.** This classifies the technology on the basis of the ability to perform a task without any human intervention, including the possibility that a task is further decomposed into an automated and humanly supervised segments
   Values: Yes; No
4. **Knowledge codification.** This classifies the technology on the basis of the ability to make all instruction explicit (i.e. codified) without the use of any tacit knowledge (so far assumed as being a human-specific characteristic)
   Values: Yes; No
5. **Works with People/Symbols/Objects.** This characterize the type of taks that can be performed in terms of what the technology works with

[18] With X depending on the number of documents retrieved, aiming to screen around 500 documents

Values: People ("in-person" services (requires high soft skills)); Things (routine production services (does not require soft skills)); Symbols ("symbolic- analytic" services (does not require soft skills))

Potential relations with employment:

6. **Skills.** Skills needed to use the technology
   Values: Low; Medium; High
7. **Substitute or complement.** Whether the technology is meant to complement or substitute the human workers. In the documents reviewed, it is also possible to identify segments of tasks (or sub-tasks) that are replaced and others within the same task that are complemented. In these cases records are identified as substitute and complements.
   Values: Complement ; Substitute
8. **Time saving or product/process innovation.** Whether the technology lead to mainly improvements on the quality of the product or service produced and/or on the production process (quantity, i.e. time saving)
   Values: Process; Product

Sectors, industry, tasks and geography categories

9. **Sector of application.** This classifies technologies according to the main intended sector of use according to the 3-digit ISIC classification for manufacturing and 2-digit for other sectors
   Values: 3-digit ISIC classification for manufacturing and 2-digit for other sectors (see Table 4)
10. **Task of application.** This classifies technologies on the basis of the tasks they are intended to perform, classified based on the ONET work activities
    Values: See Table 31-A
11. **Geographical area of provenience:** We record where the technology has been developed and mainly deployed, based on the UN M49 classification[19]
    Values: See Table 30-A
12. **Type of organizations:** This question as left open to identify the different type of organizations for which the technology is intended
13. **Size of organisations:** the typical size of the organisations for which the technology is intended

[19] This dimension is hard to codify as often the place/country of first development or application of the technology is not explicitly mentioned or deducible. When this is the case, the geographical area is attributed through the affiliation of the author. Because of this, we have not taken into consideration this category when looking at the results.

Values: Micro<10 employees; Small<50 employees; Medium<500 employees; Large>550 employees

Table 30-A: Geographical area of the technology

Geographical area
015 Northern Africa
202 Sub-Saharan Africa
419 Latin America and the Caribbean
021 Northern America
010 Antarctica
143 Central Asia
030 Eastern Asia
035 South-eastern Asia
034 Southern Asia
145 Western Asia
151 Eastern Europe
154 Northern Europe
039 Southern Europe
155 Western Europe
009 Oceania

Table 31-A: ONET work activities

ONET work activities
Getting Information
Monitor Processes, Materials, or Surroundings
Identifying Objects, Actions, and Events
Inspecting Equipment, Structures, or Material
Estimating the Quantifiable Characteristics of Products, Events, or Information
Judging the Qualities of Things, Services, or People
Processing Information
Evaluating Information to Determine Compliance with Standards
Analyzing Data or Information
Making Decisions and Solving Problems
Thinking Creatively
Updating and Using Relevant Knowledge
Developing Objectives and Strategies

Scheduling Work and Activities
Organizing, Planning, and Prioritizing Work
Performing General Physical Activities
Handling and Moving Objects
Controlling Machines and Processes
Operating Vehicles, Mechanized Devices, or Equipment
Interacting With Computers
Repairing and Maintaining Mechanical Equipment
Documenting/Recording Information
Interpreting the Meaning of Information for Others
Communicating with Supervisors, Peers, or Subordinates
Communicating with Persons Outside Organization
Establishing and Maintaining Interpersonal Relationships
Assisting and Caring for Others
Selling or Influencing Others
Resolving Conflicts and Negotiating with Others
Performing for or Working Directly with the Public
Training and Teaching Others
Guiding, Directing, and Motivating Subordinates
Coaching and Developing Others
Provide Consultation and Advice to Others
Performing Administrative Activities
Staffing Organizational Units
Monitoring and Controlling Resources

Table 32-A: Search queries by automation technology family[20]

Code	Technology family	First Query		Expansion		Notes
A	Robots	TITLE-ABS-KEY ((robot* OR "human worker") W/2 (process* OR routine OR task OR service) W/2 (automat* OR repetitive OR autonomous* OR smart OR intelligen* OR self-learn* OR interact OR recogn* OR weld OR control OR move OR clean OR walk OR carry OR detect OR drive OR predict OR detect OR identify OR determine OR control OR generate OR classify) OR cobot* OR "co-bot*" OR "collaborative robot*") AND DOCTYPE (ar OR cp OR cr OR re) AND INDEXTERMS (robot*) AND PUBYEAR > 2000		INDEXTERMS ((robotics AND robots) AND (automation OR "intelligent robots" OR "service robots" OR "mobile robots" OR "multi agent systems" OR "process control") AND ("service industry" OR crops OR "information management" OR "risk management" OR "architectural design" OR personnel OR "large scale systems" OR aircraft OR welding OR navigation OR surgery OR assembly)) AND DOCTYPE (ar OR cp OR cr OR re) AND PUBYEAR > 2000		• After 2000 • function include routine, task, process, service
B	Physical Data Acquisition Technologies	To be completed				• After 2010

[20] The queries for all the other technologies have been completed but are yet to be validated. The table is To Be Completed

C	Software-based Data management	TITLE-ABS ((("Database system*" OR "Information Management" OR "Query process*" OR "Information retrieval" OR "Search engine*" OR "Digital storage" OR "Relational Database" OR "Application Programming Interface" OR "Graph Database" OR "Cryptograph*" OR "Data Security" OR "Blockchain" OR "Data encrypt*" OR "Data privacy" OR "Network security" OR "Embedded system*" OR "Map-reduce" OR "Mapreduce") W/2 (automat* OR autonomous* OR smart OR intelligen* OR "self-learn*" OR interact OR recogn* OR clean OR detect OR predict OR identify OR generate OR classify OR acqui* OR stor* OR organi* OR access* OR retriev* OR extract* OR maintain* OR convert* OR encod* OR decod* OR encrypt* OR decrypt*) W/2 (manag* OR "problem solving" OR backoffice OR "back office" OR organis*)) OR ((data) W/1 (min* OR reduc* OR handl* OR integr* OR entr* OR enter* OR report* OR clean*) W/2 (manag* OR "problem solving" OR backoffice OR "back office" OR organis*)) OR (("big data") W/1 (analy*) W/2 (manag* OR "problem solving" OR backoffice OR "back office" OR organis*))) AND DOCTYPE (ar OR cp OR cr OR re) AND PUBYEAR > 2000 AND INDEXTERMS (data*)	**No automation, top 2%** INDEXTERMS (("data collection" OR "relational database" OR "digital storage" OR "distributed computer systems" OR "data acquisition" OR "data processing" OR "information management" OR blockchain OR "network architecture" OR "real time systems") AND ("laboratory information management system" OR "image watermarks" OR "information hiding" OR "enterprise resource planning" OR "program processors" OR "copyright protection" OR "watermarking relational databases" OR "electronic document identification systems" OR "scheduling" OR "pervasive monitoring" "interactive querying")) AND	**Automation as separate, 50%** INDEXTERMS (automation AND ("data collection" OR "relational database" OR "digital storage" OR "distributed computer systems" OR "data acquisition" OR "data processing" OR "information management" OR blockchain OR "network architecture" OR "real time systems") AND ("laboratory information management system" OR "image watermarks" OR "information hiding" OR "enterprise resource planning" OR "program processors" OR "copyright protection" OR "watermarking relational databases" OR "electronic document identification systems" OR "scheduling" OR "pervasive monitoring" OR "interactive querying")) AND	• After 2010

				DOCTYPE (ar OR cp OR cr OR re) AND PUBYEAR > 2010	DOCTYPE (ar OR cp OR cr OR re) AND PUBYEAR > 2010	
D	Computing	To be completed				
E	AI (not directly as a cloud service) & Intelligent Information System	To be completed				
F	Additive manufacturing	To be completed				
G	Networking	To be completed				
H	User interface	To be completed				
I	Other	To be completed				

Figure 6-A: Selection of documents for literature review: robots

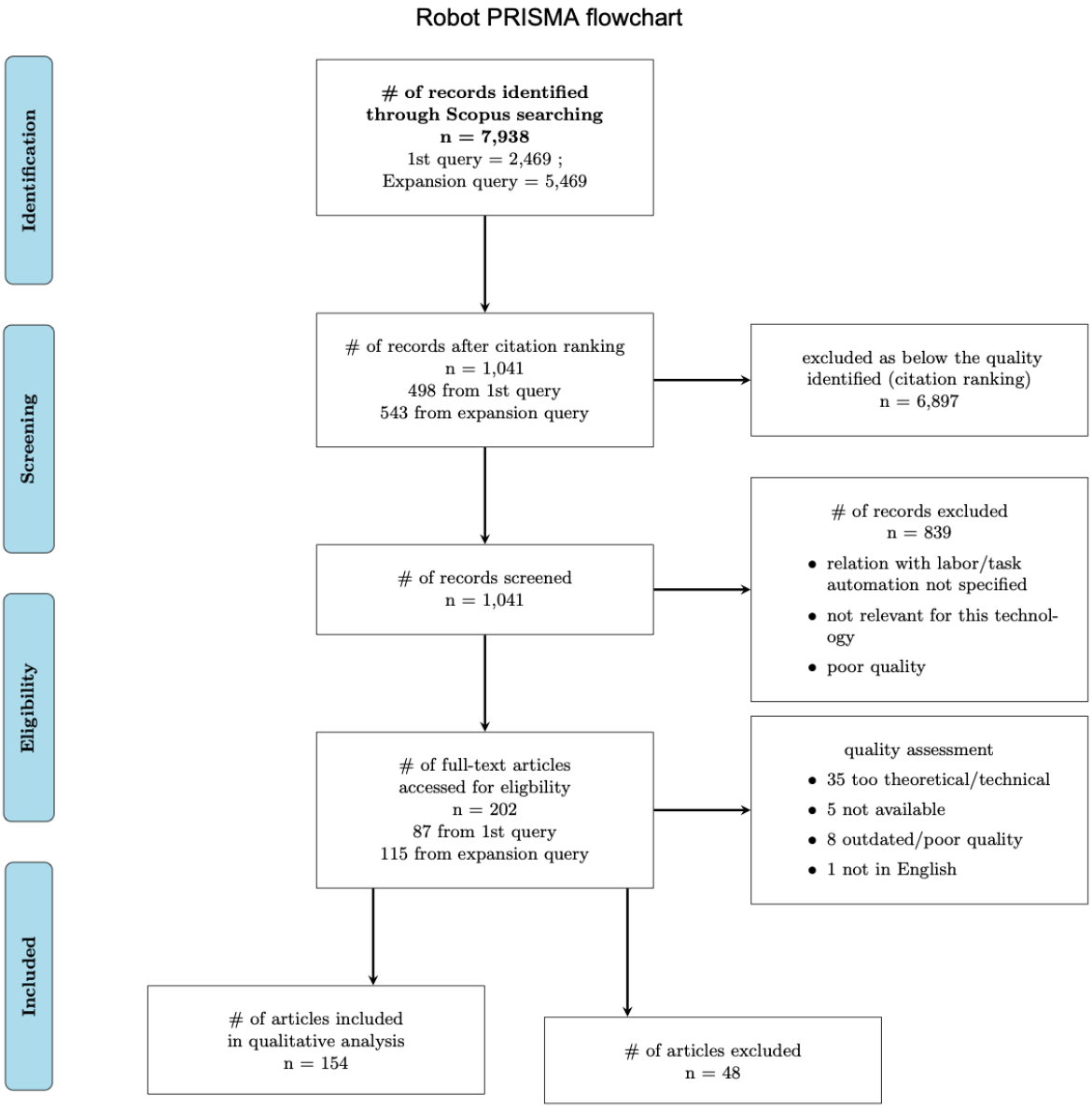


Figure 7-A: Selection of documents for literature review: software based data management